

Project Report

Tim, Xiaofan, Yanwen, Jingning

11/20/2019

Introduction

Client is investigating how foreign language teachers feel about and utilize methods from the Teacher Effectiveness for Language Learning (TELL), and seeking advice about how to improve the survey.

Our purpose for our client in this project: 1. A lot of people don't answer the survey because it's long. Can we reduce the number of questions? 2. Is the survey currently answering the research questions?

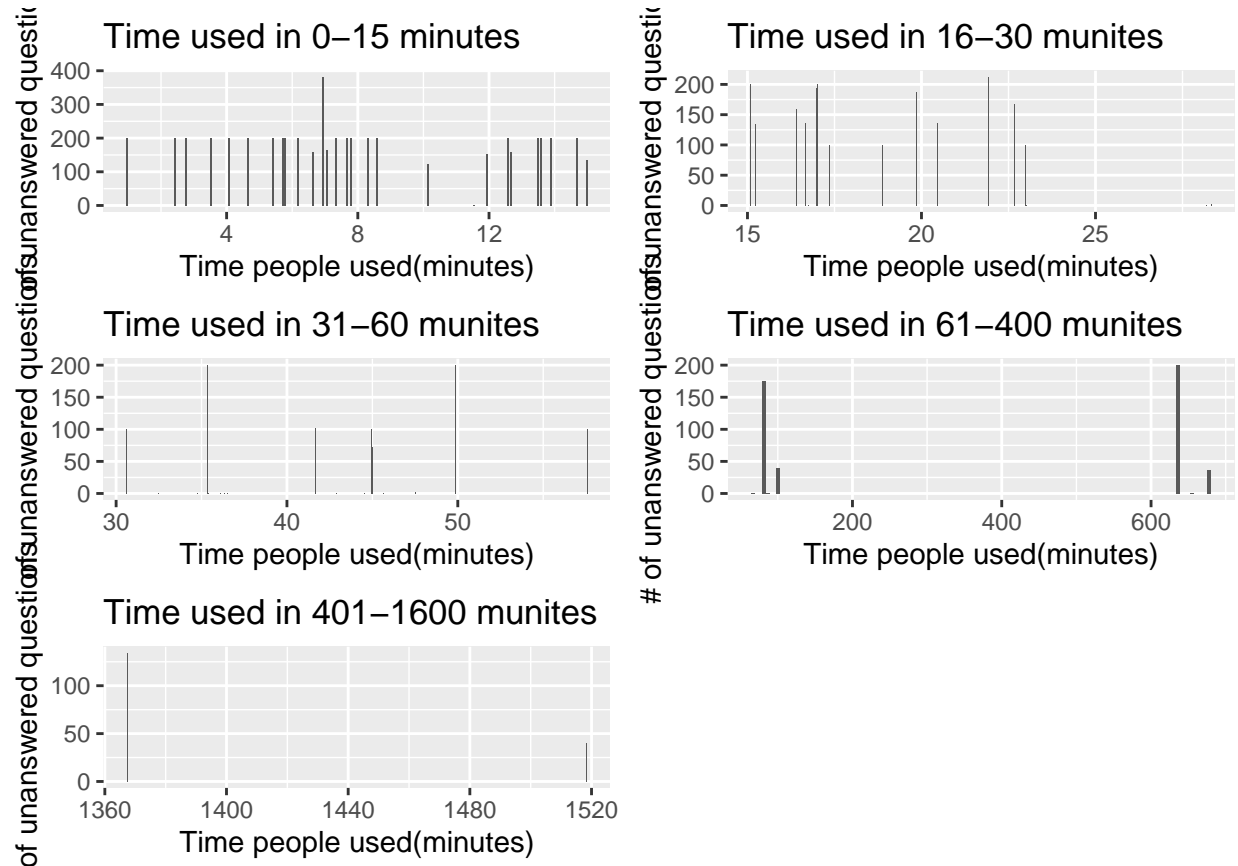
EDA & Concerns

Data Structure

We are provided the data in an excel file with 6 spreadsheets including one sheet of notes, one sheet of personal information and 4 sheets of Teacher Effectiveness for Language Learning (TELL) framework survey questions. The dataset of personal information contains questions regarding respondents' teaching language and education background. The dataset of Teacher Effectiveness for Language Learning (TELL) framework survey contains around 200 questions asking about respondents' attitudes of contribution and confidence towards each practice in the framework. There are 4 domains of Teacher Effectiveness for Language Learning (TELL) survey questions: planning, learning experience, learning tools, and performance & feedback. Each domain contains several subdomains and each subdomain contains a different number of questions. For this project, we focus on reducing the number of questions in the dataset of Teacher Effectiveness for Language Learning (TELL) framework survey.

