COMPARING READABILITY OF GRAPHICAL AND SENTENTIAL PROCESS DESIGN NOTATIONS - DATA ANALYSIS REPORT

TECHNICAL REPORT 658

AVNER OTTENSOOSER AND ALAN FEKETE

AUGUST, 2010
1. Goal of this technical report

The goal of this document is to create an audit for the calculation of the test results in a manner that will enable all interested stakeholders to review and tune the calculations. Furthermore, using the Raw data CSV file, this document, a C# programme, the Stata nalitical platform and the free R environment (available from http://www.r-project.org/) all interested stakeholders will be able to reproduce the results in part or in full. The results will then be included in a forthcoming paper.

Contents

1. Goal of this technical report .......................................................................................................................... 2

Contents .......................................................................................................................................................... 2

2. Preparation of the data .................................................................................................................................. 5

2.1. Transpose the input file ............................................................................................................................... 5

2.2. Mark the questioners ................................................................................................................................... 5

2.3. Load the transposed data into R .................................................................................................................. 8

3. Work published in the paper .......................................................................................................................... 9

3.1. Table 1: Results from the three universities ................................................................................................. 9

3.2. Table 2: P Values for one-sided Wilcoxon sign-rank tests .......................................................................... 13

3.3. Section 5.4 - Predictions .......................................................................................................................... 22

3.4. Statistical tests for normality ....................................................................................................................... 37

4. Work not published in the paper .................................................................................................................... 39

4.1. Comparing comprehension of the first questioner and Placebo ................................................................. 39

4.2. Comparing comprehension of Cockburn and BPMN ................................................................................. 65

4.3. Comparing the contribution of the 2nd design ............................................................................................ 74

Appendix – Participant work book .................................................................................................................. 102
Tables

Table 1 – The right answers to the questioners ................................................................. 5
Table 2 - TU/e Results - First notation BPMN ..................................................................10
Table 3 - TU/e Results - First notation Cockburn ...........................................................10
Table 4 - USYD Results - First notation BPMN .............................................................11
Table 5 - USYD Results - First notation Cockburn ........................................................11
Table 6 - HU/b - First notation BPMN Results ..............................................................12
Table 7 - HU/b - First notation Cockburn Results ........................................................12
Table 8 - Statistical tests for normality ..............................................................................37
Table 9 – H₀ – USYD Students don’t comprehend BPMN better than placebo – ☑ Not rejected
........................................................................................................................................40
Table 10 – H₁ – USYD student comprehend Cockburn better than placebo – ☑ Accepted (P =
0.005432) ..........................................................................................................................45
Table 11 – H₁ – TU/e Students comprehend BPMN better than placebo – ☑ Accepted (P = 1.148e-09)
............................................................................................................................................50
> Table 12 – H₁ – TU/e Students comprehend Cockburn better than placebo - ☑ Accepted (P =
0.0001207) ........................................................................................................................55
Table 13 – H₁ – HU/b Students comprehend BPMN better than placebo - ☑ Accepted (p = 0.01477)
............................................................................................................................................60
Table 14 - H₁ – HU/b Students comprehend Cockburn better than placebo – ☑ Accepted (P = 0.0209)
............................................................................................................................................64
Table 15 – H₀ – USYD Students don’t comprehend Cockburn Better than BPMN ☑ Not rejected
............................................................................................................................................67
Table 16 – H₀ – TU/e Students don’t comprehend BPMN Better than Cockburn ☑ Not rejected
............................................................................................................................................70
Table 17 – H₀ – HU/b Students don’t comprehend BPMN Better than Cockburn ☑ Not rejected
............................................................................................................................................73
Table 18 – H₁ – Presenting BPMN to USYD students who already read Cockburn improves
comprehension – ☑ Accepted (P = 0.0005168) .....................................................................78
Table 19 – H₀ – Presenting Cockburn to USYD students who already read BPMN doesn’t improve
comprehension – ☑ Not rejected ..........................................................................................80
Table 20 – H₁ – Presenting BPMN to TU/e students who already read Cockburn improves comprehension – ✅ Accepted (P = 0.01404) ............................................................................................... 85

Table 21 – H₀ – Presenting Cockburn to TU/e students who already read BPMN doesn’t reduce comprehension – ❌ Not rejected ............................................................................................... 87

Table 22 – H₀ – Presenting BPMN to HU/b students who already read Cockburn doesn’t improve comprehension – ❌ Not rejected ............................................................................................... 89

Table 23 – H₀ – Presenting Cockburn to HU/b students who already read BPMN doesn’t reduce comprehension – ❌ Not rejected ............................................................................................... 91
2. Preparation of the data

2.1. Transpose the input file
The raw data file is in CSV format, it includes column header and row header. Every row in the file contains all participants’ answers to a single question (the record structure is not normalised). The file was generated in a format easy for data capture and should be inverted for analyse purpose so that each record will contain all the answers of one participant.

Method of conversion: Excel→ Pastespecial → Transpose → Participants questioneers.CSV

2.2. Mark the questioners

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
<th>Rational</th>
</tr>
</thead>
<tbody>
<tr>
<td>QSet1,7,13</td>
<td>6</td>
<td>See use case 3.2 1a1</td>
</tr>
<tr>
<td>QSet2,8,14</td>
<td>2</td>
<td>See use case 1 items 2 and 4</td>
</tr>
<tr>
<td>Q3,9,15</td>
<td>1</td>
<td>See use case 2 item 2</td>
</tr>
<tr>
<td>Q4,10,16</td>
<td>2</td>
<td>Activity 6 at use case 3.1 is entered into as long as bugs are found, the bugs a severe enough to be x but do not that cancellation of the project.severe that thy trigger the</td>
</tr>
<tr>
<td>Q5,11,17</td>
<td>1</td>
<td>See use case 3.1 activity 6</td>
</tr>
<tr>
<td>Q6,12,18</td>
<td>1</td>
<td>While items 1-3 may trigger a build, other activities can trigger a build such as the initial build</td>
</tr>
</tbody>
</table>

Replace the answers to questions Q1 – Q18 with 0s and 1s where ‘0’ represents a wrong answer and ‘1’ represent a right answer using table #1 as a reference:

Method of conversion: a C# programme that accepts

Total the answers for each questioner for each participant and add these new fields to the table –

Score Placebo, Score QSet1, Score QSet2, Delta QSet1 - Placebo, Delta QSet2 - QSet1, Delta QSet2 - Placebo
2.2.1. Method of operation – Use a C# programme.

```csharp
using System;
using System.IO;
namespace LabScore
{
    class prgLabScore
    {
        static void Main(string[] args)
        {
            char[] AO_delimiter = (','); // Delimeter for the CSV parser
            string AO_inputLine; // one participant's set of answers
            string[] AO_Fileds; // Parsed one participant's set of answers
            string AO_outputLine; // one participant's answers and scores
            int ScorePlacebo, ScoreQSet1, ScoreQSet2, DeltaQSet1Placebo, DeltaQSet2QSet1, DeltaQSet2Placebo; // the
            varnious scores
            string[] AO_RightAnswers = {"6", "2", "1", "2", "1", "1"}; // The right answers
            File AO_Reader = new StreamReader("C:\Users\Avner\Documents\My Dropbox\Lab\Trasposed.csv");
            string[] AO_Writer = new StreamWriter("C:\Users\Avner\Documents\My Dropbox\Lab\Scored.csv");
            AO_inputLine = AO_Reader.ReadLine(); // read the header line of text
            AO_outputLine = AO_inputLine + ",Score Placebo,Score QSet1,Score QSet2,Delta QSet1 Placebo,Delta QSet2 QSet1,Delta QSet2 Placebo";
            AO_Writer.WriteLine(AO_outputLine); // write the header to the new file
            while ((AO_inputLine = AO_Reader.ReadLine()) != null) // read all the participants' records until the file's end
            {
                // parse the input line
                AO_Fileds = AO_inputLine.Split(AO_delimiter);
                ScorePlacebo = 0;
                ScoreQSet1 = 0;
            }
        }
    }
}
```
ScoreQSet2 = 0;
// mark all the 3 instances of each question and go to the next question
for (int i = 0; i < 6; i++)
{
    // compare the i's question in the 1st questioner with the i's right answer
    // Six cells - Cell#1 = QSet1 ,Cell#2 = QSet2 ,Cell#3 = Q3 ,Cell#4 = Q4 ,Cell#5 = Q5 ,Cell#6 = Q6
    if (AO_Fileds[i + 1] == AO_RightAnswers[i])
    {
        ScorePlacebo++;
    }
    // the 8th cell is used to record the first test (Cockburn or BPMN) so we jump to 9
    // compare the i's question in the 2nd questioner with the i's right answer
    // Six cells - Cell#8 = Q7 ,Cell#9 = Q8 ,Cell#10 = Q9 ,Cell#11 = QSet10 ,Cell#12 = QSet11
    ,Cell#13 = QSet12
    if (AO_Fileds[i + 8] == AO_RightAnswers[i])
    {
        ScoreQSet1++;
    }
    // compare the i's question in the 3rd questioner with the i's right answer
    // Six cells - Cell#14 = QSet13 ,Cell#15 = QSet14 ,Cell#16 = QSet15 ,Cell#17 = QSet16 ,Cell#18 = QSet17 ,Cell#19 = QSet18
    if (AO_Fileds[i + 14] == AO_RightAnswers[i])
    {
        ScoreQSet2++;
    }
}
DeltaQSet1Placebo = ScoreQSet1 - ScorePlacebo;
DeltaQSet2QSet1 = ScoreQSet2 - ScoreQSet1;
DeltaQSet2Placebo = ScoreQSet2 - ScorePlacebo;
// build the participant's scored line
AO_score = AO_inputLine + ",";
AO_score = AO_score + ScorePlacebo.ToString() + ",";
AO_score = AO_score + ScoreQSet1.ToString() + ",";
AO_score = AO_score + ScoreQSet2.ToString() + ",";
AO_score = AO_score + DeltaQSet1Placebo.ToString() + ",";
AO_score = AO_score + DeltaQSet2QSet1.ToString() + ",";
AO_score = AO_score + DeltaQSet2Placebo.ToString();

AO_Writer.WriteLine(AO_score); // write the participant's record to the new file

AO_Reader.Close(); // close the input file
AO_Writer.Close(); // close the scored file

2.3. **Load the transposed data into R**

Method of operation – Use R command:

```r
AO_lab<-read.table("C:\\Users\\Avner\\Documents\\My Dropbox\\Lab\\Scored.csv", header=T, row.names = 1, sep=",")
```
3. Work published in the paper

3.1. Table 1: Results from the three universities.

<table>
<thead>
<tr>
<th>University</th>
<th>First artifact</th>
<th>Score</th>
<th>( \bar{x} )</th>
<th>s</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>TU/e</td>
<td>BPMN</td>
<td>Placebo</td>
<td>1.96</td>
<td>1.09</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QSet1</td>
<td>3.41</td>
<td>1.32</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>QSet2</td>
<td>3.23</td>
<td>1.37</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UC</td>
<td>Placebo</td>
<td>1.78</td>
<td>1.10</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QSet1</td>
<td>2.73</td>
<td>1.38</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>QSet2</td>
<td>3.16</td>
<td>1.36</td>
<td></td>
</tr>
<tr>
<td>USYD</td>
<td>BPMN</td>
<td>Placebo</td>
<td>2.05</td>
<td>1.22</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QSet1</td>
<td>2.47</td>
<td>1.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>QSet2</td>
<td>2.53</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UC</td>
<td>Placebo</td>
<td>1.86</td>
<td>0.77</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QSet1</td>
<td>2.68</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>QSet2</td>
<td>3.55</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>HU</td>
<td>BPMN</td>
<td>Placebo</td>
<td>1.77</td>
<td>0.73</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QSet1</td>
<td>2.85</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>QSet2</td>
<td>2.54</td>
<td>1.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UC</td>
<td>Placebo</td>
<td>1.62</td>
<td>1.12</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QSet1</td>
<td>2.77</td>
<td>1.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>QSet2</td>
<td>2.85</td>
<td>1.77</td>
<td></td>
</tr>
</tbody>
</table>

3.1.1. R Function:

```r
AO_fStat <- function(AO_vSet)
{
  AO_fDescriptive <- c(ConfFrom = mean(AO_vSet) - 1.96*sd(AO_vSet)/sqrt(length(AO_vSet)),
                       Mean = mean(AO_vSet),
                       ConfTo = mean(AO_vSet) + 1.96*sd(AO_vSet)/sqrt(length(AO_vSet)),
                       sd(AO_vSet),
                       Count = length(AO_vSet))
  (AO_fDescriptive)
}
```

Figure 1 - The data from that table is justified in the following sections
Table 2 - TU/e Results - First notation BPMN

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>sd</th>
<th>Count</th>
<th>R Call</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Score.Placebo</td>
<td>1.7108980</td>
<td>1.9594590</td>
<td>2.2080210</td>
<td>1.0909240</td>
<td>74</td>
<td>1</td>
</tr>
<tr>
<td>B Score.QSet1</td>
<td>3.1038910</td>
<td>3.4054050</td>
<td>3.7069200</td>
<td>1.3233300</td>
<td>74</td>
<td>2</td>
</tr>
<tr>
<td>C Score.QSet2</td>
<td>2.9304030</td>
<td>3.2432430</td>
<td>3.5560830</td>
<td>1.3730360</td>
<td>74</td>
<td>3</td>
</tr>
<tr>
<td>D Delta.QSet1.Placebo</td>
<td>1.1041430</td>
<td>1.4459460</td>
<td>1.7877490</td>
<td>1.5001540</td>
<td>74</td>
<td>4</td>
</tr>
<tr>
<td>E Delta.QSet2.q1</td>
<td>-0.4341868</td>
<td>-0.1621622</td>
<td>0.1098625</td>
<td>1.1939003</td>
<td>74</td>
<td>5</td>
</tr>
<tr>
<td>F Delta.QSet2.Placebo</td>
<td>0.9411942</td>
<td>1.2837838</td>
<td>1.6236733</td>
<td>1.5036054</td>
<td>74</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 3 - TU/e Results - First notation Cockburn

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>sd</th>
<th>Count</th>
<th>R Call</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Score.Placebo</td>
<td>1.4910100</td>
<td>1.7818180</td>
<td>2.0726260</td>
<td>1.10352</td>
<td>55</td>
<td>8</td>
</tr>
<tr>
<td>B Score.QSet1</td>
<td>2.3624310</td>
<td>2.7272730</td>
<td>3.0921150</td>
<td>1.380480</td>
<td>55</td>
<td>9</td>
</tr>
<tr>
<td>C Score.QSet2</td>
<td>2.8048390</td>
<td>3.1636360</td>
<td>3.5224330</td>
<td>1.357607</td>
<td>55</td>
<td>10</td>
</tr>
<tr>
<td>D Delta.QSet1.Placebo</td>
<td>0.5126252</td>
<td>0.9454545</td>
<td>1.3782838</td>
<td>1.6377286</td>
<td>55</td>
<td>11</td>
</tr>
<tr>
<td>E Delta.QSet2.q1</td>
<td>0.1037576</td>
<td>0.4363636</td>
<td>0.7689697</td>
<td>1.2585064</td>
<td>55</td>
<td>12</td>
</tr>
<tr>
<td>F Delta.QSet2.Placebo</td>
<td>0.9608887</td>
<td>1.3818182</td>
<td>1.8027476</td>
<td>1.5927022</td>
<td>55</td>
<td>13</td>
</tr>
</tbody>
</table>

1 AO_fStat(subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.Placebo)
2 AO_fStat(subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.QSet1)
3 AO_fStat(subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.QSet2)
4 AO_fStat(subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Delta.QSet1.Placebo)
5 Here we see negative progress – the results after reading BPM and Cockburn are worth than the results after only reading BPM!
6 AO_fStat(subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$ Delta.QSet2.q1)
7 AO_fStat(subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$ Delta.QSet2.Placebo)
8 AO_fStat(subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$Score.Placebo)
9 AO_fStat(subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$Score.QSet1)
10 AO_fStat(subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$Score.QSet2)
11 AO_fStat(subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$ Delta.QSet1.Placebo)
12 AO_fStat(subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$ Delta.QSet2.q1)
13 AO_fStat(subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$ Delta.QSet2.Placebo)
Table 4 - USYD Results - First notation BPMN

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>Sd</th>
<th>Count</th>
<th>R Call</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Score.Placebo</td>
<td>1.502456</td>
<td>2.052632</td>
<td>2.602807</td>
<td>1.223551</td>
<td>19</td>
<td>14</td>
</tr>
<tr>
<td>B Score.QSet1</td>
<td>1.849032</td>
<td>2.473684</td>
<td>3.098336</td>
<td>1.389181</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>C Score.QSet2</td>
<td>2.043686</td>
<td>2.526316</td>
<td>3.008946</td>
<td>1.073340</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td>D Delta.QSet1.Placebo</td>
<td>0.34784090</td>
<td>0.42105260</td>
<td>1.1899462</td>
<td>1.70996390</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>E Delta.QSet2.q1</td>
<td>0.57391052</td>
<td>0.05263158</td>
<td>0.67917368</td>
<td>1.39338454</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>F Delta.QSet2.Placebo</td>
<td>0.21917380</td>
<td>0.47368420</td>
<td>1.16654220</td>
<td>1.54086630</td>
<td>19</td>
<td>20</td>
</tr>
</tbody>
</table>

Here we see very little progress. BPMN score is similar to placebo score.

Table 5 - USYD Results - First notation Cockburn

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>Sd</th>
<th>Count</th>
<th>R Call</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Score.Placebo</td>
<td>1.5400698</td>
<td>1.8636364</td>
<td>2.1872029</td>
<td>0.7743172</td>
<td>22</td>
<td>21</td>
</tr>
<tr>
<td>B Score.QSet1</td>
<td>2.2279580</td>
<td>2.6818180</td>
<td>3.1356780</td>
<td>1.0861190</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>C Score.QSet2</td>
<td>3.1869760</td>
<td>3.5454550</td>
<td>3.9039330</td>
<td>0.8578640</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>D Delta.QSet1.Placebo</td>
<td>0.2921896</td>
<td>0.818181</td>
<td>1.3441740</td>
<td>1.2587357</td>
<td>22</td>
<td>25</td>
</tr>
<tr>
<td>E Delta.QSet2.q1</td>
<td>0.5153181</td>
<td>0.8636364</td>
<td>1.2119546</td>
<td>0.8335498</td>
<td>22</td>
<td>26</td>
</tr>
<tr>
<td>F Delta.QSet2.Placebo</td>
<td>1.2279580</td>
<td>1.6818180</td>
<td>2.1356780</td>
<td>1.0861190</td>
<td>22</td>
<td>28</td>
</tr>
</tbody>
</table>

Here we see more progress. Cockburn score is better than placebo score.

Here we see the highest progress. Cockburn first and then BPMN.
### Table 6 - HU/b - First notation BPMN Results

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>sd</th>
<th>Count</th>
<th>R Call</th>
</tr>
</thead>
<tbody>
<tr>
<td>A  Score.Placebo</td>
<td>1.37511000</td>
<td>1.76923100</td>
<td>2.16335100</td>
<td>0.72501100</td>
<td>13</td>
<td>29</td>
</tr>
<tr>
<td>B  Score.QSet1</td>
<td>2.26548000</td>
<td>2.84615400</td>
<td>3.42682700</td>
<td>1.06818800</td>
<td>13</td>
<td>30</td>
</tr>
<tr>
<td>C  Score.QSet2</td>
<td>1.85029700</td>
<td>2.53846200</td>
<td>3.22662600</td>
<td>1.26592400</td>
<td>13</td>
<td>31</td>
</tr>
<tr>
<td>D  Delta.QSet1.Placebo</td>
<td>0.29350320</td>
<td>1.07692310</td>
<td>1.86034300</td>
<td>1.44115340</td>
<td>13</td>
<td>32</td>
</tr>
<tr>
<td>E  Delta.QSet2.q1</td>
<td>-0.30769230</td>
<td>0.15700970</td>
<td>0.85485040</td>
<td>13</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>F  Delta.QSet2.Placebo</td>
<td>0.03565667</td>
<td>1.57411821</td>
<td>1.48064435</td>
<td>13</td>
<td>35</td>
<td></td>
</tr>
</tbody>
</table>

### Table 7 - HU/b - First notation Cockburn Results

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>sd</th>
<th>Count</th>
<th>R Call</th>
</tr>
</thead>
<tbody>
<tr>
<td>A  Score.Placebo</td>
<td>1.00605800</td>
<td>1.61538500</td>
<td>2.22471100</td>
<td>1.12089700</td>
<td>13</td>
<td>36</td>
</tr>
<tr>
<td>B  Score.QSet1</td>
<td>1.93430900</td>
<td>2.76923100</td>
<td>3.60415300</td>
<td>1.53589500</td>
<td>13</td>
<td>37</td>
</tr>
<tr>
<td>C  Score.QSet2</td>
<td>1.88272400</td>
<td>2.84615400</td>
<td>3.80958400</td>
<td>1.77229400</td>
<td>13</td>
<td>38</td>
</tr>
<tr>
<td>D  Delta.QSet1.Placebo</td>
<td>0.19041620</td>
<td>1.15384620</td>
<td>2.11727610</td>
<td>1.77229390</td>
<td>13</td>
<td>39</td>
</tr>
<tr>
<td>E  Delta.QSet2.q1</td>
<td>-0.07692308</td>
<td>0.86034298</td>
<td>1.44115338</td>
<td>13</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>F  Delta.QSet2.Placebo</td>
<td>0.07432566</td>
<td>1.23076923</td>
<td>2.38721280</td>
<td>2.12735541</td>
<td>13</td>
<td>42</td>
</tr>
</tbody>
</table>

29 AO_fStat(subset(AO_lab,Source=="HU/b" &First.Questionnaire=="BPMN")$Score.Placebo)
30 AO_fStat(subset(AO_lab,Source=="HU/b" &First.Questionnaire=="BPMN")$Score.QSet1)
31 AO_fStat(subset(AO_lab,Source=="HU/b" &First.Questionnaire=="BPMN")$Score.QSet2)
32 AO_fStat(subset(AO_lab,Source=="HU/b" &First.Questionnaire=="BPMN")$ Delta.QSet1.Placebo)
33 Here we see that administering Cockburn second makes things worst even in Berlin
34 AO_fStat(subset(AO_lab,Source=="HU/b" &First.Questionnaire=="BPMN")$ Delta.QSet2.q1)
35 AO_fStat(subset(AO_lab,Source=="HU/b" &First.Questionnaire=="BPMN")$ Delta.QSet2.Placebo)
36 AO_fStat(subset(AO_lab,Source=="HU/b" &First.Questionnaire=="Cockburn")$Score.Placebo)
37 AO_fStat(subset(AO_lab,Source=="HU/b" &First.Questionnaire=="Cockburn")$Score.QSet1)
38 AO_fStat(subset(AO_lab,Source=="HU/b" &First.Questionnaire=="Cockburn")$Score.QSet2)
39 AO_fStat(subset(AO_lab,Source=="HU/b" &First.Questionnaire=="Cockburn")$ Delta.QSet1.Placebo)
40 Here we see that the BPMN second did not help much.
41 AO_fStat(subset(AO_lab,Source=="HU/b" &First.Questionnaire=="Cockburn")$ Delta.QSet2.q1)
42 AO_fStat(subset(AO_lab,Source=="HU/b" &First.Questionnaire=="Cockburn")$ Delta.QSet2.Placebo)
3.2. Table 2: P Values for one-sided Wilcoxon sign-rank tests

<table>
<thead>
<tr>
<th></th>
<th>BAs</th>
<th>SMEs</th>
<th>Test</th>
<th>Data set 1</th>
<th>Data set 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_1$</td>
<td>0.0000</td>
<td>0.0654</td>
<td>Within Subjects</td>
<td>Primary Contribution (UC)</td>
<td>0</td>
</tr>
<tr>
<td>$H_2$</td>
<td>0.0000</td>
<td>0.1540</td>
<td>Within Subjects</td>
<td>Primary Contribution (BPMN)</td>
<td>0</td>
</tr>
<tr>
<td>$H_3$</td>
<td>0.9462</td>
<td>0.2352</td>
<td>Between Subjects</td>
<td>Primary Contribution (UC)</td>
<td>0</td>
</tr>
<tr>
<td>$H_4$</td>
<td>0.0642</td>
<td>0.7729</td>
<td>Between Subjects</td>
<td>Primary Contribution (BPMN)</td>
<td>0</td>
</tr>
<tr>
<td>$H_5$</td>
<td>0.0066</td>
<td>0.0003</td>
<td>Within Subjects</td>
<td>Secondary Contribution (BPMN)</td>
<td>0</td>
</tr>
<tr>
<td>$H_6$</td>
<td>0.9114</td>
<td>0.4521</td>
<td>Within Subjects</td>
<td>Secondary Contribution (UC)</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 2 - Calculations of this data follows