EDA

We conduct a basic Exploratory Data Analysis (EDA) for this project. Firstly, We focus on the time for respondents to complete this survey.



Data Cleaning

Concerns

We come up with several concerns after the initial EDA. Firstly, the observations we can use in the analysis are very limited since there are many N/A in the dataset. Secondly, some respondents seem like choosing the same answer through the whole survey and if we identify these answers as non-valid, then our sample size would become even smaller. With this limited sample size, the accuracy and reference of results from our subsequent analysis could be affected.

Method we used

We will use Confirmatory Factor Analysis (CFA) to reduce the survey questions number. CFA is a special form of factor analysis and mostly used in social science research. It is used to check whether measures of a construct are consistent with a researcher's understanding of the nature of that construct.

Here we will use CFA to see if there are survey questions equivalent to each other so we can reduce those repeated questions. We will analysis each subdomain separately and will only consider the problems regarding confidence or not. Within each subdomain, there will be several questions and our null hypothesis is that all survey questions are identical to each other. Then our alternative hypothesis is that the questions are not all equal.

We will focus on the p value result we have from CFA and we will take a p value larger than 0.05 to reject our null hypothesis. When we are not able to reject our null hypothesis, we will look at our factor loading to check the correlations between questions. Then we will fit new model by dropping question with lowest factor loading and see if we will reject our null hypothesis now. We will keep doing this until we have a subdomain with an acceptable p value, which gives us a set of survey questions are not identical to each other.

Analysis

Learning Tool Domain Analysis

```
## lavaan 0.6-5 ended normally after 12 iterations
##
##      Estimator                      ML
##      Optimization method          NLMINB
##      Number of free parameters      6
##      Number of equality constraints    1
##      Row rank of the constraints matrix 1
##
##                                     Used      Total
##      Number of observations          27        84
##
## Model Test User Model:
##
##      Test statistic                  0.109
##      Degrees of freedom              1
##      P-value (Chi-square)            0.741
##
## Parameter Estimates:
##
##      Information                    Expected
##      Information saturated (h1) model Structured
##      Standard errors                Standard
##
## Latent Variables:
##
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all
##      lt1 =~
##      LT1a_Cnfdn (aa)    0.384    0.165    2.331    0.020    0.384    0.488
##      LT1b_Cnfd (aa)    0.384    0.165    2.331    0.020    0.384    0.472
##      LT1c_Cnfd         0.612    0.282    2.167    0.030    0.612    0.722
##
## Variances:
##
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all
##      .LT1a_Confidenc    0.472    0.168    2.816    0.005    0.472    0.762
##      .LT1b_Confidenc    0.515    0.177    2.914    0.004    0.515    0.777
##      .LT1c_Confidenc    0.344    0.314    1.095    0.273    0.344    0.479
##      lt1                1.000                1.000    1.000
##
##      npar      fmin      chisq
##      5.000      0.002      0.109
##      df      pvalue      baseline.chisq
##      1.000      0.741      7.404
##      baseline.df      baseline.pvalue      cfi
##      3.000      0.060      1.000
##      tli      nnfi      rfi
##      1.607      1.607      0.956
##      nfi      pnfi      ifi
##      0.985      0.328      1.139
##      rni      logl      unrestricted.logl
##      1.202      -94.804      -94.750
##      aic      bic      ntotal
##      199.609      206.088      27.000
```

```

##          bic2          rmsea      rmsea.ci.lower
##        190.555          0.000          0.000
##      rmsea.ci.upper      rmsea.pvalue          rmr
##        0.355          0.749          0.018
##      rmr_nomean          srmr          srmr_bentler
##        0.018          0.028          0.028
## srmr_bentler_nomean          crmr          crmr_nomean
##        0.028          0.028          0.028
##      srmr_mplus      srmr_mplus_nomean          cn_05
##        0.026          0.026          951.942
##        cn_01          gfi          agfi
##      1643.449          0.997          0.984
##        pgfi          mfi          ecvi
##        0.166          1.017          0.374

##          lhs op          rhs      mi      epc sepc.lv sepc.all
## 10 LT1a_Confidence ~~ LT1c_Confidence 0.109 0.053 0.053 0.133
## 11 LT1b_Confidence ~~ LT1c_Confidence 0.109 -0.053 -0.053 -0.127
##      sepc.nox
## 10      0.133
## 11     -0.127