3.2.1. R Code

```r
AO_fwilcoxTestP <- function(AO_v1, AO_v2) {
  wilcox.test(AO_v1, AO_v2, alternative="less", paired=TRUE, exact = TRUE)
}

# H1
AO_vec1 <- subset(AO_lab, (Source="HU/b" | Source="TU/e") & First.Questionnaire=="Cockburn")$Score.Placebo
AO_vec2 <- subset(AO_lab, (Source="HU/b" | Source="TU/e") & First.Questionnaire=="Cockburn")$Score.QSet1
AO_fwilcoxTestP(AO_vec1, AO_vec2)

# H2
AO_vec3 <- subset(AO_lab, (Source="USYD") & First.Questionnaire=="Cockburn")$Score.Placebo
AO_vec4 <- subset(AO_lab, (Source="USYD") & First.Questionnaire=="Cockburn")$Score.QSet1
AO_fwilcoxTestP(AO_vec3, AO_vec4)

# H3
AO_vec5 <- subset(AO_lab, (Source="HU/b" | Source="TU/e") & First.Questionnaire=="BPMN")$Score.Placebo
AO_vec6 <- subset(AO_lab, (Source="HU/b" | Source="TU/e") & First.Questionnaire=="BPMN")$Score.QSet1
AO_vec7 <- subset(AO_lab, (Source="USYD") & First.Questionnaire=="BPMN")$Score.Placebo
AO_vec8 <- subset(AO_lab, (Source="USYD") & First.Questionnaire=="BPMN")$Score.QSet1
AO_fwilcoxTestP(AO_vec5, AO_vec6)
AO_vec9 <- subset(AO_lab, (Source="HU/b" | Source="TU/e") & First.Questionnaire=="BPMN")$Delta.QSet1.Placebo
AO_vec10 <- subset(AO_lab, (Source="HU/b" | Source="TU/e") & First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo
AO_fwilcoxTestP(AO_vec9, AO_vec10)
```
AO_fwilcoxTestP (AO_vec9, AO_vec10)
AO_vec11 <- subset(AO_lab, (Source=="USYD") & First.Questionnaire=="BPMN")$Delta.QSet1.Placebo
AO_vec11
AO_vec12 <- subset(AO_lab, (Source=="USYD") & First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo
AO_vec12
AO_fwilcoxTestP (AO_vec11, AO_vec12)

#H4
AO_vec13 <- subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo
AO_vec13
AO_vec14 <- subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")$Delta.QSet1.Placebo
AO_vec14
AO_vec15 <- subset(AO_lab, (Source=="USYD") & First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo
AO_vec15
AO_vec16 <- subset(AO_lab, (Source=="USYD") & First.Questionnaire=="BPMN")$Delta.QSet1.Placebo
AO_vec16
AO_fwilcoxTestP (AO_vec15, AO_vec16)
AO_fwilcoxTestP <- function(AO_v1, AO_v2){ wilcox.test
(AO_v1, AO_v2, alternative="less", paired=TRUE, exact = TRUE) }

#H5
AO_fwilcoxTestP <- function(AO_v1, AO_v2){ wilcox.test
(AO_v1, AO_v2, alternative="less", paired=TRUE, exact = TRUE) }
AO_vec17 <- subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo
AO_vec17
AO_vec18 <- subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")$Delta.QSet2.Placebo
AO_vec18
AO_fwilcoxTestP (AO_vec17, AO_vec18)
AO_vec19 <- subset(AO_lab, (Source=="USYD") & First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo
AO_vec19
AO_vec20 <- subset(AO_lab, (Source=="USYD") & First.Questionnaire=="BPMN")$Delta.QSet2.Placebo
AO_vec20
AO_fwilcoxTestP (AO_vec19, AO_vec20)

#H6
AO_vec21 <- subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")$Delta.QSet1.Placebo
AO_vec21
AO_vec22 <- subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")$Delta.QSet2.Placebo
AO_vec22
AO_fwilcoxTestP (AO_vec21, AO_vec22)
AO_vec23 <- subset(AO_lab, (Source=="USYD") & First.Questionnaire=="BPMN")$Delta.QSet1.Placebo
AO_vec23
AO_vec24<- subset(AO_lab, (Source=="USYD") & 
First.Questionnaire=="BPMN")$Delta.QSet2.Placebo 
AO_vec24 
AO_fwilcoxTestP (AO_vec23, AO_vec24)
### 3.2.2. R Listing

```r
> AO_fwilcoxTestP <- function(AO_v1, AO_v2) { wilcox.test(AO_v1, AO_v2, alternative="less", paired=TRUE, exact = TRUE) }
> # H1
> AO_vec1 <- subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="Cockburn")$Score.Placebo
> AO_vec1
[1] 1 2 4 3 2 3 2 4 1 2 2 1 0 2 3 1 1 3 0 3 1 0 1 4 2 1 0 4 2 2 2 1 1 1 1
1 0 4
[39] 1 1 1 2 1 1 2 3 2 2 2 3 3 2 2 1 2 3 0 1 1 1 1 1 3 4 2
> AO_vec2 <- subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="Cockburn")$Score.QSet1
> AO_vec2
[1] 2 3 4 3 1 2 2 5 2 5 4 3 4 6 2 2 3 1 0 1 2 2 1 4 3 2 3 2 2 1 1 4 4 2 0 0 4
[39] 5 4 3 4 3 3 0 4 3 2 3 3 4 4 4 4 4 2 3 0 6 3 2 2 3 4 1 2
> AO_fwilcoxTestP(AO_vec1, AO_vec2)

Wilcoxon signed rank test with continuity correction
data:  AO_v1 and AO_v2
V = 245, p-value = 1.182e-05
alternative hypothesis: true location shift is less than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "less", paired = TRUE,  :
   cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "less", paired = TRUE,  :
   cannot compute exact p-value with zeroes
> AO_vec3 <- subset(AO_lab, (Source=="USYD") & First.Questionnaire=="Cockburn")$Score.Placebo
> [1] 1 2 1 1 1 2 3 3 2 3 2 2 2 3 1 2 2 1 1 1 3 2
> AO_vec4 <- subset(AO_lab, (Source=="USYD") & First.Questionnaire=="Cockburn")$Score.QSet1
> AO_vec4
[1] 3 3 3 3 3 1 2 4 5 2 2 2 4 5 2 3 2 2 2 2 1 3
> # H2
> AO_vec5 <- subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")$Score.Placebo
> AO_vec5
[1] 3 2 1 3 1 0 3 2 3 1 2 3 2 3 2 3 1 4 4 1 1 1 2 0 2 3 3 4 2 1 2 2 0 2 1
[39] 4 2 0 1 3 2 3 2 3 2 2 3 2 1 2 2 1 1 0 0 4 3 2 2 2 0 1 3 2 1 3 4 2 2 1 1 2 1
[77] 1 3 3 1 1 2 2 2 2
> AO_vec6 <- subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")$Score.QSet1
> AO_vec6
```

```
AO_vec7 <- subset(AO_lab, (Source=="USYD") & First.Questionnaire=="BPMN")$Score.Placebo
AO_vec7
[1]  1  3  1  4  3  1  3  0  4  2  3  3  2  2  3  1  0  2  1
AO_vec8 <- subset(AO_lab, (Source=="USYD") & First.Questionnaire=="BPMN")$Score.QSet1
AO_vec8
[1]  2  2  1  1  2  4  1  2  5  5  3  1  3  2  3  3  0  4  3
AO_fwilcoxTestP (AO_vec7, AO_vec8)

Wilcoxon signed rank test with continuity correction
data:  AO_v1 and AO_v2
V = 36, p-value = 0.1540
alternative hypothesis: true location shift is less than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "less", paired = TRUE,  :
   cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "less", paired = TRUE,  :
   cannot compute exact p-value with zeroes
> # H3
> AO_fwilcoxTestP <- function(AO_v1,AO_v2){ wilcox.test
     (AO_v1,AO_v2,alternative="less",paired=FALSE, exact = TRUE) }
> AO_vec9 <- subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")$Delta.QSet1.Placebo
> AO_vec9
[1]  1  2  3  2  2  0  -2  1  1  1  0  0  0  3  2  3  0  1  1  -1  2  2  0  0
  2  2
[26]  2  4  2  2  1  1  1  2  3  2  5  0  2  0  -1  3  0  2  0  3  1  0  0
  1  1
[51]  2  4  3  1  4  2  4  4  4  1  1  4  2  -1  4  3  -1  2  2  1  -2  0  0  1
  1  1
[76]  3  3  0  -1  1  1  3  2  2  0  -1  0  > AO_vec10 <- subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo
> AO_vec10
[1]  1  1  0  0  -1  -1  0  1  1  3  2  2  4  4  -1  1  2  -2  0  -2  1  2  0
  3  -1
[26]  0  2  3  -2  0  -1  -1  3  3  1  -1  0  0  4  3  1  3  2  1  2  -2  1  1
  0  1
[51]  0  1  3  1  3  2  -1  3  -1  5  2  1  1  0  0  0  0
> AO_fwilcoxTestP (AO_vec9, AO_vec10)

Wilcoxon rank sum test with continuity correction
data: AO_v1 and AO_v2  
W = 3396, p-value = 0.9462  
alternative hypothesis: true location shift is less than 0

Warning message:  
In wilcox.test.default(AO_v1, AO_v2, alternative = "less", paired = FALSE,  
:  
cannot compute exact p-value with ties

> AO_vec11 <- subset(AO_lab, (Source=="USYD") &  
First.Questionnaire=="BPMN")$Delta.QSet1.Placebo  
> AO_vec11  
[1] 1 -1 0 -3 -1 3 -2 2 1 3 0 -2 1 0 0 2 0 2 2  
> AO_vec12 <- subset(AO_lab, (Source=="USYD") &  
First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo  
> AO_vec12  
[1] 2 1 2 2 2 -1 -1 1 3 -1 0 0 2 2 1 1 0 1 1 1 -2 1  
> AO_fwilcoxTestP (AO_vec11, AO_vec12)

Wilcoxon rank sum test with continuity correction

data: AO_v1 and AO_v2  
W = 181.5, p-value = 0.2352  
alternative hypothesis: true location shift is less than 0

Warning message:  
In wilcox.test.default(AO_v1, AO_v2, alternative = "less", paired = FALSE,  
:  
cannot compute exact p-value with ties

> #H4  
> AO_vec13 <- subset(AO_lab, (Source=="HU/b" | Source=="TU/e") &  
First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo  
> AO_vec13  
[1] 1 1 0 0 -1 -1 0 1 1 3 2 2 4 4 -1 1 2 -2 0 -2 1 2 0 3 -1  
[26] 0 2 3 -2 0 -1 -1 3 3 1 -1 0 0 4 3 1 3 2 1 2 -2 1 1 0 1  
[51] 0 1 3 1 3 3 2 -1 3 -1 5 2 1 1 0 0 0 0  
> AO_vec14 <- subset(AO_lab, (Source=="HU/b" | Source=="TU/e") &  
First.Questionnaire=="BPMN")$Delta.QSet1.Placebo  
> AO_vec14  
[1] 1 2 3 2 2 0 -2 1 1 1 0 0 3 2 3 0 1 1 -1 2 2 0 0 2 2  
[26] 2 4 2 2 1 1 1 2 3 2 5 0 2 0 -1 3 0 2 0 3 1 0 0 1 1  
[51] 2 4 3 1 4 2 4 4 1 1 4 2 -1 4 3 -1 2 2 1 -2 0 0 1 1  
[76] 3 3 0 -1 1 1 3 2 2 0 -1 0  
> AO_vec15 <- subset(AO_lab, (Source=="USYD") &  
First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo  
> AO_vec15  
[1] 2 1 2 2 2 -1 -1 1 3 -1 0 0 2 2 1 1 0 1 1 1 -2 1  
> AO_vec16 <- subset(AO_lab, (Source=="USYD") &  
First.Questionnaire=="BPMN")$Delta.QSet1.Placebo
```r
> AO_vec16
[1]  1 -1  0 -3 -1  3 -2  2  1  3  0 -2  1  0  0  2  0  2  2
> AO_fwilcoxTestP (AO_vec15, AO_vec16)

Wilcoxon rank sum test with continuity correction

data:  AO_v1 and AO_v2
W = 236.5, p-value = 0.7729
alternative hypothesis: true location shift is less than 0

Warning message:
In wilcox.test.default(AO_v1, AO_v2, alternative = "less", paired = FALSE, :
  cannot compute exact p-value with ties
> AO_fwilcoxTestP <- function(AO_v1,AO_v2){ wilcox.test
  (AO_v1,AO_v2,alternative="less",paired=TRUE, exact = TRUE) }

> #H5
> AO_fwilcoxTestP <- function(AO_v1,AO_v2){ wilcox.test
  (AO_v1,AO_v2,alternative="less",paired=TRUE, exact = TRUE) }
> AO_vec17<-
  subset(AO_lab, (Source="HU/b" | Source="TU/e") &
First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo
> AO_vec17
[1]  1  1  0  0 -1 -1  0  1  1  3  2  2  4  4 -1  1  2  -2  0 -2  1  2  0
  3 -1
[26]  0  2  3 -2  0 -1 -1  3  3  1 -1  0  0  4  3  1  3  2  1  2  -2  1  1
  0  1
[51]  0  1  3  1  3  3  2 -1  3 -1  5  2  1  1  0  0  0  0
> AO_vec18<- subset(AO_lab, (Source="HU/b" | Source="TU/e") &
First.Questionnaire=="Cockburn")$Delta.QSet2.Placebo
> AO_vec18
[1]  1  2  0 -1  4  0  3  1  2  3  2  2  3  4 -2  3  2  0  2 -2  3  1  1
  3  0
[26]  2  4  3 -1  0  0  1  5  3  1  0  0  0  3  2  0  2  2  2  2  -1  1  1
  1  2
[51]  1 -1  3 -1  2  4 -2 -2  1  2 -1  5  3  3  2  0  0  1  0
> AO_fwilcoxTestP (AO_vec17, AO_vec18)

Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 246, p-value = 0.006626
alternative hypothesis: true location shift is less than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "less", paired = TRUE, :
  cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "less", paired = TRUE, :
  cannot compute exact p-value with zeroes
> AO_vec19<-
  subset(AO_lab, (Source="USYD") &
First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo
> AO_vec19
```

Page 19
> AO_vec20 <- subset(AO_lab, (Source == "USYD") & First.Questionnaire == "Cockburn")$Delta.QSet2.Placebo
> AO_vec20
> [1] 2 1 2 2 2 -1 -1 3 -1 0 0 2 2 1 1 0 1 1 1 -2 1

Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 7.5, p-value = 0.0002583
alternative hypothesis: true location shift is less than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "less", paired = TRUE,  :
   cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "less", paired = TRUE,  :
   cannot compute exact p-value with zeroes

> #H6
> AO_vec21 <- subset(AO_lab, (Source == "HU/b" | Source == "TU/e") & First.Questionnaire == "BPMN")$Delta.QSet1.Placebo
> AO_vec21
> [1] 1 2 3 2 2 0 -2 1 1 1 0 0 3 2 3 0 1 1 -1 2 2 0 0 2 2
[26] 2 4 2 2 1 1 1 2 3 2 5 0 2 0 -1 3 0 2 0 3 1 0 0 1 1
[51] 2 4 3 1 4 2 4 1 1 4 2 -1 4 3 -1 2 2 1 -2 0 0 1 1
[76] 3 3 0 -1 1 1 3 2 2 0 -1 0

Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 686, p-value = 0.9114
alternative hypothesis: true location shift is less than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "less", paired = TRUE,  :
   cannot compute exact p-value with ties
```r
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "less", paired = TRUE, :
cannot compute exact p-value with zeroes
> AO_vec23 <- subset(AO_lab, (Source=="USYD") &
First.Questionnaire=="BPMN")$Delta.QSet1.Placebo
> AO_vec23
[1] 1 -1 0 -3 -1 3 -2 2 1 3 0 -2 1 0 0 2 0 2 2
> AO_vec24 <- subset(AO_lab, (Source=="USYD") &
First.Questionnaire=="BPMN")$Delta.QSet2.Placebo
> AO_vec24
[1] 1 -2 2 -1 -1 1 -1 3 1 0 0 -1 1 -2 0 2 2 1 3
> AO_fwilcoxTestP(AO_vec23, AO_vec24)

Wilcoxon signed rank test with continuity correction
data:  AO_v1 and AO_v2
V = 37, p-value = 0.4521
alternative hypothesis: true location shift is less than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "less", paired = TRUE, :
cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "less", paired = TRUE, :
cannot compute exact p-value with zeroes
```
3.3.   Section 5.4 - Predictions

3.3.1.   Books Read

254 . *---------------------- for q23 ---------------------*

We did find one statistically significant predictor. When asked how many fiction books they had read in the past 12 month, 73 participants reported none, 52 reported one to three, 26 reported four to six, 13 reported seven to ten, and 22 reported more than ten. *Our findings suggest a strong relationship between readership and textual aptitude.* The range of the number of books that were read in the past 12 months ranged zero to more than ten. For the participants who first received written use cases as opposed to BPMN, the number of books they read were clustered via a Wards Hierarchical clustering routine, which revealed three distinct clusters of readership intensity. The first cluster comprised low readership, the third cluster comprised high readership levels. The primary contribution of the written use cases was regressed over readership intensity. Low readership significantly predicted a negative effect on the primary contribution with written use cases: $-0.2370 \ (0.1125; -2.1070)$ and high readership predicted a significant positive effect: $0.4712 \ (0.1918; 2.4560)$. We also investigated this factor among the participants who received BPMN first: high readership predicted a significant negative effect on primary contribution from the BPMN artifact: $-0.3911 \ (0.0867; -4.5092)$, but low readership was not a statistically significant predictor for success with BPMN: $-0.1250 \ (0.1814; -0.6890)$.

Figure 3 – Q23 The number of fiction books you read last 12 month is:
255 .
256 . cluster wardslinkage q23 deltaqset1placebo if first==2, measure(L2)
cluster name: _cl_1
257 . cluster dendrogram _cl_1, cutnumber(21)
258 . cluster gen cluster1=groups(3), name(_cl_
259 .
260 . reg deltaqset1placebo2 q23 if cluster1==1 & first==2

Source |       SS       df       MS              Number of obs =       56
-------------+------------------------------------------------------------
Model |  3.94484813     1  3.94484813           Prob > F      =  0.0398
Residual |  47.9837233    54  .888587468           R-squared     =  0.0760
Total |  51.9285714    55  .944155844           Root MSE      =  .94265

deltaqset1~2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+---------------------------------------------------------------
q23 |  -.2370295   .1124961  -2.11   0.040     -.4625705   -.0114885
_cons |   4.717701   .3791158    12.44   0.000     3.957619    5.477782
-------------+---------------------------------------------------------------

261 . estimates store a

262 . reg deltaqset1placebo2 q23 if cluster1==2 & first==2

Source |       SS       df       MS              Number of obs =       7
-------------+------------------------------------------------------------
Model |           0     0
Residual |  1.71428571     6  .285714286           R-squared     =  0.0000
Total |  1.71428571     6  .285714286           Root MSE      =  .53452

deltaqset1~2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+---------------------------------------------------------------
q23 |  (dropped)
_cons |   5.428571   .2020305    26.87   0.000     4.934221    5.922922
-------------+---------------------------------------------------------------

263 . estimates store b

264 . reg deltaqset1placebo2 q23 if cluster1==3 & first==2

Source |       SS       df       MS              Number of obs =      30
-------------+------------------------------------------------------------
Model |  3.07820513     1  3.07820513           Prob > F      =  0.0205
Residual | 14.2884615    28  .510302198           R-squared     =  0.1772
Total |  17.3666667    29  .598850575           Root MSE      =  .71435

deltaqset1~2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+---------------------------------------------------------------
q23 |   .4711538   .1918349     2.46   0.021     .0781978    .8641099
_cons |   5.384615   .5776323     9.32   0.000     4.201389    6.567842
-------------+---------------------------------------------------------------

265 . estimates store c

266 . mean deltaqset1placebo2, over(cluster1)

Mean estimation                      Number of obs      =       93
1: cluster1 = 1
2: cluster1 = 2
3: cluster1 = 3

-------------------------------+-----------------------
Over |   Mean   Std. Err.    [95% Conf. Interval]
-------------------------------+-----------------------
deltaqset1~2 |
267. `est table a b c, b(%9.4f) se(%9.4f) t(%9.4f) stats(N r2_a)`

| Variable |     a           b           c |
|----------|----------------|----------------|
|          |                |                |
| q23      | 0.2370        | 0.0000        | 0.4712 |
|          | 0.1125        | 0.0000        | 0.1918 |
|          | -2.1070       |                | 2.4560 |
| _cons    | 4.7177        | 5.4286        | 5.3846 |
|          | 0.3791        | 0.2020        | 0.5776 |
|          | 12.4440       | 26.8701       | 9.3219 |
| N        | 56.0000       | 7.0000        | 30.0000 |
| r2_a     | 0.0589        | 0.0000        | 0.1479 |

`legend: b/se/t`

268. `cluster drop _all`

269. `drop cluster`

270. `*------------------------ for q23 ------------------------*`

271. `cluster wardslinkage q23 deltaqset1placebo if first==1, measure(L2)`

272. `cluster name: _cl_1`

273. `// rename _cl_1 _cl_2`

274. `cluster dendrogram _cl_1, cutnumber(21)`

275. `cluster gen cluster2=groups(3), name(_cl_1)`

276. `tab cluster`

<table>
<thead>
<tr>
<th>cluster2</th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>22</td>
<td>20.56</td>
<td>20.56</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>33.64</td>
<td>54.21</td>
</tr>
<tr>
<td>3</td>
<td>49</td>
<td>45.79</td>
<td>100.00</td>
</tr>
</tbody>
</table>

278. `recode cluster 3=1 1=2 1=3`

279. `reg deltaqset1placebo2 q23 if cluster2==1 & first==1`

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 49</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>0.4942</td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>20.7998588</td>
<td>1</td>
<td>20.7998588</td>
<td>Prob &gt; F = 0.0000</td>
</tr>
<tr>
<td>Residual</td>
<td>49.5</td>
<td>47</td>
<td>1.05319149</td>
<td>R-squared = 0.0111</td>
</tr>
<tr>
<td>Total</td>
<td>50</td>
<td>48</td>
<td>1.04166667</td>
<td>Root MSE = 1.0263</td>
</tr>
</tbody>
</table>

280. `estimates store d`

281. `reg deltaqset1placebo2 q23 if cluster2==2 & first==1`

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 58</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>20.7998588</td>
<td>1</td>
<td>20.7998588</td>
<td>Prob &gt; F = 0.0000</td>
</tr>
</tbody>
</table>
Residual |  57.2863481    56   1.0229705           R-squared     =  0.2664
---------+-------------------------------------------------------------------
Total |  78.0862069    57  1.36993345          Root MSE      =  1.0114

---+-------------------------------------------------------------------
      |         Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---+-------------------------------------------------------------------
deltaqset1~2 |  g23 |  -0.3910508    0.0867231  -4.51   0.000    -0.564778   -0.2173237
      +--------------------------
      _cons |  7.673723     0.3478255   22.06   0.000     6.976945    8.370501
---+-------------------------------------------------------------------

estimates store e
// qreg deltaqset1placebo2 q23 if cluster2=3 & first==1 // insufficient obs
estimates store f
mean deltaqset1placebo2, over(cluster2)
Mean estimation                     Number of obs    =     107
1: cluster2 = 1
2: cluster2 = 2

-------------------------------------------------------------------
      Over |       Mean   Std. Err.     [95% Conf. Interval]
---------+-------------------------------------------------------------------
deltaqset1~2 |  1 | 4.000000  .1458030    3.710931    4.289069
  2 | 6.224138  .1536865    5.919439    6.528836
-------------------------------------------------------------------
est table d e, b(%9.4f) se(%9.4f) t(%9.4f) stats(N r2_a)

Variable |   d           e
---------+-------------------------------------------------------------------
g23 |  -0.1250     -0.3911
     |  0.1814     0.0867
     | -0.6890    -4.5092
     +--------------------------
     _cons |   4.3571     7.6737
     |  0.5387     0.3478
     |  8.0887    22.0620
---------+-------------------------------------------------------------------
     N |  49.0000    58.0000
     r2_a |  -0.0111     0.2533
---------+-------------------------------------------------------------------

legend: b/se/t
cluster drop _all
drop cluster
3.3.2. Comfort with charts

When asked if they were comfortable with flow charts, 82 participants responded that they strongly agree and one participant responded with strong disagreement. The remaining participants reported preferences somewhere in the middle. Participants at the higher end of the scale did not perform differently when compared to those at the lower end, with model estimates of $-0.2258 (0.1507; -1.4980)$ for comfort, and $0.0642 (0.1576; 0.4075)$ for lack of comfort.

Figure 4 Q24 You are comfortable with flow charts:
cluster wardslinkage q24 deltaqset1placebo if first==1, measure(L2)

cluster name: _cl_1

cluster dendrogram _cl_1, cutnumber(20)

cluster gen clusterq24 = groups(2), name(_cl_1)

counts:

```
Source |       SS       df       MS              Number of obs =      60
-------------+----------------------------------------------------------------
Model |  2.10752688     1  2.10752688           Prob > F      =  0.1396
Residual |  54.4758065    58  .939238042           R-squared     =  0.0372
-------------+----------------------------------------------------------------
Total |  56.5833333    59  .959039548           Root MSE      =  .96914
-------------+----------------------------------------------------------------
deltaqset1-o |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
q24 |   -.002258   .1507432    -1.50   0.140     -.5275517    .0759388
_cons |   .6854839   .4210027     1.63   0.109     -.1572447    1.528212
-------------+----------------------------------------------------------------
```

counts:

```
Source |       SS       df       MS              Number of obs =      47
-------------+----------------------------------------------------------------
Model |  .113447731     1  .113447731           Prob > F      =  0.6855
Residual |  30.7376161    45  .683058136           R-squared     =  0.0037
-------------+----------------------------------------------------------------
Total |  30.8510638    46  .670675301           Root MSE      =  .82647
-------------+----------------------------------------------------------------
deltaqset1-o |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
q24 |   .0642415   .1576328     0.41   0.686     -.2532472    .3817302
_cons |    2.47291   .4233479     5.84   0.000     1.620244    3.325577
-------------+----------------------------------------------------------------
```

counts:

```
  mean deltaqset1placebo2, over(clusterq24)

Mean estimation                     Number of obs =  107
               1: clusterq24 = 1
               2: clusterq24 = 2

Over |       Mean   Std. Err.     [95% Conf. Interval]
---------------+---------------------------------------------
deltaqset1-2 |          
  1 |   4.083333   .1264278     3.832678    4.333989
  2 |   6.638298   .1194558     6.401465    6.875131
---------------+---------------------------------------------
```

counts:

```
est table a b, b(%9.4f) se(%9.4f) t(%9.4f) stats(N r2_a)

Variable |     a           b
------------------------
q24 |    -0.2258      0.0642
     |    0.1507      0.1576
     |   -1.4980      0.4075
_cons |    0.6855      2.4729
     |    0.4210      0.4233
     |    1.6282      5.8413
     |----------------------------------------------------------------
N |   60.0000     47.0000   
r2_a |    0.0206     -0.0185
```

Page 27
307 .
308 . cluster drop _all
309 . drop cluster
310 .
3.3.3. Comfort with Use cases

When asked if they were comfortable with written use cases, 73 participants responded that they strongly agree and 16 participants strongly disagreed. The remaining participants reported preferences somewhere in the middle. Participants at the higher end of the scale did not perform differently when compared to those at the lower end, with model estimates of $-0.0823$ ($0.0986; -0.8350$) for comfort and $0.0701$ ($0.1027; 0.6830$) for lack of comfort.

Figure 5 Q26 You are comfortable with written use cases:

```plaintext
312 .
313 . cluster wardslinkage q26 deltaqset1placebo if first==2, measure(L2)
    cluster name: _cl_1
314 . cluster dendrogram _cl_1, cutnumber(20)
315 . cluster gen clusterq26 = groups(2), name(_cl_1)
316 . mean deltaqset1placebo if clusterq26==1 & first==2
    Mean estimation                     Number of obs    =      61
      1: clusterq26 = 1
      2: clusterq26 = 2
    --------------------------------------------------------------
    Over |       Mean   Std. Err.   [95% Conf. Interval]
    -------------+----------------------------------------------
    deltaqset1-2 |  
      1 |  0.69736765  .69736765
      2 |  0.995628415
    --------------------------------------------------------------
317 .
318 . reg deltaqset1placebo q26 if clusterq26==1 & first==2
    Source |       SS       df       MS              Number of obs =      61
            |            |            |                                F(  1,    59) =    0.70
            |            |            |                                Prob > F      =  0.4071
    Model |  .69736765   1  .69736765                                R-squared     =  0.0117
    Residual |  59.0399682  59  1.00067743                               Adj R-squared = -0.0051
    Total |  59.7377049  60  .995628415                                Root MSE      =  1.0003
    ------------------------------------------------------------------------------
    deltaqset1-o |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
            |            |            |            |            |                                [95% Conf. Interval]
    q26 |   -.0823248   .0985989   -0.84   0.407   -.2796026   .114953
    _cons |   .3840764   .4023593     0.95   0.344   -.4210427   1.189196
    ------------------------------------------------------------------------------
319 . estimates store a
320 . reg deltaqset1placebo q26 if clusterq26==2 & first==2
    Source |       SS       df       MS              Number of obs =      32
            |            |            |                                F(  1,    30) =  0.47
    ------------------------------------------------------------------------------
```
Model |  .282779364     1  .282779364           Pro
Residual |  18.1859706    30  .606199021           R-
          squared     =  0.0153
Total |  18.46875    31  .595766129           Root MSE =  .77859
          -------------
          Adj R-
          squared = -0.0175
          ------------------------------
deltagset1-o |      Coef.   Std. Err.      t    P>|t|   [95% Conf. Interval]
          -------------
          q26 |   .0701468    .102705     0.68   0.500   -.1396048    .2798985
          _cons |   2.453507    .412022     5.95   0.000    1.612046    3.294968
          321 . estimates store b
          322 . mean deltagset1placebo2, over(clusterq26)
          Mean estimation                     Number of obs    =      93
          1: clusterq26 = 1
          2: clusterq26 = 2
          Number of obs    =      93
          Over |       Mean   Std. Err.     [95% Conf. Interval]
          -------------
          deltagset1~2 |            1 |   4.065574   .1277567      3.811838     4.31931
          2 |    6.71875   .1364467      6.447755    6.989745
          323 . est table a b, b(%9.4f) se(%9.4f) t(%9.4f) stats(N r2_a)
          Variable |     a           b
          -------------
          q26 |  -0.0823      0.0701
          _cons |  0.0986      0.1027
          |  -0.8350      0.6830
          |  0.3841      2.4535
          |  0.4024      0.4120
          |   0.9546      5.9548
          -------------
          N |   61.0000     32.0000
          r2_a |  -0.0051     -0.0175
          legend: b/se/t
          324 .
          325 . cluster drop _all
          326 . drop cluster
          327 .
3.3.4. Experience with flow charts

When asked if they often worked with flow charts, 30 participants strongly agreed and six participants responded with strong disagreement. The remaining participants reported preferences somewhere in the middle. Participants at the higher end of the scale performed significantly better 0.2669 (0.1558; 2.3104). The performance of participants at the lower end varied −0.0865 (0.1351; −0.6404).

Figure 6 – Q25 You often work with flow charts:

329. *** coding the missing value for q25
330. recode q25 1=. 2=5 3=4 4=3 5=2 6=1
(q25: 200 changes made)

331. tab q25

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>.</td>
<td>6</td>
<td>3.05</td>
<td>3.05</td>
</tr>
<tr>
<td>a</td>
<td>40</td>
<td>20.30</td>
<td>23.35</td>
</tr>
<tr>
<td>b</td>
<td>10</td>
<td>5.08</td>
<td>28.43</td>
</tr>
<tr>
<td>c</td>
<td>111</td>
<td>56.35</td>
<td>84.77</td>
</tr>
<tr>
<td>d</td>
<td>30</td>
<td>15.23</td>
<td>100.00</td>
</tr>
<tr>
<td>Total</td>
<td>197</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

332. tab q25, nol

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>3.05</td>
<td>3.05</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>20.30</td>
<td>23.35</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>5.08</td>
<td>28.43</td>
</tr>
<tr>
<td>4</td>
<td>111</td>
<td>56.35</td>
<td>84.77</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>15.23</td>
<td>100.00</td>
</tr>
<tr>
<td>Total</td>
<td>197</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

333. label define q25l 1 "a" 2 "b" 3 "c" 4 "d" 5 "e"

334. label values q25 q25l

335. tab q25

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>6</td>
<td>3.05</td>
<td>3.05</td>
</tr>
<tr>
<td>b</td>
<td>40</td>
<td>20.30</td>
<td>23.35</td>
</tr>
<tr>
<td>c</td>
<td>10</td>
<td>5.08</td>
<td>28.43</td>
</tr>
<tr>
<td>d</td>
<td>111</td>
<td>56.35</td>
<td>84.77</td>
</tr>
<tr>
<td>e</td>
<td>30</td>
<td>15.23</td>
<td>100.00</td>
</tr>
<tr>
<td>Total</td>
<td>197</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

336.
337. cluster wardslinkage q25 deltaqset1placebo if first==1, measure(L2)
    cluster name: _cl_1

338. cluster dendrogram _cl_1, cutnumber(20)
339 . cluster gen clusterq25= groups(3), name(_cl_1)

340 . mean deltaqset1placebo2, over(clusterq25)

Mean estimation                     Number of obs    =     105

1: clusterq25 = 1
2: clusterq25 = 2
3: clusterq25 = 3

--------------------------------------------------------------
Over |       Mean   Std. Err.     [95% Conf. Interval]
------- |-----------:        :-----------------:-----------------:
1 |  3.682927   .1458324      3.393736    3.972118
2 |          7   .1516196      6.699333    7.300667
3 |        5.5   .087038       5.327399    5.672601
--------------------------------------------------------------

341 . recode clusterq25 3=2 2=3
(clusterq25: 64 changes made)

342 . tab clusterq25

clusterq25 |      Freq.     Percent
Cum.
------------|-------------------:---:
1 |         41       39.05       39.05
2 |         34       32.38       71.43
3 |         30       28.57      100.00
------------|-------------------:---:-------------------:

Total |        105      100.00

343 . reg deltaqset1placebo2 q25 if clusterq25==1 & first==1

Source |       SS       df       MS              Number of obs =      41
-------- |-----------:        :-----------:-----------------:
Model    |  .362958237     1  .362958237           Prob > F      =  0.5257
Residual |  34.5150905    39  .885002322           R-squared     =  0.0104
---------- |-----------------:-----------------:-----------------:
Adj R-squared = -0.0150
Total |  34.8780488    40  .87195122           Root MSE      =  .94075

deltaqset1-2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------|-----------:        :-----------:-------------------:
q25 |  -.0865191   .1351002     -0.64  0.526     -.359785    .1867468
_cons |   3.967807   .4684761     8.47   0.000     3.020224    4.915389

344 . reg deltaqset1placebo2 q25 if clusterq25==2 & first==1

Source |       SS       df       MS              Number of obs =      34
-------- |-----------:        :-----------:-----------------:
Model    |  .202380952     1  .202380952           Prob > F      =  0.3836
Residual |  8.29761905    32  .259300595           R-squared     =  0.0238
---------- |-----------------:-----------------:-----------------:
Adj R-squared = -0.0067
Total |          8.5    33  .257575758           Root MSE      =  .50922

deltaqset1-2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------|-----------:        :-----------:-------------------:
q25 |    .202381   .2290797     0.88   0.384     -.2642391    .669001
_cons |   4.654762   .9607219     4.85   0.000     3.020224    6.297835

345 . estimates store a

346 . reg deltaqset1placebo2 q25 if clusterq25==3 & first==1

Source |       SS       df       MS              Number of obs =      30
-------- |-----------:        :-----------:-----------------:
Model    |  3.20237213     1  3.20237213           Prob > F      =  0.0284
Residual |  16.7976279    28  .599915281           R-squared     =  0.1601
---------- |-----------------:-----------------:-----------------:
Adj R-squared =  0.1301
Total |          20    29  .689655172           Root MSE      =  .77454
349. estimates store c
350. mean deltaqset1placebo2, over(clusterq25)

Mean estimation                     Number of obs    =     105
1: clusterq25 = 1
2: clusterq25 = 2
3: clusterq25 = 3

--------------------------------------------------------------
Over |       Mean   Std. Err.     [95% Conf. Interval]
--------------------------------------------------------------
deltaqset1~2 | 3.682927   .1458324      3.393736    3.972118
1 | 5.5      .0870388      5.327399    5.672601
2 | 7       .1516196      6.699333    7.300667
--------------------------------------------------------------
351. est table a b c, b(%9.4f) se(%9.4f) t(%9.4f) stats(N r2_a)

Variable |     a           b           c
-------------
q25 | -0.0865      0.2024      0.2669
|    0.1351
|    0.6404      0.8835      2.3104
_cons |  3.9678      4.6548      6.1905
|    0.4685      0.9607      0.3778
|    8.4696      4.8451     16.3846
-------------
N |   41.0000     34.0000     30.0000
r2_a |  0.0150      0.0067      0.1301
-------------

legend: b/se/t

352.
353. cluster drop _all
354. drop cluster

355.
356. *** In Avner's paper, we are only reporting model 'a' and 'b'
357. *** only to be consistent with the dual model situation that
358. *** has manifested!
3.3.5. Experience with use cases

When asked if they often worked with written use cases, 36 participants responded that they strongly agree and 20 participant responded with strong disagreement. The remaining participants reported preferences somewhere in the middle. Participants at the higher end of the scale performed significantly better at the 0.1 level 1.1875 (0.6426; 1.8479). The performance of participants at the lower end of the scale varied 0.1714 (0.3569; 0.4803).

Figure 7 – Q27 You often work with written use cases

* *** coding the missing value for q25
362. recode q27 1=1, 2=5 3=4 4=3 5=2 6=1
 (q27: 199 changes made)

363. tab q27

<table>
<thead>
<tr>
<th>Q27</th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>.</td>
<td>20</td>
<td>10.15</td>
<td>10.15</td>
</tr>
<tr>
<td>a</td>
<td>41</td>
<td>20.81</td>
<td>30.96</td>
</tr>
<tr>
<td>b</td>
<td>16</td>
<td>8.12</td>
<td>39.09</td>
</tr>
<tr>
<td>c</td>
<td>83</td>
<td>42.13</td>
<td>81.22</td>
</tr>
<tr>
<td>d</td>
<td>36</td>
<td>18.27</td>
<td>99.49</td>
</tr>
<tr>
<td>f</td>
<td>1</td>
<td>0.51</td>
<td>100.00</td>
</tr>
</tbody>
</table>

364. tab q27, nol

<table>
<thead>
<tr>
<th>Q27</th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>10.15</td>
<td>10.15</td>
</tr>
<tr>
<td>2</td>
<td>41</td>
<td>20.81</td>
<td>30.96</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>8.12</td>
<td>39.09</td>
</tr>
<tr>
<td>4</td>
<td>83</td>
<td>42.13</td>
<td>81.22</td>
</tr>
<tr>
<td>5</td>
<td>36</td>
<td>18.27</td>
<td>99.49</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0.51</td>
<td>100.00</td>
</tr>
</tbody>
</table>

365. label define q27l 1 "a" 2 "b" 3 "c" 4 "d" 5 "e"

366. label values q27 q27l

367. tab q27

<table>
<thead>
<tr>
<th>Q27</th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>20</td>
<td>10.15</td>
<td>10.15</td>
</tr>
<tr>
<td>b</td>
<td>41</td>
<td>20.81</td>
<td>30.96</td>
</tr>
<tr>
<td>c</td>
<td>16</td>
<td>8.12</td>
<td>39.09</td>
</tr>
<tr>
<td>d</td>
<td>83</td>
<td>42.13</td>
<td>81.22</td>
</tr>
<tr>
<td>e</td>
<td>36</td>
<td>18.27</td>
<td>99.49</td>
</tr>
<tr>
<td>f</td>
<td>1</td>
<td>0.51</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Total | 197 | 100.00
cluster wardslinkage q27 deltaqset1placebo if first==2, measure(L2)
cluster name: _cl_1

cluster dendrogram _cl_1, cutnumber(20)
cluster gen clusterq27= groups(3), name(_cl_1)
mean deltaqset1placebo2, over(clusterq27)

Mean estimation                     Number of obs    =      92

Over |       Mean   Std. Err.     [95% Conf. Interval]
------------- | ---------------------- ----------------------
deltaqset1~2 |                        |
    1 |   5.882353   .1566398      5.571207    6.193499
    2 |   5.354839   .3024928      4.753974    5.955704
    3 |   3.444444   .1540834      3.138377    3.750512

recode clusterq27 1=3 3=1
(clusterq27: 61 changes made)
tab clusterq27

reg deltaqset1placebo q27 if clusterq27==1 & first==2
Source |       SS       df       MS              Number of obs =      27
------------- | ---------------------- ----------------------
Model |  .152380952     1  .152380952           Prob > F      =  0.6352
Residual |  16.5142857    25  .660571429           R-squared     =  0.0091
Total |  16.6666667    26  .641025641           Root MSE      =  .81276

deltaqset1~o |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------- | ---------------------- ----------------------
q27 |   .1714286   .3569256     0.48   0.635    -.5636736    .9065307
_cons |  -1.285714   1.528264    -0.84   0.408    -4.433234    1.861805

estimates store a

reg deltaqset1placebo q27 if clusterq27==2 & first==2
Source |       SS       df       MS              Number of obs =      31
------------- | ---------------------- ----------------------
Model |  1.78571429     1  1.78571429           Prob > F      =  0.4369
Residual |  83.3110599    29  2.87279517           R-squared     =  0.0210
Total |  85.0967742    30  2.83655914           Root MSE      =  1.6949

deltaqset1~o |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------- | ---------------------- ----------------------
q27 |  -.3571429   .4529897    -0.79   0.437    -1.283611    .5693251
_cons |   2.069124   .9557559     2.16   0.039     .114384    4.023865

estimates store b

reg deltaqset1placebo q27 if clusterq27==3 & first==2
<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 34</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2.65441176</td>
<td>1</td>
<td>2.65441176</td>
<td>F(1, 32) = 3.41</td>
</tr>
<tr>
<td>Residual</td>
<td>24.875</td>
<td>32</td>
<td>.77734375</td>
<td>R-squared = 0.0964</td>
</tr>
<tr>
<td>Total</td>
<td>27.5294118</td>
<td>33</td>
<td>.834224599</td>
<td>Root MSE = 0.88167</td>
</tr>
</tbody>
</table>

| deltaqset1-o | Coef.    | Std. Err. | t     | P>|t|  | [95% Conf. Interval] |
|--------------|----------|------------|-------|------|----------------------|
| q27          | 1.1875   | .6426226   | 1.85  | 0.074 | -1.214795 2.496479   |
| _cons        | -2.9375  | 2.612671   | -1.12 | 0.269 | -8.259337 2.384337  |

```
381. estimates store c
382. mean deltaqset1placebo2, over(clusterq27)
```

Mean estimation

```
Number of obs = 92
```

```
1: clusterq27 = 1
2: clusterq27 = 2
3: clusterq27 = 3
```

```
<table>
<thead>
<tr>
<th>deltaqset1-o</th>
<th>Over</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>3.444444</td>
<td>.1540834</td>
<td>3.138377 3.750512</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5.354839</td>
<td>.3024928</td>
<td>4.753974 5.955704</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5.882353</td>
<td>.1566398</td>
<td>5.571207 6.193499</td>
</tr>
</tbody>
</table>
```

```
383. est table a b c, b(%9.4f) se(%9.4f) t(%9.4f) stats(N r2_a)
```

```
<table>
<thead>
<tr>
<th>Variable</th>
<th>a</th>
<th>b</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>q27</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1714</td>
<td>-0.3571</td>
<td>1.1875</td>
</tr>
<tr>
<td></td>
<td>0.3569</td>
<td>0.4530</td>
<td>0.6426</td>
</tr>
<tr>
<td></td>
<td>0.4803</td>
<td>-0.7884</td>
<td>1.8479</td>
</tr>
<tr>
<td></td>
<td>-1.2857</td>
<td>2.0691</td>
<td>-2.9375</td>
</tr>
<tr>
<td></td>
<td>1.5283</td>
<td>0.9558</td>
<td>2.6127</td>
</tr>
<tr>
<td></td>
<td>-0.8413</td>
<td>2.1649</td>
<td>-1.1243</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>27.0000</td>
<td>31.0000</td>
</tr>
<tr>
<td></td>
<td>r2_a</td>
<td>-0.0305</td>
<td>-0.0128</td>
</tr>
</tbody>
</table>
```

```
384.
385. cluster drop _all
386. drop cluster
387.
388.
389. log close
   log: C:\Documents and Settings\com\Desktop\Avner Ottensooser\avner.smc
```

```
> 1
log type: smcl
closed on: 15 May 2010, 17:52:22
```
3.4. Statistical tests for normality

Test each of the statistics for each of the populations for normality using Shapiro test.

Please note that the P value is less important for this test!

Possible interpretation: the more information is reviled, the less do the scores distribute normally.

Table 8 - Statistical tests for normality

<table>
<thead>
<tr>
<th>University</th>
<th>First Questioner</th>
<th>Statistic</th>
<th>Shapiro test results</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>USYD</td>
<td>BPMN</td>
<td>Placebo</td>
<td>W = 0.9171, p-value = 0.09987</td>
<td>shapiro.test(subset(AO_lab, Source==&quot;USYD&quot; &amp; First.Questionnaire==&quot;BPMN&quot;)$Score.Placebo)</td>
</tr>
<tr>
<td>USYD</td>
<td>BPMN</td>
<td>QSet1</td>
<td>W = 0.9428, p-value = 0.2955</td>
<td>shapiro.test(subset(AO_lab, Source==&quot;USYD&quot; &amp; First.Questionnaire==&quot;BPMN&quot;)$Score.QSet1)</td>
</tr>
<tr>
<td>USYD</td>
<td>BPMN</td>
<td>QSet2</td>
<td>W = 0.8969, p-value = 0.04278</td>
<td>shapiro.test(subset(AO_lab, Source==&quot;USYD&quot; &amp; First.Questionnaire==&quot;BPMN&quot;)$Score.QSet2)</td>
</tr>
<tr>
<td>USYD</td>
<td>Cockburn</td>
<td>Placebo</td>
<td>W = 0.8065, p-value = 0.0006334</td>
<td>shapiro.test(subset(AO_lab, Source==&quot;USYD&quot; &amp; First.Questionnaire==&quot;Cockburn&quot;)$Score.Placebo)</td>
</tr>
<tr>
<td>USYD</td>
<td>Cockburn</td>
<td>QSet1</td>
<td>W = 0.8825, p-value = 0.01352</td>
<td>shapiro.test(subset(AO_lab, Source==&quot;USYD&quot; &amp; First.Questionnaire==&quot;Cockburn&quot;)$Score.QSet1)</td>
</tr>
<tr>
<td>USYD</td>
<td>Cockburn</td>
<td>QSet2</td>
<td>W = 0.8206, p-value = 0.001069</td>
<td>shapiro.test(subset(AO_lab, Source==&quot;USYD&quot; &amp; First.Questionnaire==&quot;Cockburn&quot;)$Score.QSet2)</td>
</tr>
<tr>
<td>TU/e</td>
<td>BPMN</td>
<td>Placebo</td>
<td>W = 0.9172, p-value = 0.0001304</td>
<td>shapiro.test(subset(AO_lab, Source==&quot;TU/e&quot; &amp; First.Questionnaire==&quot;BPMN&quot;)$Score.Placebo)</td>
</tr>
<tr>
<td>TU/e</td>
<td>BPMN</td>
<td>QSet1</td>
<td>W = 0.9434, p-value = 0.002397</td>
<td>shapiro.test(subset(AO_lab, Source==&quot;TU/e&quot; &amp; First.Questionnaire==&quot;BPMN&quot;)$Score.QSet1)</td>
</tr>
<tr>
<td>TU/e</td>
<td>BPMN</td>
<td>QSet2</td>
<td>W = 0.9484, p-value = 0.004424</td>
<td>shapiro.test(subset(AO_lab, Source==&quot;TU/e&quot; &amp; First.Questionnaire==&quot;BPMN&quot;)$Score.QSet2)</td>
</tr>
<tr>
<td>TU/e</td>
<td>Cockburn</td>
<td>Placebo</td>
<td>W = 0.8971, p-value = 0.0001983</td>
<td>shapiro.test(subset(AO_lab, Source==&quot;TU/e&quot; &amp; First.Questionnaire==&quot;Cockburn&quot;)$Score.Placebo)</td>
</tr>
<tr>
<td>TU/e</td>
<td>Cockburn</td>
<td>QSet1</td>
<td>W = 0.9459, p-value = 0.01524</td>
<td>shapiro.test(subset(AO_lab, Source==&quot;TU/e&quot; &amp; First.Questionnaire==&quot;Cockburn&quot;)$Score.QSet1)</td>
</tr>
<tr>
<td>University</td>
<td>First Questioner</td>
<td>Statistic</td>
<td>Shapiro test results</td>
<td>Formula</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------</td>
<td>-----------</td>
<td>----------------------</td>
<td>---------</td>
</tr>
<tr>
<td>TU/e</td>
<td>Cockburn</td>
<td>QSet2</td>
<td>$W = 0.945, p-value = 0.0139$</td>
<td><code>shapiro.test(subset(AO_lab,Source==&quot;TU/e&quot; &amp; First.Questionnaire==&quot;Cockburn&quot;)$Score.QSet2)</code></td>
</tr>
<tr>
<td>HU/b</td>
<td>BPMN Placebo</td>
<td>W = 0.8089, p-value = 0.0087</td>
<td><code>shapiro.test(subset(AO_lab,Source==&quot;HU/b&quot; &amp; First.Questionnaire==&quot;BPMN&quot;)$Score.Placebo)</code></td>
<td></td>
</tr>
<tr>
<td>HU/b</td>
<td>BPMN QSet1</td>
<td>W = 0.8297, p-value = 0.01569</td>
<td><code>shapiro.test(subset(AO_lab,Source==&quot;HU/b&quot; &amp; First.Questionnaire==&quot;BPMN&quot;)$Score.QSet1)</code></td>
<td></td>
</tr>
<tr>
<td>HU/b</td>
<td>BPMN QSet2</td>
<td>W = 0.8933, p-value = 0.1082</td>
<td><code>shapiro.test(subset(AO_lab,Source==&quot;HU/b&quot; &amp; First.Questionnaire==&quot;BPMN&quot;)$Score.QSet2)</code></td>
<td></td>
</tr>
<tr>
<td>HU/b</td>
<td>Cockburn Placebo</td>
<td>W = 0.8459, p-value = 0.02523</td>
<td><code>shapiro.test(subset(AO_lab,Source==&quot;HU/b&quot; &amp; First.Questionnaire==&quot;Cockburn&quot;)$Score.Placebo)</code></td>
<td></td>
</tr>
<tr>
<td>HU/b</td>
<td>Cockburn QSet1</td>
<td>W = 0.9554, p-value = 0.6813</td>
<td><code>shapiro.test(subset(AO_lab,Source==&quot;HU/b&quot; &amp; First.Questionnaire==&quot;Cockburn&quot;)$Score.QSet1)</code></td>
<td></td>
</tr>
<tr>
<td>HU/b</td>
<td>Cockburn QSet2</td>
<td>W = 0.9474, p-value = 0.5598</td>
<td><code>shapiro.test(subset(AO_lab,Source==&quot;HU/b&quot; &amp; First.Questionnaire==&quot;Cockburn&quot;)$Score.QSet2)</code></td>
<td></td>
</tr>
</tbody>
</table>
4. Work not published in the paper

The following work was not found sufficiently important to be published.

4.1. Comparing comprehension of the first questioner and Placebo

4.1.1. USYD Students, BPMN or Placebo

$H_0$ – USYD Students don’t comprehend BPMN better than placebo – ✔ Not rejected

$H_1$ – USYD students comprehend BPMN better than placebo – ✗ Rejected

4.1.1.1. USYD Results - First notation BPMN

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>Sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score.Placebo</td>
<td>1.50245600</td>
<td>2.05263200</td>
<td>2.60280700</td>
<td>1.22355100</td>
<td>19</td>
</tr>
<tr>
<td>Score.QSet1</td>
<td>1.84903200</td>
<td>2.47368400</td>
<td>3.09833600</td>
<td>1.38918100</td>
<td>19</td>
</tr>
<tr>
<td>Delta.QSet1.Placebo</td>
<td>-0.34784090</td>
<td>0.42105260</td>
<td>43</td>
<td>1.18994620</td>
<td>19</td>
</tr>
</tbody>
</table>

43 Here we see very little progress. BPMN score is similar to placebo score.
4.1.1.2. R Code

AO_ftTestP <- function(AO_v1, AO_v2) {
  t.test(AO_v1, AO_v2, alternative = "greater", paired = TRUE)
}

AO_ftTestP (subset(AO_lab, Source == "USYD" & First.Questionnaire == "BPMN")$Score.QSet1, subset(AO_lab, Source == "USYD" & First.Questionnaire == "BPMN")$Score.Placebo)

AO_fwilcoxTestP <- function(AO_v1, AO_v2) {
  wilcox.test(AO_v1, AO_v2, alternative = "greater", paired = TRUE, exact = TRUE)
}

AO_fwilcoxTestP (subset(AO_lab, Source == "USYD" & First.Questionnaire == "BPMN")$Score.QSet1, subset(AO_lab, Source == "USYD" & First.Questionnaire == "BPMN")$Score.Placebo)

Table 9 – H₀ – USYD Students don’t comprehend BPMN better than placebo – ☒ Not rejected

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1: USYD students who got BPMN first</td>
<td>answer to QSet1</td>
</tr>
<tr>
<td>Set 2: USYD students who got BPMN first</td>
<td>answer to Placebo</td>
</tr>
</tbody>
</table>

Assuming normality

Paired t-test

data:  AO_v1 and AO_v2

t = 1.0733, df = 18, \( p\)-value = 0.1487

alternative hypothesis: true difference in means is greater than 0

95 percent confidence interval: -0.2592077 Inf

sample estimates: mean of the differences 0.4210526

Not assuming normality

Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2

V = 69, \( p\)-value = 0.1540

alternative hypothesis: true location shift is greater than 0

Warning messages: 1: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, : cannot compute exact \( p\)-value with ties

2: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, : cannot compute exact \( p\)-value with zeroes

4.1.1.3. Formal results

Students from the general university population did not show statistically significant increases in their understanding of the workflows after reading a BPMN set compared to their initial understanding using only background knowledge of the domain.

4.1.2. USYD Students, Cockburn or Placebo

\( H₀ \) – USYD Students don’t comprehend Cockburn better than placebo – ☒ Rejected

\( H₁ \) – USYD student comprehend BPMN better than placebo – ☑ Accepted
4.1.2.1. USYD Results - First notation Cockburn

Here we see more progress. Cockburn score is better than placebo score.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>Sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>a  Score.Placebo</td>
<td>1.5400698</td>
<td>1.863636</td>
<td>2.1872029</td>
<td>0.7743172</td>
<td>22</td>
</tr>
<tr>
<td>b  Score.QSet1</td>
<td>2.2279580</td>
<td>2.681818</td>
<td>3.1356780</td>
<td>1.0861190</td>
<td>22</td>
</tr>
<tr>
<td>d  Delta.QSet1.Placebo</td>
<td>0.2921896</td>
<td><strong>0.818181</strong></td>
<td>1.3441740</td>
<td>1.2587357</td>
<td>22</td>
</tr>
</tbody>
</table>

*44 Here we see more progress. Cockburn score is better than placebo score.*
### 4.1.2.2. R Code

```r
## paired two sided T-Test
AO_ftTestP <- function(AO_v1, AO_v2)
  t.test(AO_v1, AO_v2, alternative="two.sided", paired=TRUE)
AO_ftTestP (subset(AO_lab, Source=="USYD" & First.Questionnaire=="Cockburn")$Score.QSet1, subset(AO_lab, Source=="USYD" & First.Questionnaire=="Cockburn")$Score.Placebo)

## paired one sided T-Test
AO_ftTestP <- function(AO_v1, AO_v2)
  t.test(AO_v1, AO_v2, alternative="greater", paired=TRUE)
AO_ftTestP (subset(AO_lab, Source=="USYD" & First.Questionnaire=="Cockburn")$Score.QSet1, subset(AO_lab, Source=="USYD" & First.Questionnaire=="Cockburn")$Score.Placebo)

## paired two sided wilcox
AO_fwilcoxTestP <- function(AO_v1, AO_v2)
  wilcox.test
  (AO_v1, AO_v2, alternative="two.sided", paired=TRUE, exact = TRUE)
AO_fwilcoxTestP (subset(AO_lab, Source=="USYD" & First.Questionnaire=="Cockburn")$Score.QSet1, subset(AO_lab, Source=="USYD" & First.Questionnaire=="Cockburn")$Score.Placebo)

## paired one sided wilcox
AO_fwilcoxTestP <- function(AO_v1, AO_v2)
  wilcox.test
  (AO_v1, AO_v2, alternative="greater", paired=TRUE, exact = TRUE)
AO_fwilcoxTestP (subset(AO_lab, Source=="USYD" & First.Questionnaire=="Cockburn")$Score.QSet1, subset(AO_lab, Source=="USYD" & First.Questionnaire=="Cockburn")$Score.Placebo)

AO_fCor <- function(AO_v1, AO_v2)
  cor (AO_v1, AO_v2, use = "everything", method= "spearman")
AO_fCor (subset(AO_lab, Source=="USYD" & First.Questionnaire=="Cockburn")$Score.QSet1, subset(AO_lab, Source=="USYD" & First.Questionnaire=="Cockburn")$Score.Placebo)
```

4.1.2.3. R Output

> ## paired two sided T-Test
> AO_ftTestP <- function(AO_v1,AO_v2){
t.test(AO_v1,AO_v2,alternative="two.sided",paired=TRUE)
> AO_ftTestP (subset(AO_lab,Source="USYD" &
First.Questionnaire="Cockburn")$Score.QSet1,subset(AO_lab,Source="USYD"
&First.Questionnaire="Cockburn")$Score.Placebo)

Paired t-test

data:  AO_v1 and AO_v2
t = 3.0488, df = 21, p-value = 0.006101
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.2600897 1.3762740
sample estimates:
mean of the differences
 0.8181818

> ## paired one sided T-Test
> AO_ftTestP <- function(AO_v1,AO_v2){
t.test(AO_v1,AO_v2,alternative="greater",paired=TRUE)
> AO_ftTestP (subset(AO_lab,Source="USYD" &
First.Questionnaire="Cockburn")$Score.QSet1,subset(AO_lab,Source="USYD"
&First.Questionnaire="Cockburn")$Score.Placebo)

Paired t-test

data:  AO_v1 and AO_v2
t = 3.0488, df = 21, p-value = 0.003050
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.3563975       Inf
sample estimates:
mean of the differences
 0.8181818

> ## paired two sided wilcoxon
> AO_fwilcoxTestP <- function(AO_v1,AO_v2){ wilcox.test
(AO_v1,AO_v2,alternative="two.sided",paired=TRUE, exact = TRUE) }
> AO_fwilcoxTestP (subset(AO_lab,Source="USYD" &
First.Questionnaire="Cockburn")$Score.QSet1,subset(AO_lab,Source="USYD"
&First.Questionnaire="Cockburn")$Score.Placebo)

Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 157, p-value = 0.01086
alternative hypothesis: true location shift is not equal to 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", :
cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", : 
cannot compute exact p-value with zeroes
> ## paired one sided wilcox
> AO_fwilcoxTestP <- function(AO_v1,AO_v2){ wilcox.test
(AO_v1,AO_v2,alternative="greater",paired=TRUE, exact = TRUE) }
> AO_fwilcoxTestP (subset(AO_lab,Source=="USYD" &
First.Questionnaire=="Cockburn")$Score.QSet1,subset(AO_lab,Source=="USYD" &
First.Questionnaire=="Cockburn")$Score.Placebo)

Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 157, p-value = 0.005432
alternative hypothesis: true location shift is greater than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired =
TRUE, : 
cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired =
TRUE, : 
cannot compute exact p-value with zeroes
>
Table 10 – $H_1$ – USYD student comprehend Cockburn better than placebo – $\checkmark$ Accepted ($p = 0.005432$)

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1: USYD students who got Cockburn first</td>
<td>answer to QSet1</td>
</tr>
<tr>
<td>Set 2: USYD students who got Cockburn first</td>
<td>answer to Placebo</td>
</tr>
</tbody>
</table>

assuming normality

- Paired t-test (sample size less than 30)
  - $t = 3.0488$, $df = 21$, $p-value = 0.003050$  
  - alternative hypothesis: true difference in means is greater than 0  
  - 95 percent confidence interval: $0.3563975$ Inf  
  - sample estimates: mean of the differences $0.8181818$  

not assuming normality

- Wilcoxon signed rank test with continuity correction  
  - $V = 157$, $p-value = 0.005432$  
  - alternative hypothesis: true location shift is greater than 0  
  - Warning messages:
    1: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, : cannot compute exact p-value with ties
    2: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, : cannot compute exact p-value with zeroes

4.1.2.4. Formal results

Students from the general university population, showed statistically significant increases in their understanding of the workflows ($P = 0.003$) after reading a written use case set compared to their understanding using only background knowledge of the domain.

USYD student comprehend Cockburn better than placebo – the two sided paired T-Test P Value for was $0.006101$; the one sided paired T-Test that showed that QSet1 $>$ Placebo had a P value of $0.003050$. However, knowing that sample size does not lend itself to a normality assumption, we further confirmed the differences using a non-parametric test, namely the Wilcoxon matched-pairs sign-rank test. The one two paired Wilcoxon matched-pairs sign-rank test had a P value of $0.01086$; the one sided paired Wilcoxon matchedpairs sign-rank test had a P value of $0.005432$. 
4.1.3. TU/e Students, BPMN or Placebo

H₀ – TU/e Students don’t comprehend BPMN better than placebo – ☑ Rejected

H₁ – TU/e Students comprehend BPMN better than placebo – ☑ Accepted

4.1.3.1. TU/e Results - First notation BPMN

subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.Placebo

A Score.Placebo 1.7108980 1.9594590 2.2080210 1.0909240 74
B Score.QSet1 3.1038910 3.4054050 3.7069200 1.3233300 74

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>Sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score.Placebo</td>
<td>1.7108980</td>
<td>1.9594590</td>
<td>2.2080210</td>
<td>1.0909240</td>
<td>74</td>
</tr>
<tr>
<td>Score.QSet1</td>
<td>3.1038910</td>
<td>3.4054050</td>
<td>3.7069200</td>
<td>1.3233300</td>
<td>74</td>
</tr>
</tbody>
</table>
4.1.3.2. R Code

z.test requires the BSDA package which need to be downloaded from [http://cran.r-project.org/web/packages/BSDA/index.html](http://cran.r-project.org/web/packages/BSDA/index.html) and installed using the R GUI:

Packages → Install package from local zip file Packages → Load Package …

```r
## Paired two sided Z-Test
AO_fzTestP <- function(AO_v1, AO_v2) {
  z.test(AO_v1, AO_v2, alternative = "two.sided", mu = 0, sigma.x = sd(AO_v1), sigma.y = sd(AO_v2), conf.level = 0.95)
}
AO_fzTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.QSet1, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.Placebo)

## Paired one sided Z-Test
AO_fzTestP <- function(AO_v1, AO_v2) {
  z.test(AO_v1, AO_v2, alternative = "greater", mu = 0, sigma.x = sd(AO_v1), sigma.y = sd(AO_v2), conf.level = 0.95)
}
AO_fzTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.QSet1, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.Placebo)

## paired two sided wilcox
AO_fwilcoxTestP <- function(AO_v1, AO_v2) {
  wilcox.test(AO_v1, AO_v2, alternative="two.sided", paired=TRUE, exact = TRUE)
}
AO_fwilcoxTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.QSet1, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.Placebo)

## paired one sided wilcox
AO_fwilcoxTestP <- function(AO_v1, AO_v2) {
  wilcox.test(AO_v1, AO_v2, alternative="greater", paired=TRUE, exact = TRUE)
}
AO_fwilcoxTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.QSet1, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.Placebo)
```
4.1.3.3. R Results

## Paired two sided Z-Test

```r
dontr普
AO_fzTestP <- function(AO_v1, AO_v2){
  z.test(AO_v1, AO_v2, alternative = "two.sided", mu = 0, sigma.x = sd(AO_v1), sigma.y = sd(AO_v2), conf.level = 0.95)
}
AO_fzTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.QSet1, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.Placebo)
```

Two-sample z-Test

data:  AO_v1 and AO_v2
z = 7.2527, p-value = 4.087e-13
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  1.055192 1.836700
sample estimates:
mean of x  mean of y
  3.405405  1.959459

## Paired one sided Z-Test

```r
dontr普
AO_fzTestP <- function(AO_v1, AO_v2){
  z.test(AO_v1, AO_v2, alternative = "greater", mu = 0, sigma.x = sd(AO_v1), sigma.y = sd(AO_v2), conf.level = 0.95)
}
AO_fzTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.QSet1, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.Placebo)
```

Two-sample z-Test

data:  AO_v1 and AO_v2
z = 7.2527, p-value = 2.044e-13
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
  1.118015       NA
sample estimates:
mean of x  mean of y
  3.405405  1.959459

## paired two sided wilcox

```r
dontr普
AO_fwilcoxTestP <- function(AO_v1, AO_v2){
  wilcox.test(AO_v1, AO_v2, alternative="two.sided",paired=TRUE, exact = TRUE)
}
AO_fwilcoxTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.QSet1, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$Score.Placebo)
```

Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 1717, p-value = 2.296e-09
alternative hypothesis: true location shift is not equal to 0
Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", :
   cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", :
   cannot compute exact p-value with zeroes

> ## paired one sided wilcox
> AO_fwilcoxTestP  <- function(AO_v1,AO_v2){ wilcox.test
  (AO_v1,AO_v2,alternative="greater",paired=TRUE, exact = TRUE) }
> AO_fwilcoxTestP  (subset(AO_lab,Source=="TU/e" &
  First.Questionnaire=="BPMN")$Score.QSet1,subset(AO_lab,Source=="TU/e"
  &First.Questionnaire=="BPMN")$Score.Placebo)

 Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 1717, p-value = 1.148e-09
alternative hypothesis: true location shift is greater than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired =
   TRUE, :
   cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired =
   TRUE, :
   cannot compute exact p-value with zeroes

>
Table 11 – H₁ – TU/e Students comprehend BPMN better than placebo – ✅ Accepted (P = 1.148e-09)

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Set 1:</strong> TU/e students who got BPMN first</td>
<td>answer to QSet1</td>
</tr>
<tr>
<td><strong>Set 2:</strong> TU/e students who got BPMN first</td>
<td>answer to Placebo</td>
</tr>
</tbody>
</table>

assuming normality

Two-sample z-Test
- data: AO_v1 and AO_v2
- z = 7.2527, p-value = 2.044e-13
- alternative hypothesis: true difference in means is greater than 0
- 95 percent confidence interval: 1.118015       NA
- sample estimates: mean of x mean of y  3.405405  1.959459

not assuming normality

Wilcoxon signed rank test with continuity correction
- data: AO_v1 and AO_v2
- V = 1717, p-value = 1.148e-09
- alternative hypothesis: true location shift is greater than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE,  : cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE,  : cannot compute exact p-value with zeroes

4.1.3.4. Formal results

Students who study business process management, showed statistically significant increases in their understanding of the workflows (P = 2.0 e-13) after reading a BPMN set compared to their understanding using only background knowledge of the domain.
4.1.4. TU/e Students, Cockburn or Placebo

H$_0$ – TU/e Students don’t comprehend Cockburn better than placebo - ✗ Rejected

H$_1$ – TU/e Students comprehend Cockburn better than placebo - ✓ Accepted

4.1.4.1. TU/e Results - First notation Cockburn

subset(AO_lab, Source=="TU/e" 
& First.Questionnaire=="Cockburn")$Score.Placebo

1 2 4 3 2 3 2 4 1 2 2 1 0 2 3 1 1 3 0 3 1 0 1 1 4 2 1 0 4 2 2 2 1 1 1 0 4
1 1 2 1 1 2 1 2 3 2 2 2 3 3 1 2 1

subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$Score.QSet1

2 3 4 3 1 2 2 5 2 5 4 3 4 6 2 2 3 1 0 1 2 2 1 4 3 2 3 3 2 2 1 1 4 4 2 0 0 4
5 4 3 4 3 3 3 0 4 3 2 3 3 4 4 3 4

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Score.Placebo</td>
<td>1.4910100</td>
<td>1.781818</td>
<td>2.0726260</td>
<td>1.100352</td>
<td>55</td>
</tr>
<tr>
<td>B Score.QSet1</td>
<td>2.3624310</td>
<td>2.727273</td>
<td>3.092150</td>
<td>1.380480</td>
<td>55</td>
</tr>
<tr>
<td>D Delta.QSet1.Placebo</td>
<td>0.5126252</td>
<td>0.9454545</td>
<td>1.3782838</td>
<td>1.6377286</td>
<td>55</td>
</tr>
</tbody>
</table>
4.1.4.2. R Code

z.test requires the BSDA package which need to be downloaded from http://cran.r-project.org/web/packages/BSDA/index.html and installed using the R GUI:
Packages → Install package from local zip file Packages → Load Package ...

```r
## Paired two sided Z-Test
AO_fzTestP <- function(AO_v1, AO_v2){
  z.test(AO_v1, AO_v2, alternative = "two.sided", mu = 0, sigma.x = sd(AO_v1), sigma.y = sd(AO_v2), conf.level = 0.95)
}

AO_fzTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$Score.QSet1, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$Score.Placebo)

## Paired one sided Z-Test
AO_fzTestP <- function(AO_v1, AO_v2){
  z.test(AO_v1, AO_v2, alternative = "greater", mu = 0, sigma.x = sd(AO_v1), sigma.y = sd(AO_v2), conf.level = 0.95)
}

AO_fzTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$Score.QSet1, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$Score.Placebo)

## paired two sided wilcox
AO_fwilcoxTestP <- function(AO_v1, AO_v2){
  wilcox.test(AO_v1, AO_v2, alternative="two.sided",paired=TRUE, exact = TRUE)
}

AO_fwilcoxTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$Score.QSet1, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$Score.Placebo)

## paired one sided wilcox
AO_fwilcoxTestP <- function(AO_v1, AO_v2){
  wilcox.test(AO_v1, AO_v2, alternative="greater",paired=TRUE, exact = TRUE)
}

AO_fwilcoxTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$Score.QSet1, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$Score.Placebo)
```
4.1.4.3. R Test Log

```r
# Paired two sided Z-Test

AO_fzTestP <- function(AO_v1, AO_v2) {
  z.test(AO_v1, AO_v2, alternative = "two.sided", mu = 0, sigma.x = sd(AO_v1),
          sigma.y = sd(AO_v2), conf.level = 0.95)
}

AO_fzTestP (subset(AO_lab, Source == "TU/e" & First.Questionnaire == "Cockburn")$Score.QSet1,
            subset(AO_lab, Source == "TU/e" & First.Questionnaire == "Cockburn")$Score.Placebo)

Two-sample z-Test
data:  AO_v1 and AO_v2
z = 3.9718, p-value = 7.133e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  0.4789024 1.4120067
sample estimates:
mean of x mean of y
  2.727273  1.781818
```

```r
# Paired one sided Z-Test

AO_fzTestP <- function(AO_v1, AO_v2) {
  z.test(AO_v1, AO_v2, alternative = "greater", mu = 0, sigma.x = sd(AO_v1),
          sigma.y = sd(AO_v2), conf.level = 0.95)
}

AO_fzTestP (subset(AO_lab, Source == "TU/e" & First.Questionnaire == "Cockburn")$Score.QSet1,
            subset(AO_lab, Source == "TU/e" & First.Questionnaire == "Cockburn")$Score.Placebo)

Two-sample z-Test
data:  AO_v1 and AO_v2
z = 3.9718, p-value = 3.566e-05
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
  0.5539116        NA
sample estimates:
mean of x mean of y
  2.727273  1.781818
```

```r
# Paired two sided wilcox

AO_fwilcoxTestP <- function(AO_v1, AO_v2) {
  wilcox.test(AO_v1, AO_v2, alternative="two.sided", paired=TRUE, exact = TRUE)
}
```

```r
```
Wilcoxon signed rank test with continuity correction

data: AO_v1 and AO_v2
V = 805, p-value = 0.0002413
alternative hypothesis: true location shift is not equal to 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", : cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", : cannot compute exact p-value with zeroes

Wilcoxon signed rank test with continuity correction

data: AO_v1 and AO_v2
V = 805, p-value = 0.0001207
alternative hypothesis: true location shift is greater than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, : cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, : cannot compute exact p-value with zeroes
Table 12 – H1 – TU/e Students comprehend Cockburn better than placebo - ✐ Accepted (P= 0.0001207)

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1: TU/e students who got Cockburn first</td>
<td>answer to QSet1</td>
</tr>
<tr>
<td>Set 2: TU/e students who got Cockburn first</td>
<td>answer to Placebo</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>assuming normality</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-sample z-Test</td>
<td>data: AO_v1 and AO_v2</td>
</tr>
<tr>
<td>z = 3.9718, <strong>p-value = 3.566e-05</strong></td>
<td>alternative hypothesis: true difference in means is greater than 0</td>
</tr>
<tr>
<td>95 percent confidence interval: 0.5539116–NA</td>
<td>sample estimates: mean of x mean of y 2.727273 1.781818</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>not assuming normality</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilcoxon signed rank test with continuity correction</td>
<td>data: AO_v1 and AO_v2 V = 805, <strong>p-value = 0.0001207</strong></td>
</tr>
<tr>
<td>alternative hypothesis: true location shift is greater than 0</td>
<td></td>
</tr>
<tr>
<td>Warning messages: 1: In wilcox.test.default(AO_v1, AO_v2, alternative = &quot;greater&quot;, paired = TRUE, ) : cannot compute exact p-value with ties 2: In wilcox.test.default(AO_v1, AO_v2, alternative = &quot;greater&quot;, paired = TRUE, ) : cannot compute exact p-value with zeroes</td>
<td></td>
</tr>
</tbody>
</table>

4.1.4.4. Formal results
Students who study business process management, showed statistically significant increases in their understanding of the workflows after reading a written use case set compared to their understanding using only background knowledge of the domain (**p-value = 0.0001207**).
4.1.5. **HU/b Students, BPMN or Placebo**

H₀ – HU/b Students don’t comprehend BPMN better than placebo - ✗ Rejected

H₁ – HU/b Students comprehend BPMN better than placebo - ✔ Accepted

### 4.1.5.1. HU/b - First notation BPMN Results

\( \text{subset(AO}\_\text{lab,Source=="HU/b" \&First.Questionnaire=="BPMN")}\$\text{Score.Placebo} \\
2 1 1 3 3 1 1 1 2 2 2 2

\( \text{subset(AO}\_\text{lab,Source=="HU/b" \&First.Questionnaire=="BPMN")}\$\text{Score.QSet1} \\
3 4 4 3 2 2 2 4 4 4 2 1 2

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>mean</th>
<th>Conf 0.05 to</th>
<th>Sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>a Score.Placebo</td>
<td>1.37511000</td>
<td>1.76923100</td>
<td>2.16335100</td>
<td>0.72501100</td>
<td>13</td>
</tr>
<tr>
<td>b Score.QSet1</td>
<td>2.26548000</td>
<td>2.84615400</td>
<td>3.42682700</td>
<td>1.06818800</td>
<td>13</td>
</tr>
<tr>
<td>d Delta.QSet1.Placebo</td>
<td>0.29350320</td>
<td>1.07692310</td>
<td>1.86034300</td>
<td>1.44115340</td>
<td>13</td>
</tr>
</tbody>
</table>
4.1.5.2. R Code

## -----------------------
## paired two sided T-Test
## -----------------------

AO_ftTestP <- function(AO_v1, AO_v2) {
  t.test(AO_v1, AO_v2, alternative="two.sided", paired=TRUE)
}

AO_ftTestP(subset(AO_lab, Source=="HU/b" & First.Questionnaire=="BPMN")$Score.QSet1, subset(AO_lab, Source=="HU/b" & First.Questionnaire=="BPMN")$Score.Placebo)

## paired one sided T-Test

AO_ftTestP <- function(AO_v1, AO_v2) {
  t.test(AO_v1, AO_v2, alternative="greater", paired=TRUE)
}

AO_ftTestP(subset(AO_lab, Source=="HU/b" & First.Questionnaire=="BPMN")$Score.QSet1, subset(AO_lab, Source=="HU/b" & First.Questionnaire=="BPMN")$Score.Placebo)

## paired two sided wilcox

AO_fwilcoxTestP <- function(AO_v1, AO_v2) {
  wilcox.test(AO_v1, AO_v2, alternative="two.sided", paired=TRUE, exact = TRUE)
}

AO_fwilcoxTestP(subset(AO_lab, Source=="HU/b" & First.Questionnaire=="BPMN")$Score.QSet1, subset(AO_lab, Source=="HU/b" & First.Questionnaire=="BPMN")$Score.Placebo)

## paired one sided wilcox

AO_fwilcoxTestP <- function(AO_v1, AO_v2) {
  wilcox.test(AO_v1, AO_v2, alternative="greater", paired=TRUE, exact = TRUE)
}

AO_fwilcoxTestP(subset(AO_lab, Source=="HU/b" & First.Questionnaire=="BPMN")$Score.QSet1, subset(AO_lab, Source=="HU/b" & First.Questionnaire=="BPMN")$Score.Placebo)
4.1.5.3. R Log

```r
> ## -----------------------
> ## paired two sided T-Test
> ## -----------------------
> AO_ftTestP <- function(AO_v1,AO_v2){
  t.test(AO_v1,AO_v2,alternative="two.sided",paired=TRUE)
} 
> AO_ftTestP (subset(AO_lab,Source=="HU/b" &
First.Questionnaire=="BPMN")$Score.QSet1,subset(AO_lab,Source=="HU/b" &
&First.Questionnaire=="BPMN")$Score.Placebo)

Paired t-test

data:  AO_v1 and AO_v2
t = 2.6943, df = 12, p-value = 0.01951
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  0.2060428 1.9478034
sample estimates:
  mean of the differences
    1.076923

> ## paired one sided T-Test
> AO_ftTestP <- function(AO_v1,AO_v2){
  t.test(AO_v1,AO_v2,alternative="greater",paired=TRUE)
} 
> AO_ftTestP (subset(AO_lab,Source=="HU/b" &
First.Questionnaire=="BPMN")$Score.QSet1,subset(AO_lab,Source=="HU/b" &
&First.Questionnaire=="BPMN")$Score.Placebo)

Paired t-test

data:  AO_v1 and AO_v2
t = 2.6943, df = 12, p-value = 0.009757
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
  0.3645356       Inf
sample estimates:
  mean of the differences
    1.076923

> ## paired two sided wilcox
> AO_fwilcoxTestP <- function(AO_v1,AO_v2) { wilcox.test
(AO_v1,AO_v2,alternative="two.sided",paired=TRUE, exact = TRUE) }
> AO_fwilcoxTestP (subset(AO_lab,Source=="HU/b" &
First.Questionnaire=="BPMN")$Score.QSet1,subset(AO_lab,Source=="HU/b" &
&First.Questionnaire=="BPMN")$Score.Placebo)
```
Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 49, p-value = 0.02955
alternative hypothesis: true location shift is not equal to 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", : cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", : cannot compute exact p-value with zeroes

> ## paired one sided wilcox
>
> AO_fwilcoxTestP  <- function(AO_v1,AO_v2){ wilcox.test(AO_v1,AO_v2,alternative="greater",paired=TRUE, exact = TRUE) }
> AO_fwilcoxTestP  (subset(AO_lab,Source=="HU/b" & First.Questionnaire=="BPMN")$Score.QSet1,subset(AO_lab,Source=="HU/b" &First.Questionnaire=="BPMN")$Score.Placebo)

Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 49, p-value = 0.01477
alternative hypothesis: true location shift is greater than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, : cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, : cannot compute exact p-value with zeroes

>
Students comprehend BPMN better than placebo -Accepted (p = 0.01477)

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1: HU/b students who got BPMN first</td>
<td>answer to QSet1</td>
</tr>
<tr>
<td>Set 2: HU/b students who got BPMN first</td>
<td>answer to Placebo</td>
</tr>
</tbody>
</table>

assuming normality

One Sample t-test

data: AO_v1
t = 3.635, df = 12, p-value = 0.001710
alternative hypothesis: true mean is greater than 1.769231
95 percent confidence interval: 2.318130 Inf
sample estimates: mean of x 2.846154

not assuming normality

data: AO_v1 and AO_v2 V = 49, p-value = 0.01477
alternative hypothesis: true location shift is greater than 0
Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, :cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, :cannot compute exact p-value with zeroes

4.1.5.4. Formal results

Students who study business process management, showed statistically significant increases in their understanding of the workflows after reading a BPMN set compared to their understanding using only background knowledge of the domain p-value = 0.01477.

4.1.6. HU/b Students, Cockburn or Placebo

H₀ – HU/b Students don’t comprehend Cockburn better than placebo – Rejected

H₁ – HU/b Students comprehend Cockburn better than placebo – Accepted

4.1.6.1. HU/b - First notation Cockburn Results

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>Sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Score.Placebo</td>
<td>1.00605800</td>
<td>1.61538500</td>
<td>2.22471100</td>
<td>1.12089700</td>
</tr>
<tr>
<td>b</td>
<td>Score.QSet1</td>
<td>1.93430900</td>
<td>2.76923100</td>
<td>3.60415300</td>
<td>1.53589500</td>
</tr>
<tr>
<td>d</td>
<td>Delta.QSet1.Placebo</td>
<td>0.19041620</td>
<td>1.15384620</td>
<td>2.11727610</td>
<td>1.77229390</td>
</tr>
</tbody>
</table>
4.1.6.2. R Code

```r
## -----------------------  
## paired two sided T-Test  
## -----------------------  

AO_ftTestP <- function(AO_v1,AO_v2){  
t.test(AO_v1,AO_v2,alternative="two.sided",paired=TRUE)
AO_ftTestP (subset(AO_lab,Source=="HU/b" &  
First.Questionnaire="Cockburn")$Score.QSet1,subset(AO_lab,Source=="HU/b"  
&First.Questionnaire="Cockburn")$Score.Placebo)

## paired one sided T-Test

AO_ftTestP <- function(AO_v1,AO_v2){  
t.test(AO_v1,AO_v2,alternative="greater",paired=TRUE)
AO_ftTestP (subset(AO_lab,Source=="HU/b" &  
First.Questionnaire="Cockburn")$Score.QSet1,subset(AO_lab,Source=="HU/b"  
&First.Questionnaire="Cockburn")$Score.Placebo)

## paired two sided wilcox

AO_fwilcoxTestP <- function(AO_v1,AO_v2){ wilcox.test  
(AO_v1,AO_v2,alternative="two.sided",paired=TRUE, exact = TRUE) }
AO_fwilcoxTestP (subset(AO_lab,Source=="HU/b" &  
First.Questionnaire="Cockburn")$Score.QSet1,subset(AO_lab,Source=="HU/b"  
&First.Questionnaire="Cockburn")$Score.Placebo)

## paired one sided wilcox

AO_fwilcoxTestP <- function(AO_v1,AO_v2){ wilcox.test  
(AO_v1,AO_v2,alternative="greater",paired=TRUE, exact = TRUE) }
AO_fwilcoxTestP (subset(AO_lab,Source=="HU/b" &  
First.Questionnaire="Cockburn")$Score.QSet1,subset(AO_lab,Source=="HU/b"  
&First.Questionnaire="Cockburn")$Score.Placebo)
```
4.1.6.3. R Log

```r
> ## -----------------------
> ## paired two sided T-Test
> ## -----------------------
> AO_ftTestP <- function(AO_v1,AO_v2){
t.test(AO_v1,AO_v2,alternative="two.sided",paired=TRUE)
>
> AO_ftTestP (subset(AO_lab,Source=="HU/b" &
First.Questionnaire=="Cockburn")$Score.QSet1,subset(AO_lab,Source=="HU/b"
&First.Questionnaire=="Cockburn")$Score.Placebo)

Paired t-test
data:  AO_v1 and AO_v2
t = 2.3474, df = 12, p-value = 0.03689
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.08285967  2.22483263
sample estimates:
mean of the differences
  1.153846

> ## paired one sided T-Test
> AO_ftTestP <- function(AO_v1,AO_v2){
t.test(AO_v1,AO_v2,alternative="greater",paired=TRUE)
>
> AO_ftTestP (subset(AO_lab,Source=="HU/b" &
First.Questionnaire=="Cockburn")$Score.QSet1,subset(AO_lab,Source=="HU/b"
&First.Questionnaire=="Cockburn")$Score.Placebo)

Paired t-test
data:  AO_v1 and AO_v2
t = 2.3474, df = 12, p-value = 0.01844
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
  0.2777700         Inf
sample estimates:
mean of the differences
  1.153846

> ## paired two sided wilcox
> AO_fwilcoxTestP <- function(AO_v1,AO_v2){ wilcox.test
(AO_v1, AO_v2, alternative="two.sided", paired=TRUE, exact = TRUE )
>
> AO_fwilcoxTestP (subset(AO_lab,Source=="HU/b" &
First.Questionnaire=="Cockburn")$Score.QSet1,subset(AO_lab,Source=="HU/b"
&First.Questionnaire=="Cockburn")$Score.Placebo)

Wilcoxon signed rank test with continuity correction
data:  AO_v1 and AO_v2
V = 40, p-value = 0.0418
alternative hypothesis: true location shift is not equal to 0
```
Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", :
   cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", :
   cannot compute exact p-value with zeroes

> ## paired one sided wilcox
> AO_fwilcoxTestP  <- function(AO_v1,AO_v2){wilcox.test
(AO_v1, AO_v2,alternative="greater",paired=TRUE, exact = TRUE) }
> AO_fwilcoxTestP  (subset(AO_lab,Source=="HU/b" &
First.Questionnaire=="Cockburn")$Score.QSet1,subset(AO_lab,Source=="HU/b"
&First.Questionnaire=="Cockburn")$Score.Placebo)

Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 40, p-value = 0.0209
alternative hypothesis: true location shift is greater than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired =
   TRUE, :
   cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired =
   TRUE, :
   cannot compute exact p-value with zeroes

>
Table 14 - H$_1$ – HU/b Students comprehend Cockburn better than placebo – ✅ Accepted (P = 0.0209)

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Set 1:</strong> HU/b students who got Cockburn first</td>
<td>answer to QSet1</td>
</tr>
<tr>
<td><strong>Set 2:</strong> HU/b students who got Cockburn first</td>
<td>answer to Placebo</td>
</tr>
</tbody>
</table>

**not assuming normality**

Paired t-test  
- Data: AO_v1 and AO_v2  
- t = 2.3474, df = 12, \( p\text{-value} = 0.01844 \)  
- Alternative hypothesis: true difference in means is greater than 0  
- 95 percent confidence interval: 0.2777700 Inf  
- Sample estimates: mean of the differences 1.153846

**assuming normality**

Wilcoxon signed rank test with continuity correction  
- Data: AO_v1 and AO_v2 V = 40, \( p\text{-value} = 0.0209 \)  
- Alternative hypothesis: true location shift is greater than 0  
- Warning messages:
  1: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE,  : cannot compute exact p-value with ties  
  2: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE,  : cannot compute exact p-value with zeroes

4.1.6.4. Formal results

Students who study business process management, **showed** statistically significant increases in their understanding of the workflows (P = 0.018) after reading a written use case set compared to their understanding using only background knowledge of the domain.
4.2. Comparing comprehension of Cockburn and BPMN

4.2.1. USYD Students, Cockburn or BPMN ▲

H₀ – USYD Students don’t comprehend Cockburn Better than BPMN ✔ Not rejected

H₁ – USYD Students comprehend Cockburn Better than BPMN — ☒ Rejected

4.2.1.1. USYD Results - First notation Cockburn

subset(AO_lab, Source=="USYD" & First.Questionnaire=="Cockburn")$Score.Placebo
1 2 1 1 1 2 3 3 2 3 2 2 2 3 1 2 2 1 1 1 3 2

subset(AO_lab, Source=="USYD" & First.Questionnaire=="Cockburn")$Score.QSet1
3 3 3 3 3 1 2 4 5 2 2 2 4 5 2 3 2 2 2 2 1 3

subset(AO_lab, Source=="USYD" & First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo
2 1 2 2 2 -1 -1 1 3 -1 0 0 2 2 1 1 0 1 1 1 -2 1

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to Mean</th>
<th>sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score.Placebo</td>
<td>1.5400698</td>
<td>1.8636364</td>
<td>2.1872029</td>
<td>0.7743172</td>
<td>22</td>
</tr>
<tr>
<td>Score.QSet1</td>
<td>2.2279580</td>
<td>2.6818180</td>
<td>3.1356780</td>
<td>1.0861190</td>
<td>22</td>
</tr>
<tr>
<td>Delta.QSet1.Placebo</td>
<td>0.2921896</td>
<td>0.8181818</td>
<td>1.3441740</td>
<td>1.2587357</td>
<td>22</td>
</tr>
</tbody>
</table>

45 Here we see more progress. Cockburn score is better than placebo score.
4.2.1.1. USYD Results - First notation BPMN

subset(AO_lab, Source=="USYD" \& First.Questionnaire=="BPMN")$Score.Placebo
1 3 1 4 3 1 3 0 4 2 3 3 2 2 3 1 0 2 1
subset(AO_lab, Source=="USYD" \& First.Questionnaire=="BPMN")$Score.QSet1
2 2 1 1 2 4 1 2 5 3 1 3 2 3 3 0 4 3
subset(AO_lab, Source=="USYD" \& First.Questionnaire=="BPMN")$Delta.QSet1.Placebo
1 -1 0 -3 -1 3 -2 2 1 3 0 -2 1 0 0 2 0 2 2

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>mean</th>
<th>Conf 0.05 to</th>
<th>Sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>a Score.Placebo</td>
<td>1.50245600</td>
<td>2.05263200</td>
<td>2.60280700</td>
<td>1.22355100</td>
<td>19</td>
</tr>
<tr>
<td>b Score.QSet1</td>
<td>1.84903200</td>
<td>2.47368400</td>
<td>3.0983600</td>
<td>1.38918100</td>
<td>19</td>
</tr>
<tr>
<td>d Delta.QSet1.Placebo</td>
<td>-0.34784090</td>
<td><strong>0.42105260</strong> *</td>
<td>1.18994620</td>
<td>1.70996390</td>
<td>19</td>
</tr>
</tbody>
</table>

4.2.1.2. R Code

AO_ftTestNp<- function(AO_v1,AO_v2) {
  t.test(AO_v1,AO_v2,alternative="two.sided",paired=FALSE) }
AO_ftTestNp (subset(AO_lab,Source=="USYD" \& First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo,subset(AO_lab,Source=="USYD" \& First.Questionnaire=="BPMN")$Delta.QSet1.Placebo)
AO_fwilcoxTestP  <- function(AO_v1,AO_v2){ wilcox.test (AO_v1,AO_v2,alternative="two.sided",paired= FALSE, exact = TRUE) }
AO_fwilcoxTestP (subset(AO_lab,Source=="USYD" \& First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo,subset(AO_lab,Source=="USYD" \& First.Questionnaire=="BPMN")$Delta.QSet1.Placebo)

46 Here we see very little progress. BPMN score is similar to placebo score.
Table 15 – H₀ – USYD Students don’t comprehend Cockburn Better than BPMN  Not rejected

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Set 1:</strong> USYD students who got Cockburn first</td>
<td>answer to QSet1-Placebo</td>
</tr>
<tr>
<td><strong>Set 2:</strong> USYD students who got BPMN first</td>
<td>answer to QSet1-Placebo</td>
</tr>
</tbody>
</table>

**assuming normality**

Welch Two Sample t-test
data: AO_v1 and AO_v2
t = 0.8355, df = 32.659, p-value = 0.4095
alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: -0.5702649 1.3645233
sample estimates: mean of x mean of y 0.8181818 0.4210526

**not assuming normality**
data: AO_v1 and AO_v2
W = 236.5, p-value = 0.4705
alternative hypothesis: true location shift is not equal to 0
Warning message: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", :  cannot compute exact p-value with ties

4.2.1.3. Formal results
4.2.2. TU/e Students, BPMN or Cockburn

H₀ – TU/e Students don’t comprehend BPMN Better than Cockburn [Not rejected]

H₁ – TU/e Students comprehend BPMN Better than Cockburn — Rejected

4.2.2.1. TU/e Results - First notation BPMN

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>Sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>a Score.Placebo</td>
<td>1.7108980</td>
<td>1.9594590</td>
<td>2.2080210</td>
<td>1.0909240</td>
<td>74</td>
</tr>
<tr>
<td>b Score.QSet1</td>
<td>3.1038910</td>
<td>3.4054050</td>
<td>3.7069200</td>
<td>1.3233300</td>
<td>74</td>
</tr>
<tr>
<td>d Delta.QSet1.Placebo</td>
<td>1.1041430</td>
<td>1.4459460</td>
<td>1.7877490</td>
<td>1.5001540</td>
<td>74</td>
</tr>
</tbody>
</table>
4.2.2.1.1. TU/e Results - First notation Cockburn

```r
subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$Score.Placebo
1 2 4 3 2 3 2 4 1 2 2 1 0 2 3 1 1 3 0 3 1 0 1 1 4 2 1 0 4 2 2 2 1 1 1 1 1 0 4
1 1 2 1 1 1 2 1 2 3 2 2 2 3 3 1 2 1
subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$Score.QSet1
2 3 4 3 1 2 2 5 2 5 4 3 4 6 2 2 3 1 0 1 2 2 1 4 3 2 3 2 2 1 1 4 4 2 0 0 4
5 4 3 4 3 3 0 4 3 2 3 3 4 4 3 4
subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo
1 1 0 0 -1 -1 0 1 1 3 2 2 4 4 -1 1 2 -2 0 -2 1 2 0 3 -1
0 2 3 -2 0 -1 -1 3 3 1 -1 0 0 4 3 1 3 2 1 2 -2 1 1 0 1 0
1 3 1 3
```

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>Sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>a Score.Placebo</td>
<td>1.4910100</td>
<td>1.7818180</td>
<td>2.0726260</td>
<td>1.100352</td>
<td>55</td>
</tr>
<tr>
<td>b Score.QSet1</td>
<td>2.3624310</td>
<td>2.7272730</td>
<td>3.0921150</td>
<td>1.380480</td>
<td>55</td>
</tr>
<tr>
<td>d Delta.QSet1.Placebo</td>
<td>0.5126252</td>
<td>0.9454545</td>
<td>1.3782838</td>
<td>1.6377286</td>
<td>55</td>
</tr>
</tbody>
</table>
4.2.2.2. R Code

z.test requires the BSDA package which need to be downloaded from http://cran.r-project.org/web/packages/BSDA/index.html and installed using the R GUI:

Packages → Install package from local zip file Packages → Load Package ...

AO_fzTestP <- function(AO_v1,AO_v2){ z.test(AO_v1, AO_v2, alternative ="two.sided",mu = 0, sigma.x = sd(AO_v1), sigma.y = sd(AO_v2), conf.level = 0.95) }

AO_fzTestP (subset(AO_lab,Source=="TU/e" & First.Questionnaire="BPMN")$Delta.QSet1.Placebo,subset(AO_lab,Source=="TU/e" & First.Questionnaire="Cockburn")$Delta.QSet1.Placebo)

AO_fwilcoxTestP <- function(AO_v1,AO_v2){ wilcox.test(AO_v1,AO_v2,alternative="two.sided",paired= FALSE, exact = TRUE) }

AO_fwilcoxTestP (subset(AO_lab,Source=="TU/e" & First.Questionnaire="BPMN")$Delta.QSet1.Placebo,subset(AO_lab,Source=="TU/e" & First.Questionnaire="Cockburn")$Delta.QSet1.Placebo)

Table 16 – H₀ – TU/e Students don’t comprehend BPMN Better than Cockburn ☒ Not rejected

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1: TU/e students who got Cockburn first</td>
<td>answer to QSet2-Placebo</td>
</tr>
<tr>
<td>Set 2: TU/e students who got BPMN first</td>
<td>answer to QSet2-Placebo</td>
</tr>
</tbody>
</table>

assuming normality

Two-sample z-Test data: AO_v1 and AO_v2

\[
z = 1.7787, \quad p\text{-value} = 0.0753
\]

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval: -0.05101513 1.05199793

sample estimates: mean of x mean of y 1.4459459 0.9454545

not assuming normality

Wilcoxon rank sum test with continuity correction data: AO_v1 and AO_v2

\[
W = 2386.5, \quad p\text{-value} = 0.0888
\]

alternative hypothesis: true location shift is not equal to 0

Warning message: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", : cannot compute exact p-value with ties

4.2.2.3. Formal results

Compensating for domain knowledge, students who study business process management, did not show statistically significant better understanding of a workflows articulated as a BPMN set compared to their understanding of identical workflow articulated as written use case.

4.2.3. HU/b Students, BPMN or Cockburn ↑

H₀ – HU/b Students don’t comprehend BPMN Better than Cockburn ☒ Not rejected

H₁ – HU/b Students comprehend BPMN Better than Cockburn – ☐ Rejected

4.2.3.1. HU/b Results - First notation BPMN

subset(AO_lab,Source="HU/b" & First.Questionnaire="BPMN")$Score.Placebo
2 1 1 3 3 1 1 1 2 2 2 2 2 subset(AO_lab,Source="HU/b" & First.Questionnaire="BPMN")$Score.QSet1
2 1 1 3 3 1 1 1 2 2 2 2 2
Here we see that administering Cockburn second makes things worst even in Berlin
4.2.3.2. HU/b Results - First notation Cockburn

subset(AO_lab, Source=="HU/b" & First.Questionnaire=="Cockburn")$Score.Placebo
1 2 3 0 1 1 1 1 1 3 4 1 2
subset(AO_lab, Source=="HU/b" & First.Questionnaire=="Cockburn")$Score.QSet1
4 4 2 3 0 6 3 2 2 3 4 1 2
subset(AO_lab, Source=="HU/b" & First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo
3 2 -1 3 -1 5 2 1 1 0 0 0 0

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>a Score.Placebo</td>
<td>1.00605800</td>
<td>1.61538500</td>
<td>2.22471100</td>
<td>1.12089700</td>
<td>13</td>
</tr>
<tr>
<td>b Score.QSet1</td>
<td>1.93430900</td>
<td>2.76923100</td>
<td>3.60415300</td>
<td>1.53589500</td>
<td>13</td>
</tr>
<tr>
<td>d Delta.QSet1.Placebo</td>
<td>0.19041620</td>
<td>1.15384620</td>
<td>2.11727610</td>
<td>1.77229390</td>
<td>13</td>
</tr>
</tbody>
</table>
4.2.3.3. R Code

AO_ftTestNp <- function(AO_v1, AO_v2) {
  t.test(AO_v1, AO_v2, alternative="two.sided", paired=FALSE)
}

AO_ftTestNp(subset(AO_lab, Source=="HU/b" & First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo, subset(AO_lab, Source=="HU/b" & First.Questionnaire=="BPMN")$Delta.QSet1.Placebo)

AO_fwilcoxTestP <- function(AO_v1, AO_v2) {
  wilcox.test(AO_v1, AO_v2, alternative="two.sided", paired=FALSE), exact=TRUE
}

AO_fwilcoxTestP(subset(AO_lab, Source=="HU/b", First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo, subset(AO_lab, Source=="HU/b", First.Questionnaire=="BPMN")$Delta.QSet1.Placebo)

Table 17 – H₀ – HU/b Students don’t comprehend BPMN Better than Cockburn ☑ Not rejected

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Set 1:</strong> HU/b students who got Cockburn first</td>
<td>answer to QSet2-Placebo</td>
</tr>
<tr>
<td><strong>Set 2:</strong> HU/b students who got BPMN first</td>
<td>answer to QSet2-Placebo</td>
</tr>
</tbody>
</table>

assuming normality

Welch Two Sample t-test
data:  AO_v1 and AO_v2
t = 0.1214, df = 23.042, p-value = 0.9044
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-1.233536 1.387382
sample estimates: mean of x mean of y 1.153846 1.076923

not assuming normality

Wilcoxon rank sum test with continuity correction
data:  AO_v1 and AO_v2
W = 83, p-value = 0.9583
alternative hypothesis: true location shift is not equal to 0
Warning message: In wilcox.test.default(AO_v1, AO_v2, alternative="two.sided", : cannot compute exact p-value with ties

4.2.3.4. Formal results

Compensating for domain knowledge, students who study business process management, did not show statistically significant better understanding of a workflows articulated as a BPMN set compared to their understanding of identical workflow articulated as written use case.
4.3. Comparing the contribution of the 2nd design

4.3.1. USYD Students, Cockburn and then BPMN

H₀ – Presenting BPMN to USYD students who already read Cockburn doesn’t improve comprehension – **Rejected**

H₁ – Presenting BPMN to USYD students who already read Cockburn improves comprehension – **Accepted**

4.3.1.1. USYD Results - First notation Cockburn

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Score.Placebo</td>
<td>1.5400698</td>
<td>1.8636364</td>
<td>2.1872029</td>
</tr>
<tr>
<td>b</td>
<td>Score.QSet1</td>
<td>2.2279580</td>
<td>2.6818180</td>
<td>3.1356780</td>
</tr>
<tr>
<td>c</td>
<td>Score.QSet2</td>
<td>3.1869760</td>
<td>3.5455550</td>
<td>3.9039330</td>
</tr>
<tr>
<td>d</td>
<td>Delta.QSet1.Placebo</td>
<td>0.2921896</td>
<td>0.8181818</td>
<td><strong>48</strong></td>
</tr>
<tr>
<td>e</td>
<td>Delta.QSet2.q1</td>
<td>0.5153181</td>
<td>0.8636364</td>
<td>1.2119546</td>
</tr>
<tr>
<td>f</td>
<td>Delta.QSet2.Placebo</td>
<td>1.2279580</td>
<td><strong>1.6818180</strong></td>
<td><strong>49</strong></td>
</tr>
</tbody>
</table>

---

**48** Here we see more progress. Cockburn score is better than placebo score.

**49** Here we see the highest progress. Cockburn first and then BPMN.
4.3.1.2. R Code

## -----------------------
## paired two sided T-Test
## ...
subset(AO_lab,Source=='USYD'
&First.Questionnaire=='Cockburn')$Delta.QSet2.Placebo ,
subset(AO_lab,Source=='USYD'
&First.Questionnaire=='Cockburn')$Delta.QSet1.Placebo)

## paired one sided T-Test
AO_ftTestP  <- function(AO_v1,AO_v2){
t.test(AO_v1,AO_v2,alternative='greater',paired=TRUE)}
AO_ftTestP  (subset(AO_lab,Source=='USYD'
&First.Questionnaire=='Cockburn')$Delta.QSet2.Placebo ,
subset(AO_lab,Source=='USYD'
&First.Questionnaire=='Cockburn')$Delta.QSet1.Placebo)

## paired two sided wilcoxon
AO_fwilcoxTestP  <- function(AO_v1,AO_v2){
wilcox.test
(AO_v1,AO_v2,alternative='two.sided',paired= TRUE, exact=TRUE) }
AO_fwilcoxTestP  (subset(AO_lab,Source=='USYD'
&First.Questionnaire=='Cockburn')$Delta.QSet2.Placebo ,
subset(AO_lab,Source=='USYD'
&First.Questionnaire=='Cockburn')$Delta.QSet1.Placebo)

## paired one sided wilcoxon
AO_fwilcoxTestP  <- function(AO_v1,AO_v2){
wilcox.test
(AO_v1,AO_v2,alternative='greater',paired= TRUE, exact=TRUE) }
AO_fwilcoxTestP  (subset(AO_lab,Source=='USYD'
&First.Questionnaire=='Cockburn')$Delta.QSet2.Placebo ,
subset(AO_lab,Source=='USYD'
&First.Questionnaire=='Cockburn')$Delta.QSet1.Placebo)
R Log

> ## -----------------------
> ## paired two sided T-Test
> ## -----------------------
> AO_ftTestP <- function(AO_v1,AO_v2){
  t.test(AO_v1,AO_v2,alternative="two.sided",paired=TRUE)
>
> AO_ftTestP (subset(AO_lab,Source=="USYD" &First.Questionnaire="Cockburn")$Delta.QSet2.Placebo ,
subset(AO_lab,Source=="USYD" &First.Questionnaire="Cockburn")$Delta.QSet1.Placebo)

  Paired t-test

data:  AO_v1 and AO_v2
t = 4.8597, df = 21, p-value = 8.353e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  0.4940611 1.2332116
sample estimates:
  mean of the differences
  0.8636364
>
> ## paired one sided T-Test
>
> AO_ftTestP <- function(AO_v1,AO_v2){
  t.test(AO_v1,AO_v2,alternative="greater",paired=TRUE)
>
> AO_ftTestP (subset(AO_lab,Source=="USYD" &First.Questionnaire="Cockburn")$Delta.QSet2.Placebo ,
subset(AO_lab,Source=="USYD" &First.Questionnaire="Cockburn")$Delta.QSet1.Placebo)

  Paired t-test

data:  AO_v1 and AO_v2
t = 4.8597, df = 21, p-value = 4.177e-05
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
  0.5578373       Inf
sample estimates:
  mean of the differences
  0.8636364
>
> ## paired two sided wilcoxon
>
> AO_fwilcoxTestP <- function(AO_v1,AO_v2){
  wilcox.test(AO_v1,AO_v2,alternative="two.sided",paired= TRUE, exact=TRUE) }

  Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
W = 75, n = 22
p-value = 8.353e-05
alternative hypothesis: true median of differences is not equal to 0
95 percent confidence interval:
  0.3457206 1.4152012
sample estimates:
  median of the differences
  0.8636364

>
Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 145.5, p-value = 0.0005166
alternative hypothesis: true location shift is not equal to 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided",  :
cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided",  :
cannot compute exact p-value with zeroes

Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 145.5, p-value = 0.0002583
alternative hypothesis: true location shift is greater than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, exact=TRUE)  :
cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE,  :
cannot compute exact p-value with zeroes
Table 18 – $H_1$ – Presenting BPMN to USYD students who already read Cockburn improves comprehension – ☑ Accepted ($P = 0.0005166$)

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Set 1:</strong> USYD students who got Cockburn first</td>
<td>answer to QSet2</td>
</tr>
<tr>
<td><strong>Set 2:</strong> USYD students who got Cockburn first</td>
<td>answer to QSet1</td>
</tr>
</tbody>
</table>

assuming normality
- Paired t-test data: AO_v1 and AO_v2
  - $t = 4.8597$, df = 21, $p$-value = $8.353 \times 10^{-5}$
  - alternative hypothesis: true difference in means is not equal to 0
  - 95 percent confidence interval: 0.4940611 1.2332116
  - sample estimates: mean of the differences 0.8636364

assuming normality
- Wilcoxon signed rank test with continuity correction data: AO_v1 and AO_v2
  - $V = 145.5$, $p$-value = $0.0005166$
  - alternative hypothesis: true location shift is not equal to 0
  - Warning messages:
    1: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", : cannot compute exact $p$-value with ties
    2: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", : cannot compute exact $p$-value with zeroes

4.3.1.3. Formal results
Compensating for domain knowledge, students from the general university population showed significantly better understanding of a workflows articulated as a written use case set compared to their initial understanding of identical workflow articulated as a set of BPMN diagrams ($p=0.0005166$).

4.3.2. USYD Students, BPMN and then Cockburn

$H_0$ – Presenting Cockburn to USYD students who already read BPMN doesn’t improve comprehension – ☑ Not rejected

$H_0$ – Presenting Cockburn to USYD students who already read BPMN improves comprehension – ☑ Rejected
### 4.3.2.1. USYD Results - First notation BPMN

\[
\text{subset(AO.lab, Source=="USYD" & First.Questionnaire=="BPMN")$Delta.QSet1.Placebo}
\]

\[
1 -1 0 -3 -1 3 -2 2 1 3 0 -2 1 0 0 2 0 2 2 2
\]

\[
\text{subset(AO.lab, Source=="USYD" & First.Questionnaire=="BPMN")$Delta.QSet2.Placebo}
\]

\[
1 -2 2 -1 -1 1 -1 3 1 0 0 -1 1 -2 0 2 2 1 3
\]

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>mean</th>
<th>Conf 0.05 to</th>
<th>Sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Score.Placebo</td>
<td>1.50245600</td>
<td>2.05263200</td>
<td>2.60280700</td>
<td>1.22355100</td>
<td>19</td>
</tr>
<tr>
<td>B Score.QSet1</td>
<td>1.84903200</td>
<td>2.47368400</td>
<td>3.09833600</td>
<td>1.38918100</td>
<td>19</td>
</tr>
<tr>
<td>C Score.QSet2</td>
<td>2.04368600</td>
<td>2.52631600</td>
<td>3.00894600</td>
<td>1.07333400</td>
<td>19</td>
</tr>
<tr>
<td>D Delta.QSet1.Placebo</td>
<td>-0.34784090</td>
<td><strong>0.42105260</strong></td>
<td>0.42105260</td>
<td>1.18994620</td>
<td>19</td>
</tr>
<tr>
<td>E Delta.QSet2.q1</td>
<td>-0.57391052</td>
<td>0.05263158</td>
<td>0.67917368</td>
<td>1.39338454</td>
<td>19</td>
</tr>
<tr>
<td>F Delta.QSet2.Placebo</td>
<td>-0.21917380</td>
<td><strong>0.47368420</strong></td>
<td>0.47368420</td>
<td>1.16654220</td>
<td>19</td>
</tr>
</tbody>
</table>

\[50\] Here we see very little progress. BPMN score is similar to placebo score.
### 4.3.2.2. R code

```r
AO_ftTestP <- function(AO_v1,AO_v2){
t.test(AO_v1,AO_v2,alternative="two.sided",paired=TRUE)
}
AO_ftTestP (subset(AO_lab,Source="USYD" &First.Questionnaire="BPMN")$Delta.QSet2.Placebo, subset(AO_lab,Source="USYD" &First.Questionnaire="BPMN")$Delta.QSet1.Placebo)
AO_fwilcoxTestP <- function(AO_v1,AO_v2){ wilcox.test(AO_v1,AO_v2,alternative="two.sided",paired= TRUE, exact=TRUE) }
AO_fwilcoxTestP (subset(AO_lab,Source="USYD" &First.Questionnaire="BPMN")$Delta.QSet2.Placebo, subset(AO_lab,Source="USYD" &First.Questionnaire="BPMN")$Delta.QSet1.Placebo)
```

Table 19 – H₀—Presenting Cockburn to USYD students who already read BPMN doesn’t improve comprehension —✓ Not rejected

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1: USYD students who got BPMN first</td>
<td>Delta.QSet1.Placebo</td>
</tr>
<tr>
<td>Set 2: USYD students who got BPMN first</td>
<td>Delta.QSet2.Placebo</td>
</tr>
</tbody>
</table>

#### assuming normality
- Paired t-test
  - data: AO_v1 and AO_v2
  - t = 0.1646, df = 18, p-value = 0.871
  - alternative hypothesis: true difference in means is not equal to 0
  - 95 percent confidence interval: -0.6189583  0.7242214
  - sample estimates: mean of the differences 0.05263158

#### not assuming normality
- Wilcoxon signed rank test with continuity correction
  - data: AO_v1 and AO_v2
  - V = 41, p-value = 0.9043
  - alternative hypothesis: true location shift is not equal to 0
  - Warning messages:
    1: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided",: cannot compute exact p-value with ties
    2: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", :cannot compute exact p-value with zeroes

### 4.3.2.3. Formal results

Compensating for domain knowledge, students from the general university population did not show significantly better understanding of a workflows articulated as a BPMN set compared to their initial understanding of identical workflow articulated as written use case.
4.3.3. TU/e Students, Cockburn and then BPMN

$H_0$ – Presenting BPMN to TU/e students who already read Cockburn doesn’t improve comprehension – **Rejected**

$H_1$ – Presenting BPMN to TU/e students who already read Cockburn improves comprehension – **Accepted**

### 4.3.3.1. TU/e Results - First notation Cockburn

![Statistical data and table](image-url)
4.3.3.2. R Code

z.test requires the BSDA package which need to be downloaded from [http://cran.r-project.org/web/packages/BSDA/index.html](http://cran.r-project.org/web/packages/BSDA/index.html) and installed using the R GUI:

Packages → Install package from local zip file Packages → Load Package ...

```r
## --------------------------------------
## paired two sided Z-Test
## --------------------------------------

AO_fzTestP <- function(AO_v1, AO_v2) { 
z.test(AO_v1, AO_v2, alternative = "two.sided", mu = 0, sigma.x = sd(AO_v1), sigma.y = sd(AO_v2), conf.level = 0.95)
}

AO_fzTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$ Delta.QSet2.Placebo, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$ Delta.QSet1.Placebo)

## --------------------------------------
## paired one sided Z-Test
## --------------------------------------

AO_fzTestP <- function(AO_v1, AO_v2) { 
z.test(AO_v1, AO_v2, alternative = "greater", mu = 0, sigma.x = sd(AO_v1), sigma.y = sd(AO_v2), conf.level = 0.95)
}

AO_fzTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$ Delta.QSet2.Placebo, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$ Delta.QSet1.Placebo)

## --------------------------------------
## paired two sided wilcox
## --------------------------------------

AO_fwilcoxTestP <- function(AO_v1, AO_v2) { 
wilcox.test (AO_v1, AO_v2, alternative="two.sided", paired= TRUE, exact=TRUE) }

AO_fwilcoxTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$ Delta.QSet2.Placebo, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$ Delta.QSet1.Placebo)

## --------------------------------------
## paired one sided wilcox
## --------------------------------------

AO_fwilcoxTestP <- function(AO_v1, AO_v2) { 
wilcox.test (AO_v1, AO_v2, alternative="greater", paired= TRUE, exact=TRUE) }

AO_fwilcoxTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$ Delta.QSet2.Placebo, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="Cockburn")$ Delta.QSet1.Placebo)
```
R Test Log

> ## -----------------------
> ## paired two sided Z-Test
> ## -----------------------
> AO_fzTestP <- function(AO_v1,AO_v2){ z.test(AO_v1, AO_v2, alternative = "two.sided", mu = 0, sigma.x = sd(AO_v1), sigma.y = sd(AO_v2), conf.level = 0.95)}
> AO_fzTestP (subset(AO_lab,Source=="TU/e" & First.Questionnaire=="Cockburn")$Delta.QSet2.Placebo,subset(AO_lab,Source=="TU/e" & First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo)

Two-sample z-Test
data:  AO_v1 and AO_v2
z = 1.4166, p-value = 0.1566
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.167382  1.0401098
sample estimates:
mean of x mean of y
1.3818182 0.9454545

> > ## paired one sided Z-Test
> > AO_fzTestP <- function(AO_v1,AO_v2){ z.test(AO_v1, AO_v2, alternative = "greater", mu = 0, sigma.x = sd(AO_v1), sigma.y = sd(AO_v2), conf.level = 0.95)}
> AO_fzTestP (subset(AO_lab,Source=="TU/e" & First.Questionnaire=="Cockburn")$Delta.QSet2.Placebo,subset(AO_lab,Source=="TU/e" & First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo)

Two-sample z-Test
data:  AO_v1 and AO_v2
z = 1.4166, p-value = 0.0783
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -0.0703161         NA
sample estimates:
mean of x mean of y
1.3818182 0.9454545

> > ## paired two sided wilcox
> > AO_fwilcoxTestP <- function(AO_v1,AO_v2){ wilcox.test
(AO_v1,AO_v2,alternative="two.sided",paired= TRUE, exact=TRUE) }
> AO_fwilcoxTestP (subset(AO_lab,Source=="TU/e" & First.Questionnaire=="Cockburn")$Delta.QSet2.Placebo,subset(AO_lab,Source=="TU/e" & First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo)

Wilcoxon signed rank test with continuity correction
data:  AO_v1 and AO_v2
V = 436.5, p-value = 0.01404
alternative hypothesis: true location shift is not equal to 0
Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided",  :
cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided",  :
cannot compute exact p-value with zeroes
Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 436.5, p-value = 0.007018
alternative hypothesis: true location shift is greater than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE,  :
cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE,  :
cannot compute exact p-value with zeroes

>
Table 20 – H1 – Presenting BPMN to TU/e students who already read Cockburn improves comprehension – **✓ Accepted (P = 0.01404)**

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Set 1:</strong> TU/e students who got Cockburn first</td>
<td>answer to QSet2 - Placebo</td>
</tr>
<tr>
<td><strong>Set 2:</strong> TU/e students who got Cockburn first</td>
<td>answer to QSet1 - Placebo</td>
</tr>
</tbody>
</table>

assuming normality:

- Two-sample z-Test
  - data: AO_v1 and AO_v2
  - \( z = 1.4166, \)  **p-value = 0.1566**
  - alternative hypothesis: true difference in means is not equal to 0
  - 95 percent confidence interval: \(-0.1673825 \text{ to } 1.0401098\)
  - sample estimates: mean of x mean of y \(1.3818182 \text{ and } 0.9454545\)

not assuming normality:

- Wilcoxon signed rank test with continuity correction
  - data: AO_v1 and AO_v2
  - \( V = 436.5, \)  **p-value = 0.01404**
  - alternative hypothesis: true location shift is not equal to 0

**Warning messages:**
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", : cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", : cannot compute exact p-value with zeroes

### 4.3.3.3. Formal results

Compensating for domain knowledge, BPM students **showed significantly better** understanding of a workflows articulated as a BPMN set compared to their initial understanding of identical workflow articulated as written use case (\(P=0.01\)).
4.3.4. TU/e students, BPMN and then Cockburn

H₀ – Presenting Cockburn to TU/e students who already read BPMN doesn’t reduce comprehension
- ✔ Not rejected

H₁ – Presenting Cockburn to TU/e students who already read BPMN reduces comprehension
- ✔ Accepted

### 4.3.4.1. TU/e Results - First notation Cockburn

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Score.Placebo</td>
<td>1.9594590</td>
<td>2.080210</td>
<td>1.0909240</td>
<td>74</td>
</tr>
<tr>
<td>B Score.QSet1</td>
<td>3.4054050</td>
<td>3.7069200</td>
<td>1.3233300</td>
<td>74</td>
</tr>
<tr>
<td>C Score.QSet2</td>
<td>3.5560830</td>
<td>3.7303630</td>
<td>1.3730360</td>
<td>74</td>
</tr>
<tr>
<td>D Delta.QSet1.Placebo</td>
<td>1.4459460</td>
<td>1.7877490</td>
<td>1.5001540</td>
<td>74</td>
</tr>
<tr>
<td>E Delta.QSet2.q1</td>
<td>-0.1621622</td>
<td>0.1098625</td>
<td>1.193903</td>
<td>74</td>
</tr>
<tr>
<td>F Delta.QSet2.Placebo</td>
<td>1.2837838</td>
<td>1.6263733</td>
<td>1.5036054</td>
<td>74</td>
</tr>
</tbody>
</table>

51 Here we see negative progress – the results after reading BPM and Cockburn are worth than the results after only reading BPM!
### 4.3.4.2. R Code

`z.test` requires the BSDA package which need to be downloaded from [http://cran.r-project.org/web/packages/BSDA/index.html](http://cran.r-project.org/web/packages/BSDA/index.html) and installed using the R GUI:

Packages → Install package from local zip file Packages → Load Package …

```r
AO_fzTestP <- function(AO_v1, AO_v2) { 
  z.test(AO_v1, AO_v2, alternative = "two.sided", 
        mu = 0, sigma.x = sd(AO_v1), sigma.y = sd(AO_v2), 
        conf.level = 0.95)
}
AO_fzTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$ Delta.QSet1.Placebo, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$ Delta.QSet2.Placebo)
```

```r
AO_fwilcoxTestP <- function(AO_v1, AO_v2) { 
  wilcox.test(AO_v1, AO_v2, alternative="two.sided", paired=TRUE, exact=TRUE) 
}
AO_fwilcoxTestP (subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$ Delta.QSet1.Placebo, subset(AO_lab, Source=="TU/e" & First.Questionnaire=="BPMN")$ Delta.QSet2.Placebo)
```

### Table 21 – H₀ – Presenting Cockburn to TU/e students who already read BPMN doesn’t reduce comprehension —✓ Not rejected

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1: TU/e students who got BPMN first</td>
<td>answer to QSet2 - Placebo</td>
</tr>
<tr>
<td>Set 2: TU/e students who got BPMN first</td>
<td>answer to QSet1 - Placebo</td>
</tr>
</tbody>
</table>

#### assuming normality
- Two-sample z-Test
  - data:  AO_v1 and AO_v2 z = 0.6568, \( p\)-value = 0.5113
  - alternative hypothesis: true difference in means is not equal to 0
  - 95 percent confidence interval: \([-0.3217680, 0.6460923]\)
  - sample estimates: mean of x mean of y \(1.445946, 1.283784\)

#### not assuming normality
- Wilcoxon signed rank test with continuity correction
  - data:  AO_v1 and AO_v2  \(V = 482\), \( p\)-value = 0.3192
  - alternative hypothesis: true location shift is not equal to 0
- Warning messages:
  1: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", : cannot compute exact p-value with ties
  2: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", :cannot compute exact p-value with zeroes

### 4.3.4.3. Formal results

Compensating for domain knowledge, BPM students did not show significantly worth understanding of a workflows articulated as a written use case set compared to their initial understanding of identical workflow articulated as BPMN.
4.3.5. HU/b Students, Cockburn then BPMN

$H_0$ – Presenting BPMN to HU/b students who already read Cockburn doesn’t improve comprehension – Not rejected

$H_1$ – Presenting BPMN to HU/b students who already read Cockburn improves comprehension – Rejected

4.3.5.1. HU/b - First notation Cockburn Results

\[
\text{subset(AO_lab,Source=="HU/b" \\
&First.Questionnaire=="Cockburn")}\$\text{Delta.QSet1.Placebo}
3 2 -1 3 -1 5 2 1 1 0 0 0 0
\]

\[
\text{subset(AO_lab,Source=="HU/b" \\
&First.Questionnaire=="Cockburn")}\$\text{Delta.QSet2.Placebo}
4 -2 -1 2 -1 5 3 3 2 0 0 1 0
\]

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>mean</th>
<th>Conf 0.05 to</th>
<th>sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>a Score.Placebo</td>
<td>1.00605800</td>
<td>1.61538500</td>
<td>2.22471100</td>
<td>1.12089700</td>
<td>13</td>
</tr>
<tr>
<td>b Score.QSet1</td>
<td>1.93430900</td>
<td>2.76923100</td>
<td>3.60415300</td>
<td>1.53589500</td>
<td>13</td>
</tr>
<tr>
<td>c Score.QSet2</td>
<td>1.88272400</td>
<td>2.84615400</td>
<td>3.80958400</td>
<td>1.77229400</td>
<td>13</td>
</tr>
<tr>
<td>d Delta.QSet1.Placebo</td>
<td>0.19041620</td>
<td>1.15384620</td>
<td>2.11727610</td>
<td>1.77229390</td>
<td>13</td>
</tr>
<tr>
<td>e Delta.QSet2.q1</td>
<td>-0.70649683</td>
<td>0.07692308</td>
<td>0.07692308</td>
<td>0.07692308</td>
<td>13</td>
</tr>
<tr>
<td>f Delta.QSet2.Placebo</td>
<td>0.07432566</td>
<td>1.23076923</td>
<td>2.38721280</td>
<td>2.12735541</td>
<td>13</td>
</tr>
</tbody>
</table>

Here we see that the BPMN second did not help much.
4.3.5.2. R Code

AO_ftTestP <- function(AO_v1, AO_v2){
t.test(AO_v1, AO_v2, alternative="two.sided", paired=TRUE)
}
AO_ftTestP (subset(AO_lab, Source=="HU/b" & First.Questionnaire="Cockburn")$Delta.QSet2.Placebo, subset(AO_lab, Source=="HU/b" & First.Questionnaire="Cockburn")$Delta.QSet1.Placebo)

AO_fwilcoxTestP <- function(AO_v1, AO_v2){
  wilcox.test(AO_v1, AO_v2, alternative="two.sided", paired= TRUE, exact = TRUE)
}
AO_fwilcoxTestP (subset(AO_lab, Source=="HU/b" & First.Questionnaire="Cockburn")$Delta.QSet2.Placebo, subset(AO_lab, Source=="HU/b" & First.Questionnaire="Cockburn")$Delta.QSet1.Placebo)

Table 22 – H0 – Presenting BPMN to HU/b students who already read Cockburn doesn’t improve comprehension – ☑ Not rejected

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1: HU/b students who got Cockburn first</td>
<td>answer to QSet2 – Placebo</td>
</tr>
<tr>
<td>Set 2: HU/b students who got Cockburn first</td>
<td>answer to QSet1 – Placebo</td>
</tr>
</tbody>
</table>

- **assuming normality**
  - Paired t-test
  - data: AO_v1 and AO_v2
  - t = 0.1925, df = 12, \( p\)-value = 0.8506
  - alternative hypothesis: true difference in means is not equal to 0
  - 95 percent confidence interval: -0.7939572 0.9478034
  - sample estimates: mean of the differences 0.07692308

- **not assuming normality**
  - Wilcoxon signed rank test with continuity correction
  - data: AO_v1 and AO_v2
  - V = 18, \( p\)-value = 0.5393
  - alternative hypothesis: true location shift is not equal to 0
  - Warning messages:
    1: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", :cannot compute exact p-value with ties
    2: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", :cannot compute exact p-value with zeroes

4.3.5.3. Formal results

Compensating for domain knowledge, BPM students **did not show** significantly worth understanding of a workflows articulated as a written use case set compared to their initial understanding of identical workflow articulated as BPMN.
4.3.6. **HU/b Students, BPMN and then Cockburn**

**H₀** – Presenting Cockburn to HU/b students who already read BPMN doesn’t reduce comprehension

- ☑ **Not rejected**

**H₁** – Presenting Cockburn to HU/b students who already read BPMN reduces comprehension

- ☑ **rejected**

**4.3.6.1. HU/b - First notation BPMN Results**

```
subset(AO_lab, Source=="HU/b" & First.Questionnaire=="BPMN")$Delta.QSet2.Placebo
1  3  2  1  -1  -1  2  3  1  1  0  -1  -1
```

```
subset(AO_lab, Source=="HU/b" & First.Questionnaire=="BPMN")$Delta.QSet1.Placebo
1  3  3  0  -1  1  1  3  2  2  0  -1  0
```

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Conf 0.05 from</th>
<th>Mean</th>
<th>Conf 0.05 to</th>
<th>sd</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>a Score.Placebo</td>
<td>1.37511000</td>
<td>1.76923100</td>
<td>2.16335100</td>
<td>0.72501100</td>
<td>13</td>
</tr>
<tr>
<td>b Score.QSet1</td>
<td>2.26548000</td>
<td>2.84615400</td>
<td>3.42682700</td>
<td>1.06818800</td>
<td>13</td>
</tr>
<tr>
<td>c Score.QSet2</td>
<td>1.85029700</td>
<td>2.53846200</td>
<td>3.22662600</td>
<td>1.26592400</td>
<td>13</td>
</tr>
<tr>
<td>d Delta.QSet1.Placebo</td>
<td>0.29350320</td>
<td>1.07692310</td>
<td>1.86034300</td>
<td>1.44115340</td>
<td>13</td>
</tr>
<tr>
<td>e Delta.QSet2.q1</td>
<td>-0.77239430</td>
<td>-0.30769230</td>
<td>0.15700970</td>
<td>0.85485040</td>
<td>13</td>
</tr>
<tr>
<td>f Delta.QSet2.Placebo</td>
<td>-0.03565667</td>
<td>0.76923077</td>
<td>1.57411821</td>
<td>1.48064435</td>
<td>13</td>
</tr>
</tbody>
</table>

Here we see that administering Cockburn second makes things worst even in Berlin.  

---

53 Here we see that administering Cockburn second makes things worst even in Berlin.
4.3.6.2. R Code
AO_ftTestP <- function(AO_v1, AO_v2) {
t.test(AO_v1, AO_v2, alternative="two.sided", paired=TRUE)
} 
AO_ftTestP(subset(AO_lab, Source=="HU/b" 
& First.Questionnaire=="BPMN")$Delta.QSet1.Placebo, 
subset(AO_lab, Source=="HU/b" 
& First.Questionnaire=="BPMN")$Delta.QSet2.Placebo)
AO_fwilcoxTestP <- function(AO_v1, AO_v2) { wilcox.test 
(AO_v1, AO_v2, alternative="two.sided", paired= TRUE, exact = TRUE) 
} 
AO_fwilcoxTestP(subset(AO_lab, Source=="HU/b" 
& First.Questionnaire=="BPMN")$Delta.QSet1.Placebo, 
subset(AO_lab, Source=="HU/b" 
& First.Questionnaire=="BPMN")$Delta.QSet2.Placebo)

Table 23 – H₀ – Presenting Cockburn to HU/b students who already read BPMN doesn’t reduce comprehension – Not rejected

<table>
<thead>
<tr>
<th>Population</th>
<th>Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Set 1:</strong></td>
<td>HU/b students who got BPMN first</td>
</tr>
<tr>
<td><strong>Set 2:</strong></td>
<td>HU/b students who got BPMN first</td>
</tr>
</tbody>
</table>

assuming normality
Paired t-test
data:  AO_v1 and AO_v2
t = 1.2978, df = 12, p-value = 0.2188
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:  -0.2088886  0.8242732
sample estimates: mean of the differences: 0.3076923

not assuming normality
Wilcoxon signed rank test with continuity correction
data:  AO_v1 and AO_v2
V = 21, p-value = 0.2402
alternative hypothesis: true location shift is not equal to 0
Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = 
"two.sided", :cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = 
"two.sided", :cannot compute exact p-value with zeroes

4.3.6.3. Formal results
Compensating for domain knowledge, BPM students did not show significantly better understanding of a workflows articulated as a written use case set compared to their initial understanding of identical workflow articulated as BPMN.
4.3.7. All students prefer Cockburn over placebo

*R Code*

```r
AO_ftTestP <- function(AO_v1, AO_v2){
  t.test(AO_v1, AO_v2, alternative="two.sided", paired=TRUE)
}

AO_ftTestP
  (subset(AO_lab, First.Questionnaire=="Cockburn")$Score.QSet1,
   subset(AO_lab, First.Questionnaire=="Cockburn")$Score.Placebo)

Paired t-test

data:  AO_v1 and AO_v2
  t = 6.0212, df = 92, p-value = 3.508e-08
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  0.6557377 1.3012516
sample estimates:
mean of the differences
  0.9784946

AO_ftTestP <- function(AO_v1, AO_v2){
  t.test(AO_v1, AO_v2, alternative="greater", paired=TRUE)
}

AO_ftTestP
  (subset(AO_lab, First.Questionnaire=="Cockburn")$Score.QSet1,
   subset(AO_lab, First.Questionnaire=="Cockburn")$Score.Placebo)

Paired t-test

data:  AO_v1 and AO_v2
  t = 6.0212, df = 92, p-value = 1.754e-08
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
  0.708472      Inf
sample estimates:
mean of the differences
  0.9784946
```
AO_fwilcoxTestP  <- function(AO_v1,AO_v2){ wilcox.test (AO_v1,AO_v2,alternative="two.sided",paired=TRUE, exact = TRUE) }

AO_fwilcoxTestP (subset(AO_lab, First.Questionnaire=="Cockburn")$Score.QSet1,subset(AO_lab,First.Questionnaire=="Cockburn")$Score.Placebo)

Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 2320.5, p-value = 3.17e-07
alternative hypothesis: true location shift is not equal to 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", :
cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", :
cannot compute exact p-value with zeroes

AO_fwilcoxTestP  <- function(AO_v1,AO_v2){ wilcox.test (AO_v1,AO_v2,alternative="greater",paired=TRUE, exact = TRUE) }

AO_fwilcoxTestP (subset(AO_lab, First.Questionnaire=="Cockburn")$Score.QSet1,subset(AO_lab,First.Questionnaire=="Cockburn")$Score.Placebo)

Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 2320.5, p-value = 1.585e-07
alternative hypothesis: true location shift is greater than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, :
cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, :
cannot compute exact p-value with zeroes

AO_fwilcoxTestP  <- function(AO_v1,AO_v2){ wilcox.test (AO_v1,AO_v2,alternative="greater",paired=TRUE, exact = TRUE) }

AO_fwilcoxTestP (subset(AO_lab, First.Questionnaire=="Cockburn")$Score.QSet1,subset(AO_lab,First.Questionnaire=="Cockburn")$Score.Placebo)
4.3.8. TU/e and HU/b students prefer BPMN over placebo

1.
AO_ftTestP <- function(AO_v1,AO_v2){
t.test(AO_v1,AO_v2,alternative="two.sided",paired=TRUE) }

AO_ftTestP (subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")$Score.QSet1,subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")$Score.Placebo)

Paired t-test
data:  AO_v1 and AO_v2
t = 8.7113, df = 86, p-value = 1.904e-13
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  1.073422 1.708187
sample estimates:
mean of the differences
               1.390805

2.
AO_ftTestP <- function(AO_v1,AO_v2){
t.test(AO_v1,AO_v2,alternative="greater",paired=TRUE) }

AO_ftTestP (subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")$Score.QSet1,subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")$Score.Placebo)

Paired t-test
data:  AO_v1 and AO_v2
t = 8.7113, df = 86, p-value = 9.522e-14
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
  1.125337      Inf
sample estimates:
mean of the differences
               1.390805
3.

AO_fwilcoxTestP <- function(AO_v1,AO_v2) {
  wilcox.test(AO_v1,AO_v2,alternative="two.sided",paired=TRUE, exact = TRUE)
}

AO_fwilcoxTestP (subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")$Score.QSet1,subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")$Score.Placebo)

Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 2322, p-value = 1.604e-10
alternative hypothesis: true location shift is not equal to 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", : cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", : cannot compute exact p-value with zeroes

4.

AO_fwilcoxTestP <- function(AO_v1,AO_v2) { wilcox.test(AO_v1,AO_v2,alternative="greater",paired=TRUE, exact = TRUE)
}

AO_fwilcoxTestP (subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")$Score.QSet1,subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")$Score.Placebo)

Wilcoxon signed rank test with continuity correction

data:  AO_v1 and AO_v2
V = 2322, p-value = 8.021e-11
alternative hypothesis: true location shift is greater than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, : cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, : cannot compute exact p-value with zeroes
Presenting BPMN to students who already read Cockburn further improved comprehension

1. 
AO_ftTestP <- function(AO_v1, AO_v2){
t.test(AO_v1, AO_v2, alternative="two.sided", paired=TRUE) }

AO_ftTestP (subset(AO_lab, 
First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo, subset(AO_lab, 
First.Questionnaire=="Cockburn")$Delta.QSet2.Placebo)

Paired t-test

data:  AO_v1 and AO_v2
t = -3.8772, df = 92, p-value = 0.0001981 (0.0002)
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.7479942 -0.2412531
sample estimates:
mean of the differences
-0.4946237

2. 
AO_ftTestP <- function(AO_v1, AO_v2){
t.test(AO_v1, AO_v2, alternative="greater", paired=TRUE) }

AO_ftTestP (subset(AO_lab, 
First.Questionnaire=="Cockburn")$Delta.QSet2.Placebo, subset(AO_lab, 
First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo)

Paired t-test

data:  AO_v1 and AO_v2
t = 3.8772, df = 92, p-value = 9.904e-05
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.2826506 Inf
sample estimates:
mean of the differences
 0.4946237
3. 

AO_fwilcoxTestP <- function(AO_v1, AO_v2) {
  wilcox.test(AO_v1, AO_v2, alternative = "two.sided", paired = TRUE, exact = TRUE)
}

AO_fwilcoxTestP(subset(AO_lab, First.Questionnaire == "Cockburn")$Delta.QSet2.Placebo, subset(AO_lab, First.Questionnaire == "Cockburn")$Delta.QSet1.Placebo)

Wilcoxon signed rank test with continuity correction

data: AO_v1 and AO_v2
V = 1415, \textit{p-value} = 0.0001199 (0.0001)
alternative hypothesis: true location shift is not equal to 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", : cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "two.sided", : cannot compute exact p-value with zeroes

4. 

AO_fwilcoxTestP <- function(AO_v1, AO_v2) {
  wilcox.test(AO_v1, AO_v2, alternative = "greater", paired = TRUE, exact = TRUE)
}

AO_fwilcoxTestP(subset(AO_lab, First.Questionnaire == "Cockburn")$Delta.QSet2.Placebo, subset(AO_lab, First.Questionnaire == "Cockburn")$Delta.QSet1.Placebo)

Wilcoxon signed rank test with continuity correction

data: AO_v1 and AO_v2
V = 1415, \textit{p-value} = 5.993e-05
alternative hypothesis: true location shift is greater than 0

Warning messages:
1: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, : cannot compute exact p-value with ties
2: In wilcox.test.default(AO_v1, AO_v2, alternative = "greater", paired = TRUE, : cannot compute exact p-value with zeroes
4.3.9. Plotting

# generating values to be plotted
dat <- AO_lab$Delta.QSet1.Placebo

# plotting function
histplot <- function(dat, breaks="Sturges", ncurve=TRUE, ...)
{
    # compute the histogram and density of "dat"
    hdat <- hist(dat, breaks=breaks, plot=F)
    ddat <- density(dat)

    # compute the xlim and ylim of the plot
    # i.e. the min and max of the different superimposed
    # plots (hist, density and normal curves)
    xlim <- range(ddat$x)
    ylim <- c(0, max(hdat$density, ddat$y))
    if(ncurve)
    {
        # max of the normal curve
        maxnorm <- pnorm(mean(dat), mean=mean(dat), sd=sd(dat))
        ylim <- c(0, max(hdat$density, ddat$y, maxnorm))
    }
    else
    {
        ylim <- c(0, max(hdat$density, ddat$y))
    }

    # plotting
    plot(hdat,
         freq=F,
         xlim=xlim, ylim=ylim, ...)
    lines(ddat)
    if (ncurve) curve(dnorm(x, mean=mean(dat), sd=(sd(dat))),
                       lty=3, add=TRUE)
}

# usage
histplot(dat)
histplot(dat, ncurve=F)
histplot(dat, col="blue") # arguments are passed to the hist plotting function
4.3.10. All students prefer BPMN over Cockburn

**R Code**

AO_ftTestP <- function(AO_v1, AO_v2) {
  t.test(AO_v1, AO_v2, alternative="two.sided", paired=FALSE) }

AO_ftTestP (subset(AO_lab, First.Questionnaire=="BPMN")$Delta.QSet1.Placebo, subset(AO_lab, First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo)

Welch Two Sample t-test

data:  AO_v1 and AO_v2
  t = 1.023, df = 194.047, p-value = 0.3076
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.2107418  0.6649675
sample estimates:
mean of x mean of y
1.2056075 0.9784946

AO_ftTestP <- function(AO_v1, AO_v2) {
  t.test(AO_v1, AO_v2, alternative="greater", paired=FALSE) }

AO_ftTestP (subset(AO_lab, First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo, subset(AO_lab, First.Questionnaire=="Cockburn")$Delta.QSet1.Placebo)

Welch Two Sample t-test

data:  AO_v1 and AO_v2
  t = 1.023, df = 194.047, p-value = 0.1538
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -0.1398063        Inf
sample estimates:
mean of x mean of y
1.2056075 0.9784946
4.3.11. TU/e and HU/b students get more from Cockburn than USYD students

AO_ftTestP <- function(AO_v1,AO_v2){
  t.test(AO_v1,AO_v2,alternative="greater", paired=FALSE) }

AO_ftTestP (subset(AO_lab, (Source=="HU/b" | Source=="TU/e") &
First.Questionnaire=="Cockburn")$Score.QSet1,subset(AO_lab, (Source=="USYD") & First.Questionnaire=="Cockburn")$Score.Placebo)

Welch Two Sample t-test

data:  AO_v1 and AO_v2

  t = 3.6815, df = 65.809, p-value = 0.0002349
 alternative hypothesis: true difference in means is greater than 0
 95 percent confidence interval:
    0.4766522       Inf
 sample estimates:
 mean of x mean of y
  2.735294   1.863636
4.3.12. TU/e and HU/b students get more from BPMN than USYD students

AO_ftTestP  <- function(AO_v1,AO_v2){
t.test(AO_v1,AO_v2,alternative="greater", paired=FALSE) }

AO_ftTestP (subset(AO_lab, (Source=="HU/b" | Source=="TU/e") & First.Questionnaire=="BPMN")(Source=="USYD") & First.Questionnaire=="BPMN")$Score.Placebo)

Welch Two Sample t-test

data:  AO_v1 and AO_v2
t = 4.0508, df = 27.593, p-value = 0.0001873
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
  0.7359339       Inf
sample estimates:
mean of x mean of y
  3.321839  2.052632
Appendix – Participant work book
1 Participant’s Information Statement

What is the study about?

We are doing some research in software engineering, a field which aims to improve the way software is developed. We are studying design notations that are used to communicate between IT professionals and the stakeholders such as managers for whom software is developed. In particular, we look at notations which describe the details of how a business process takes place. Different development teams use different notations when they explain this aspect to stakeholders; some of the notations are graphical and others use only text. While there has been a lot of research comparing the expressive power, accuracy and ambiguity of such notations, there is a gap, which we try to fill, in comparing them for usability and clarity. Notations that are not clear when read by stakeholders may cause unsatisfactory results, even if the notations are very expressive for the IT professional who writes the descriptions.

What does the study involve?

To conduct the research we have documented a business process using three different business process design notations. We imagine a fictitious company XYZ, and we document the process it uses to manage delivery of technological solutions to business problems. The process was written with one organisation in mind, so it is not generic, and probably does not fit many other organisations. As documenting a process is always a work in progress, we deliberately included some errors, omissions and inconsistencies. The inconsistencies and incompleteness simulate real life business modeling situations where designs are at times incomplete, ambiguous or even wrong.

In this experiment, you will be acting like a business stakeholder. You will be asked to read the descriptions of the business process, and we will explore how well the notation communicated to you. You have received a booklet that contains the experiment’s material. As you read the material you will be asked three times to answer a short questionnaire. The questionnaires are identical. Our
experiment measures the difference between the answers, not the answers themselves. Please answer the questionnaires using all the information you know, rather than only the information we have given you most recently. Do not guess answers, if you do not know an answer to a question please tick the "I do not know" option. We have also included a "There is insufficient information" option. We are not assessing your skill, so you cannot be right or wrong; rather we are trying to judge the impact of the different design notations. Please read one section at a time and then answer its questionnaire. You may return to old sections.

Following all of this, we will ask you some questions to investigate your preference for communicating, whether you like graphical or verbal approaches.

Please do not write your name on the questionnaire, as we intend to keep the responses so they are not connected with your identity, in order to protect your privacy.

Who is carrying out the study?

The study is being conducted by PhD student Avner Ottensooser, and will form part of the basis for his thesis for the degree of Doctor of Philosophy at The University of Sydney under the supervision of Associate Professor Alan Fekete.

How much time will the study take?

The session will last approximately one hour. The session will consist of approximately 15 minutes for the background and post-experiment questionnaires, and approximately 45 minutes to complete the practical tasks.

Can I withdraw from the study?

Being in this study is completely voluntary - you are not under any obligation to participate. If you do consent to participate, you can later withdraw at any time until the booklet is handed in. Withdrawal will not affect your relationship with the researchers or the University of Sydney in any way. You may stop the experiment at any time if you do not wish to continue. However, once you have submitted your questionnaire you cannot withdraw as we will not be able to identify yours

Will anyone else know the results?

All aspects of the study, including individual results, will be strictly confidential even the researchers will not have access to information on participants. All recorded data and responses will not be associated with any name. While the envelope you received is numbered the booklet is not. A report of the study will be submitted for publication, but individual participants will not be identifiable in such a report.
Will the study benefit me?
You will receive AU$20.00 as compensation for your time. You are also making a valuable contribution to our ongoing research on design notations. The experience you have here may also be practice that can make your work easier if you pursue a career as manager or similar stakeholder who needs to interact with software development projects.

Can I tell other people about the study?
The study is not confidential and you are free to tell others.

What if I require further information?
If you are interested in the result of the research, of which you are now an important part, please drop us a note at avner@it.usyd.nsw.au. We will gladly share the research results with you in due course. When you have read this information, the lab supervisor will discuss it with you further and answer any questions you may have. If you would like to know more at any stage, please feel free to contact Avner Ottensooser (e-mail: avner@it.usyd.edu.au, ph: 0402 798 460), or Professor Alen Fekete (e-mail: fekete@it.usyd.edu.au, ph: 02 9351 4287).

What if I have a complaint or concerns?
Any person with concerns or complaints about the conduct of a research study can contact the Manager, Ethics Administration, University of Sydney on (02) 8627 8175 (Telephone); (02) 8627 8180 (Facsimile) or gbriody@usyd.edu.au (Email).

This information sheet is for you to keep
2 Instrument I – PARTICIPANT CONSENT FORM

This page is intentionally left empty.
Participant’s Consent Form

I, .............................................................................[PRINT NAME], give consent to my participation in the research project titled:  Process Notation – Usability Study.

In giving my consent I acknowledge that:

1. I am at least 18 years old.

2. The procedures required for the project and the time involved have been explained to me, and any questions I have about the project have been answered to my satisfaction.

3. I have read the Participant Information Statement and have been given the opportunity to discuss the information and my involvement in the project with the researcher/s.

4. I understand that I can withdraw from the study at any time until I hand in the questionnaire, without affecting my relationship with the researcher(s) or the University of Sydney now or in the future. If I withdraw, my contact details will be removed from the files.

5. I understand that my involvement is strictly confidential and no information about me will be used in any way that reveals my identity. I understand that I may optionally give permission for my answered questioners and survey to be used in future research publications arising from this study.

6. I understand that being in this study is completely voluntary - I am not under any obligation to consent.
7. I consent to the use of my answered questionnaire and survey for research publications and presentations:

YES NO

Signed:..............................................................

Name:..............................................................

Date:...............................................................
3 Instrument II – Business Problem

With its global presence, company XYZ is the well established leader in the Australian financial services industry.

Company XYZ’s management acknowledges that its core competency is not in the field of software development. However, due to the importance of software in the financial services industry, management had established a two tier strategy: usage of acquired software for general operations, and usage of in-house software innovations for the timely provision of solutions for core operations, so important for their competitive advantage.

Having witnessed several project failures, the company established a risk management culture. The company’s management encourages the cancellation of projects that lag behind in cost justification measures, in order to free scarce resources.

Aiming at a repeatable process, the company instituted in 1998 a Solution Delivery Framework (SDF). The SDF articulates activities, deliverables and milestones mandated for the delivery of every software solution, whether acquired or developed in-house. The SDF’s nature is that of a classical waterfall framework, gradually flowing between the development phases from requirement gathering, to specification, procurement, testing and deployment. Company XYZ adopted lessons from the contemporary iterative frameworks by concentrating on a multitude of small projects, close customer liaison, short development cycles and early testing, early delivery of pilots, and constant assessment and adjustment of upstream deliverables as solutions evolve.

To coordinate its investment in software, a project office had been established. Over the years the project office developed a spreadsheet based solution that supports the Solution Delivery Framework. The number and size of projects, let alone the complexity of supporting multiple teams have stretched the limits of the current solution and a new solution is required.
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDF</td>
<td>Solution Delivery Framework</td>
<td>A process describing the activities that have to be performed, artifacts that have to be created and milestones that have to be met in order to deliver a software solution.</td>
</tr>
<tr>
<td>SSF</td>
<td>Strategic Solution Framework</td>
<td>A five year plan articulating the emerging business needs of company XYZ, solutions company XYZ will acquire to meet these needs, a time table, and build//buy classification.</td>
</tr>
<tr>
<td>RMP</td>
<td>Risk Management Profile</td>
<td>A description of the risks that could prevent success in the project, classified for seriousness, likelihood, ways of mitigating them, etc.</td>
</tr>
<tr>
<td>SCD</td>
<td>Solution Concept Description</td>
<td>A very high level account of one way to provide the business with software that solves its problem.</td>
</tr>
<tr>
<td>PSD</td>
<td>Project Scope Document</td>
<td>A description of the aspects that are covered by the solution, and those that are left to other projects.</td>
</tr>
<tr>
<td>IR</td>
<td>Incident (bug) Report</td>
<td>A description of the circumstances in which the software did not behave as it was supposed to.</td>
</tr>
<tr>
<td>HLD</td>
<td>High Level Design</td>
<td>A high level description, showing the most important aspects only.</td>
</tr>
<tr>
<td>LLD</td>
<td>Low level design</td>
<td>The most detailed specification, issued to developers for coding.</td>
</tr>
<tr>
<td>TBA</td>
<td>To be advised</td>
<td>This acronym is used to acknowledged that an element in the design has not been decided yet.</td>
</tr>
<tr>
<td>Role</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Lead planner</td>
<td>A person who articulates all of the systems company XYZ will need over the planning horizon</td>
<td></td>
</tr>
<tr>
<td>Solution planner</td>
<td>A person who designs in detail one system that company XYZ will need in the short term</td>
<td></td>
</tr>
<tr>
<td>Project Manager</td>
<td>A person who has the overall responsibility for all resources needed to deliver one system that company XYZ needs in the short term.</td>
<td></td>
</tr>
<tr>
<td>Technical Lead</td>
<td>A person who has the overall responsibility for the technical aspects of one system that company XYZ needs in the short term.</td>
<td></td>
</tr>
<tr>
<td>Project Board</td>
<td>A body that has the overall authority for the full development lifecycle of a solution</td>
<td></td>
</tr>
<tr>
<td>Project Office</td>
<td>A body that has the central authority for the full development lifecycle of all the solutions company XYZ has</td>
<td></td>
</tr>
<tr>
<td>Development Team</td>
<td>A body that has the responsibility to implement a single solution</td>
<td></td>
</tr>
<tr>
<td>Stakeholders</td>
<td>All people that will be effected by the system, directly or indirectly, whether they use it or not.</td>
<td></td>
</tr>
</tbody>
</table>
4 Instrument III -- Company XYZ Product offering

Research notes

This is the first design notation we evaluate. The philosophy of this design notation is that one must understand the products of a company to understand its processes.

Overview

Company XYZ has three families of products which it sells through three distribution channels.

The products are:

- A basket of real estate properties
- A family of indexed linked funds that represent equities and bonds in different markets.

These products are distributed through three channels:

- Managed funds issued to retail customers
- Retirement products issued to individuals either during the accumulation stage or the retirement stage
- A hedge fund issued by invitation only to selected high net worth individuals.

The XYZ real estate basket

With US$250 Billion under management company XYZ’s controls basket of real estate properties ranging from the high speed rail (Beijing to Shanghai), airports (Denver, Frankfurt and Dubai), shopping malls and toll roads.

The XYZ family of index funds

The XYZ index funds aim to replicate the movements of an index of a specific financial market, or a set of rules of ownership that are held constant, regardless of market conditions. Tracking is achieved by trying to hold all of the securities in the index, in the same proportions as the index. Other methods include statistically sampling the market and holding "representative" securities. Some of company XYZ’s index funds rely on a computer model with little or no human input in the decision as to which securities are purchased or sold and is therefore a form of passive management.

The lack of active management (stock picking and market timing) gives the
advantage of lower fees and lower taxes in taxable accounts. However, the fees will generally reduce the return to the investor relative to the index. Company XYZ finds that it is impossible to precisely mirror the index as the models for sampling and mirroring, by their nature, cannot be 100% accurate. The difference between the index performance and the fund performance is known as the 'tracking error' or informally 'jitter'.

The XYZ index funds track some common indices including the S&P 500, the Wilshire 5000, the FTSE 100 and the FTSE All-Share Index.

The XYZ loan book
A book of loans issued to 23 of the S&P 500 companies.

The XYZ wholesale funds
Managed funds issued to retail customers. The fund is liquid and customers who apply for units by 9:30 am will have same day value. Customers who request a redemption by 2pm will have the funds transferred to their bank accounts by close of business same day.

The XYZ retirement fund
Retirement products issued to individuals either during the accumulation stage or the retirement stage.

The XYZ Hedge fund
Company XYZ offers to wealthy individuals the opportunity to participate in a hedge fund that is designed to be profitable regardless of the directions of movements in the market. The fund is less liquid than the wholesale product. The hedge funds are offered as either Capital Guaranteed flavor or naked flavor.

Capital Guaranteed The Capital Guaranteed flavor is a basket of high quality debt picked to mature at the instrument expiration date, and a basket of high beta long and short position, designed to achieve high return (albeit at high risk). The Capital Guaranteed product has a lifespan of five years during which redemptions are possible annually, at a high cost.

Naked The naked flavor is designed to have better performance over the medium range. Redemptions are processed every three months. Application for new units are currently closed.
5 Instrument IV – Comprehension Assessment Questionnaire

Q1 What happens if stake-holders change the project scope?

1. The project stops
2. The project manager revisits the project plan
3. Scope creep is addressed in subsequent projects
4. The stake holders talk with developers who alter their code
5. I do not know
6. The handout does not supply information needed to answer the question.
7. None of the above.

Q2 Who reviews the SCD?

1. People nominated by the project office review the SCD
2. The Project Office review the SCD
3. The SCD is rarely reviewed
4. The project office reviews the SCD, and if the SCD is deemed valid, the project office nominates other approvers
5. I do not know
6. The handout does not supply information needed to answer the question.
7. None of the above.

Q3 What document does the PSD supplement?

1. The PSD supplements the SCD
2. The PSD is a core document? It supplements no other documents.
3. The PSD supplements the HLD and the LLD
4. None of the above
5. I do not know
6. The handout does not supply information needed to answer the question.
7. None of the above.
Q4 How many test cycles does the SDF allow?
1. One at the project end
2. As many as needed until no severe bugs are found or the project is cancelled.
3. Several, one at the end of every project stage
4. None - tests are not done in cycles but end users test all the time
5. I do not know
6. The handout does not supply information needed to answer the question.
7. None of the above.

Q5 What happens immediately after a solution is documented?
1. The testing team tests the solution
2. The customer tests the solution
3. The testing team writes detailed use-cases
4. The customer writes detailed use-cases
5. I do not know
6. The handout does not supply information needed to answer the question.
7. None of the above.

Q6 What condition determines when a build is reiterated?
1. Whether a severe bug has been found
2. Whether a not-severe bug has been found
3. Whether a severe bug has been fixed
4. Whether a build was completed
5. I do not know
6. The handout does not supply information needed to answer the question.
7. None of the above.
6 Instrument V – Solution expressed as written use cases

Research notes

This is another design notation we evaluate. The philosophy of this design notation is that words are good at managing complexity.

Use Case #0 Deliver a software solution

Use case Scope: High level

Trigger: General Manager nominates a lead planner.

Primary Actor: Lead Planner who wishes to establish a five year Solution acquisition plan.

Actors: Project team, Technical team, Project Manager, Operation control

1: Lead Planner creates Strategic Solution Framework (SSF).
2: Project team analyses requirements and designs a solution.
3: Technical team builds a solution and tests it.
4: Project manager performs ongoing management review.
5: Technical team deploys the solution to Operation Control.
6: Project Manager decommissions the project, retaining a skeleton maintenance team.
7: Operations Control conducts ongoing maintenance.

Figure 1. Use Case #0 Deliver a software solution
Use Case #1 Create a Solution Concept Document

Use case Scope: High level

Triggers: Lead Planner creates a Strategic Solution Framework, SSF, articulating core solutions and a five year project plan.

Risk management lead defines a company wide risk management framework, RMF, articulating project governance structures depending on Risk/Cost profile and adherence with project plan.

Brief description: A solution planner creates a Solution Concept Document

Primary Actor: A solution planner who wishes to plan a solution.

Actors: Project Office, Approvers, Development Lead, the Lead Planner

Minimal Guarantee: A strategy is written and reviewed.

1: A solution planner, concentrating on one solution in the SSF, creates a Solution Concept Document, SCD, articulating the current state, problems in the current state, requirements from a new solution, top down milestones and estimated benefits. Order = 1

2: Project Office reviews the SCD. Order = 2

3: Project Office assigns approvers depending on the projects Risk/Cost profile. Order = 3

4: Approvers review the SCD and approve the project. Order = 4

5: Development lead assigns a development slot. Order = 5

6: Project Office establishes a Solution Delivery Entry in the company’s projects portfolio. Order = 6

Extensions:

1a: The solution planner finds that the SCD cannot be implemented with current technology / resources with acceptable risk, time frame or cost. 1a1: The development of the solution stops.

2a: Project office review finds the project concept document requires improvement 2a1: Technical lead repeats step 1.

2b: Project office review finds the project’s benefits to be insufficient or the project risk to be too high. 2b1: The project is cancelled.

4a: Approver finds the project not worth pursuing. 4a1: The project is cancelled.

4b: Approvers do not approve the project, specifying assumption errors. 4b1: The use case restarts at action step 1.

5a: Resources are not unavailable. 5a1: The project is cancelled

5b: The project will be possible only with less demand for resources 5b1: The use case restarts at action step 1.

Figure 2. Use Case #1 Create a Solution Concept Document
Use Case #2 Detail analysis and design

Use case Scope: High level

Trigger: Solution Concept Document (SCD) is completed and approved

Primary Actor: Project Manager who has to assess the project cost in detail

Actors: Project team, Technical Lead

Main success scenario

1: General Manager nominates a Project Manager. Order = 1

2: Project Manager supplements the solution concept document with detail, writing a Project Scope Document (PSD). Order = 2

3: Project Manager assembles the project team, including direct reports, customers, suppliers and auditors. Order = 3

4: Project team reviews Project Concept Document. Order = 4

5: Technical Lead writes High Level Design (HLD) evaluating several design avenues. Order = 5

6: Technical Lead writes an assessment of technology and methods selecting one of the design alternatives. Order = 6

7: Technical lead conducts vendor selection. Order = 7


9: Project manager writes Detailed Project Plan using input from the LLD (Time Line and Resources in Microsoft Project, Cash flow, Risk Management Plan). Order = 9

Extensions:

2a: The project manager finds the Solution Concept Document unrealistic
   2a1: The project ends

4a: The project team finds errors in the Project Concept Document
   4a1: The use case starts at action step 2

8a: While writing the LLD the technical lead finds errors in the HLD.
   8a1: The technical use case continues at action step 5.

Figure 3. Use Case #2 Detail analysis and design
Use Case #3.1 Build and test

Use case Scope: High level

Trigger: Low Level design completed

Description: This is the technical part of the project where the solution is crafted and verified.

Primary Actor: Technical Lead who wants to build the solution.

Actors: Development team, Testers, Project Manager, customers

1: Technical lead to write detailed use cases. Order = 1
2: Technical lead conducts procurement. Order = 2
3: Technical lead establishes development environments. Order = 3
4: Development team bring the solution to a build state. Order = 4
5: Development team to document the solution (Operational manual, training material, maintenance guide). Order = 5
6: Testers test the solution. Order = 6
7: Customers learn the solution and approve the solution. Order = 7

Extensions:

2a: Technical lead finds that the negotiated costs invalidate the project benefits.
   2a1: The use case is repeated from step 1.

2b: Technical lead finds that the negotiated costs invalidate the project beyond repair.
   2b1: The project is resumed from use case 2 step 4 (review project concept).

4a: The development team identifies extensions not handled in the use cases
   4a1: The use case continues at action step 1

4b: The changes to the use case are significant
   4b1: The Technical Lead refines the LLD (Reiterates use case 2 action steps 8 and 9).

6a: Tester finds a bug and writes an IR
   6a1: The project lead finds the bug worth fixing
   6a2: The technical lead finds the bug too minor to be fixed.
   6a2a: The project continues in step 6.
   6a3: The technical lead finds the bug and finds that the bug constitutes a change to the use cases.
   6a3a: The use case continues from step 1.
   6a4: The technical lead finds the bug a major scope change.
   6a4a: The use case continues in use case 1 step 2 (back to the drawing board).

Figure 4. Use Case #3.1 Build and test
Use Case #3.2 Monitor Building

Use case Scope: High level

Primary Actor: The project manager who wants to track progress against plan.

Actor: Project Board, Project Office

1: Project manager conducts communication with stake-holders Order = 1
2: Project manager maintains and reviews risk log Order = 1
3: Project manager monitors milestone adherence. Order = 1
4: Project manager monitors budget compliance. Order = 1
5: Project manager writes a weekly project status report. Order = 2
6: Project Board reviews project progress. Order = 3
7: The project office collates the minutes of the project board’s meetings in the company’s projects portfolio. Order = 4

Extensions:
1a: stake holders change the project scope 1a1: TBA
3a: New risk is identified
   3a1: The project manager adds the new risk to the risk log
3b: Existing risk becomes a reality
   3b1: The project continues from use case 3.1 step 2.
3c: Existing risk is rectified
   3c1: The project manager erases the risk from the risk log
7a: A project did not submit a weekly report
   7a1: TBA

Figure 5. Use Case #3.2 Monitor Building
Use Case #4 Deployment

Use case Scope: High level

Primary Actor: Development team that wants users to start using the system.

Actors: Users, Service Operation

1: Development team writes a user manual.  Order = 1
2: Development team conduct user training.  Order = 1
3: Development team deploys the software.  Order = 2
4: Development team hands over maintenance responsibility to Service Operations.  Order = 3
5: After the end of the warranty period, development team ceases to maintain the project.  Order = 4

Extensions:

2a: New requirements are identified while training users
2a1: The project resumes from use case 3.1 step 5.

2a: New bugs are identified while training users
2a1: The project resumes from use case 3.1 step 5.

5a: Too many bugs are found during the warranty period.
5a1: The Warranty period is extended.

Figure 6. Use Case #4 Deployment
Use Case #5 Close Project

Use case Scope: High level

Primary actor: Project manager who wishes to bring the project to closure.

Actors: Technical Manager, Solution Maintenance Team

1: Technical Manager reviews vendor’s contracts. Order = 1
2: Project Manager conducts a post deployment review. Order = 2
3: Project Manager conducts knowledge harvesting Order = 3
4: Project Manager decommissions the project Order = 4
5: Project Manager establishes a solution maintenance skeleton Order = 5
6: The project office removes the project from the project list Order = 6

Figure 7. Use Case #5 Close Project
Use Case #6 Post deployment ongoing activities

Main Success scenario

1: Operation management monitors the solution. Order = 1
2: Help desk monitors users’ activities. Order = 1
3: Project manager reviews enhancement requests Order = 1
4: Technical team develops enhancements / bug fixes Order = 1
5: Operations Management deploys enhancements Order = 1

Figure 8. Use Case #6 Post deployment ongoing activities
7 Instrument VI – Comprehension Assessment Questionaire

Q7 What happens if stake-holders change the project scope?
1. The project stops
2. The project manager revisits the project plan
3. Scope creep is addressed in subsequent projects
4. The stakeholders talk with developers who alter their code
5. I do not know
6. The handout does not supply information needed to answer the question.
7. None of the above.

Q8 Who reviews the SCD?
1. People nominated by the project office review the SCD
2. The Project Office review the SCD
3. The SCD is rarely reviewed
4. The project office reviews the SCD, and if the SCD is deemed valid, the project office nominates other approvers
5. I do not know
6. The handout does not supply information needed to answer the question.
7. None of the above.

Q9 What document does the PSD supplement?
1. The PSD supplements the SCD
2. The PSD is a core document? It supplements no other documents.
3. The PSD supplements the HLD and the LLD
4. None of the above
5. I do not know
6. The handout does not supply information needed to answer the question.
7. None of the above.
Q10 How many test cycles does the SDF allow?
1 One at the project end
2 As many as needed until no severe bugs are found or the project is cancelled.
3 Several, one at the end of every project stage
4 None - tests are not done in cycles but end users test all the time
5 I do not know
6 The handout does not supply information needed to answer the question.
7 None of the above.

Q11 What happens immediately after a solution is documented?
1 The testing team tests the solution
2 The customer tests the solution
3 The testing team writes detailed use-cases
4 The customer writes detailed use-cases
5 I do not know
6 The handout does not supply information needed to answer the question.
7 None of the above.

Q12 What condition determines when a build is reiterated?
1 Whether a severe bug has been found
2 Whether a not-severe bug has been found
3 Whether a severe bug has been fixed
4 Whether a build was completed
5 I do not know
6 The handout does not supply information needed to answer the question.
7 None of the above.
8  Instrument VII – Solution expressed as BPMN style flow charts

Research notes

This is another design notation we evaluate. The philosophy of this design notation is that pictures are good at managing complexity (a picture is worth a thousand words).
Use Case #0 - Deliver a SW solution

<table>
<thead>
<tr>
<th>Lead Planner</th>
<th>Project Team</th>
<th>Technical Team</th>
<th>Project Manager</th>
<th>Operations Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Create Strategic Solution Framework

2. SSF
   - analyse requirements and design a solution

3. Build and tests

4. ongoing management review

5. deploys

6. decommission the project

7. conduct ongoing maintenance

Figure 9. Use Case #0 Deliver a SW solution
Use Case #1 Create a Solution Concept Document

<table>
<thead>
<tr>
<th>solution planner</th>
<th>Project office</th>
<th>Approvers</th>
<th>Development Lead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Write a Solution Concept Document</td>
<td>Review the SCD</td>
<td>assign approvers</td>
<td>Insufficient Benefits</td>
</tr>
<tr>
<td>Solution Possible?</td>
<td>SCD needs improvement</td>
<td>The project will be possible only with less demand for resources</td>
<td>Solution Delivery Entry in the company’s projects portfolio</td>
</tr>
<tr>
<td>1a:</td>
<td>2a:</td>
<td>3a:</td>
<td>4a:</td>
</tr>
<tr>
<td>1a1: Solution not feasible</td>
<td>2b: Solution Benefits do not add up</td>
<td>3b: Insufficient Benefits</td>
<td>4b: Solution not feasible</td>
</tr>
<tr>
<td>2:</td>
<td>3:</td>
<td>4:</td>
<td>5:</td>
</tr>
<tr>
<td>5a: Resources are unavailable</td>
<td>5b: Project Cancelled</td>
<td>6: Normal Successful end</td>
<td>SSF</td>
</tr>
</tbody>
</table>

Figure 10. Use Case #1 Create a Solution Concept Document
Use Case #2 Detail analysis and design

<table>
<thead>
<tr>
<th>General Manager</th>
<th>Project Manager</th>
<th>Project team</th>
<th>Technical Lead</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Nominate a Project Manager</td>
<td>Supplement the solution concept</td>
<td></td>
<td>SCD</td>
</tr>
<tr>
<td></td>
<td>PSD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2: Solution unrealistic</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3: Assemble project team</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4: Review project concept document</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5: Solution unrealistic</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6: Write HLD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7: Write technologic assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8: Conduct vendor selection</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9: Write low level design</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10: Solution unrealistic</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11: Write detailed project plan</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 11. Use Case #2 Detail analysis and design
Use case #3.1 Build and Test

<table>
<thead>
<tr>
<th>Use case</th>
<th>Technical Lead</th>
<th>Development team</th>
<th>Testing team</th>
<th>Customer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Write detailed use cases</td>
<td>2</td>
<td>Conduct procurement</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Cost overrun</td>
<td>2a1: Reasonable</td>
<td>2a2: Beyond repair</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Establish development environment</td>
<td>4</td>
<td>Build</td>
<td>4a: Beyond repair</td>
</tr>
<tr>
<td>4</td>
<td>Unhandled extension</td>
<td>4a1: Use Cases cannot be fixed</td>
<td>5</td>
<td>Document the solution</td>
</tr>
<tr>
<td>5</td>
<td>IR</td>
<td>5a: No. or low severity</td>
<td>6</td>
<td>Test</td>
</tr>
<tr>
<td>6</td>
<td>IR</td>
<td>6a: Bugs found</td>
<td>6a1: Learn to use the solution</td>
<td>6a2: Restart use case 2</td>
</tr>
<tr>
<td>6a3: Bug Changes Use case</td>
<td>6a4: Log the bug (But do not fix it)</td>
<td>6a2: Log the bug (But do not fix it)</td>
<td>6a1: Log the bug (But do not fix it)</td>
<td>6a2: Log the bug (But do not fix it)</td>
</tr>
</tbody>
</table>

Figure 12. Use case #3.1 Build and Test
Use case #3.2 Build Monitoring

Project Manager

1. Contact Stakeholder

2. Review Risk Log
   - New Risk?
     - Risk Rectified
     - Risk becomes reality

3. Monitor Milestone Adherence

4. Monitor Budget Compliance

5. weekly status report

Project Board

1. Stakeholder communication

2. Change Project Scope

3. TBA

Stakeholders

1. Continue at use case 3.1 step 2

4. Monitor Budget Compliance

5. Review Status Report

Minutes

Collate Minutes

Project Office

1. TBA

2. NO!

3. Minute Submitted?

4. +

Figure 13. Use case #3.2 Build Monitoring
Use case #4 Deploy the new system

1: Develop a User manual

2: Conduct User Training

2a: New Requirements

2b: Bugs Identified

3: Deploy the solution

3a: Minor Bug

4: Hand over maintenance

4a: Warranty expired

5: Cease Maintaining

Figure 14. Use case #4 Deploy the new system
Use case #5 Project closure

<table>
<thead>
<tr>
<th>Technical Manager</th>
<th>Project Manager</th>
<th>Project Office</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review vendors’ Contract</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conduct Post deployment Review</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conduct knowledge harvesting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decommission project resources</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Establish solution maintenance skeleton</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remove the project from the project list</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 15. Use case #5 Project closure
Use case #6 Post deployment ongoing activities

<table>
<thead>
<tr>
<th>Operation Management</th>
<th>Help Desk</th>
<th>Project Manager</th>
<th>Technical Team</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Monitor the solution</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Monitor users’ activities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Review enhancement Requests</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Enhance / Fix Bugs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Deploy enhancements</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 16. Use case #6 Post deployment ongoing activities
9 Instrument VIII – Comprehension Assessment Questionnaire

Q13 What happens if stake-holders change the project scope?
1 The project stops
2 The project manager revisits the project plan
3 Scope creep is addressed in subsequent projects
4 The stake holders talk with developers who alter their code
5 I do not know
6 The handout does not supply information needed to answer the question.
7 None of the above.

Q14 Who reviews the SCD?
1 People nominated by the project office review the SCD
2 The Project Office review the SCD
3 The SCD is rarely reviewed
4 The project office reviews the SCD, and if the SCD is deemed valid, the project office nominates other approvers
5 I do not know
6 The handout does not supply information needed to answer the question.
7 None of the above.

Q15 What document does the PSD supplement?
1 The PSD supplements the SCD
2 The PSD is a core document? It supplements no other documents.
3 The PSD supplements the HLD and the LLD
4 None of the above
5 I do not know
6 The handout does not supply information needed to answer the question.
7 None of the above.
Q16 **How many test cycles does the SDF allow?**

1. One at the project end
2. As many as needed until no severe bugs are found or the project is cancelled.
3. Several, one at the end of every project stage
4. None - tests are not done in cycles but end users test all the time
5. I do not know
6. The handout does not supply information needed to answer the question.
7. None of the above.

Q17 **What happens immediately after a solution is documented?**

1. The testing team tests the solution
2. The customer tests the solution
3. The testing team writes detailed use-cases
4. The customer writes detailed use-cases
5. I do not know
6. The handout does not supply information needed to answer the question.
7. None of the above.

Q18 **What condition determines when a build is reiterated?**

1. Whether a severe bug has been found
2. Whether a not-severe bug has been found
3. Whether a severe bug has been fixed
4. Whether a build was completed
5. I do not know
6. The handout does not supply information needed to answer the question.
7. None of the above.
10 Instrument IX – Preference Survey

Q19 When asking for directions, would you prefer to receive:
A a map
B verbal instructions
C neither, or not applicable

Q20 When asking for directions, you most often receive:
A a map
B verbal instructions
C neither, or not applicable

Q21 When your team is asked to develop software, would you prefer:
A to design the solution yourself, and hand the design to a team member for coding.
B to receive a design from a team member, and code the solution yourself.
C neither, or not applicable

Q22 When your team is asked to develop software, most often you:
A design the solution yourself, and hand the design to a team member for coding.
B receive a design from a team member, and code the solution yourself.
C neither, or not applicable

Q23 The number of fiction books you read last 12 month is:
A 0
B 1-3
C 4-6
D 7-10
E more than 10
Q24 You are comfortable with flow charts:
A Strongly agree
B Somewhat agree
C I do not know
D Somewhat disagree
E Strongly disagree

Q25 You often work with flow charts:
A Strongly agree
B Somewhat agree
C I do not know
D Somewhat disagree
E Strongly disagree

Q26 You are comfortable with written use cases:
A Strongly agree
B Somewhat agree
C I do not know
D Somewhat disagree
E Strongly disagree

Q27 You often work with written use cases:
A Strongly agree
B Somewhat agree
C I do not know
D Somewhat disagree
E Strongly disagree
11 Instrument X – Demographic Survey

Q28 What is your gender?:
A female
B male
C I do not wish to answer this question

Q29 Which role best describes your occupation (Select all that apply)?
A student
B professional
C academic
D I do not wish to answer this question

Q30 Is English your native language?
A Yes.
B No.
C I do not wish to answer this question.

Q31 Do you have work experience related to Business Process Management and/or Modelling?
A Yes.
B No.
C I do not wish to answer this question.

Q32 Do you have any knowledge related to the software delivery life cycle?:
A Yes.
B No.
C I do not wish to answer this question.

Q33 Do you have any knowledge related to the financial industry?:
A Yes.
B No.
C I do not wish to answer this question.
Thank you for your participation