```

Table 1: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
lt1	LT1a_Confidence	0.384	0.165	2.331	0.02	0.488
lt1	LT1b_Confidence	0.384	0.165	2.331	0.02	0.472
lt1	LT1c_Confidence	0.612	0.282	2.167	0.03	0.722

```

## lavaan 0.6-5 ended normally after 12 iterations
##
##      Estimator          ML
##      Optimization method      NLMINB
##      Number of free parameters          6
##      Number of equality constraints      1
##      Row rank of the constraints matrix  1
##
##                                Used      Total
##      Number of observations          28      84
##
## Model Test User Model:
##
##      Test statistic          0.003
##      Degrees of freedom          1
##      P-value (Chi-square)      0.953
##
## Parameter Estimates:
##
##      Information          Expected
##      Information saturated (h1) model      Structured
##      Standard errors          Standard
##
## Latent Variables:
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all

```

```

##      lt2 =~
##      LT2_Cnfdn (aa)    0.443    0.126    3.531    0.000    0.443    0.587
##      LT2b_Cnfd (aa)    0.443    0.126    3.531    0.000    0.443    0.603
##      LT2c_Cnfd        0.776    0.222    3.499    0.000    0.776    0.817
##
## Variances:
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all
##      .LT2a_Confidenc 0.373    0.127    2.930    0.003    0.373    0.655
##      .LT2b_Confidenc 0.345    0.122    2.838    0.005    0.345    0.637
##      .LT2c_Confidenc 0.300    0.270    1.110    0.267    0.300    0.333
##      lt2            1.000
##
##      npar          fmin          chisq
##      5.000          0.000          0.003
##      df            pvalue      baseline.chisq
##      1.000          0.953          15.764
##      baseline.df    baseline.pvalue      cfi
##      3.000          0.001          1.000
##      tli            nnfi          rfi
##      1.234          1.234          0.999
##      nfi            pnfi          ifi
##      1.000          0.333          1.067
##      rni            logl    unrestricted.logl
##      1.078          -93.396          -93.394
##      aic            bic          ntotal
##      196.792          203.453          28.000
##      bic2            rmsea      rmsea.ci.lower
##      187.908          0.000          0.000
##      rmsea.ci.upper    rmsea.pvalue      rmr
##      0.000          0.955          0.003
##      rmr_nomean      srmr          srmr_bentler
##      0.003          0.005          0.005
##      srmr_bentler_nomean      crmr          crmr_nomean
##      0.005          0.004          0.004
##      srmr_mplus      srmr_mplus_nomean      cn_05
##      0.005          0.005          31204.347
##      cn_01            gfi          agfi
##      53894.843          1.000          1.000
##      pgfi            mfi          ecvi
##      0.167          1.018          0.357
##
##      lhs op          rhs      mi      epc sepc.lv sepc.all
## 10 LT2a_Confidence ~~ LT2c_Confidence 0.003 -0.009 -0.009 -0.027
## 11 LT2b_Confidence ~~ LT2c_Confidence 0.003 0.009 0.009 0.028
##      sepc.nox
## 10      -0.027
## 11      0.028

```

Table 2: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
lt2	LT2a_Confidence	0.443	0.126	3.531	0	0.587
lt2	LT2b_Confidence	0.443	0.126	3.531	0	0.603
lt2	LT2c_Confidence	0.776	0.222	3.499	0	0.817

```

## lavaan 0.6-5 ended normally after 14 iterations
##
## Estimator ML
## Optimization method NLMINB
## Number of free parameters 6
## Number of equality constraints 1
## Row rank of the constraints matrix 1
##
## Used Total
## Number of observations 27 84
##
## Model Test User Model:
##
## Test statistic 0.017
## Degrees of freedom 1
## P-value (Chi-square) 0.897
##
## Parameter Estimates:
##
## Information Expected
## Information saturated (h1) model Structured
## Standard errors Standard
##
## Latent Variables:
## Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## lt3 =~
## LT3_Cnfdn 0.931 0.164 5.660 0.000 0.931 0.959
## LT3b_Cnfd (aa) 0.706 0.144 4.892 0.000 0.706 0.802
## LT3d_Cnfd (aa) 0.706 0.144 4.892 0.000 0.706 0.629
##
## Variances:
## Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## .LT3a_Confidenc 0.075 0.170 0.442 0.659 0.075 0.080
## .LT3b_Confidenc 0.276 0.122 2.259 0.024 0.276 0.357
## .LT3d_Confidenc 0.761 0.228 3.330 0.001 0.761 0.604
## lt3 1.000 1.000 1.000
##
## npar fmin chisq
## 5.000 0.000 0.017
## df pvalue baseline.chisq
## 1.000 0.897 36.819
## baseline.df baseline.pvalue cfi
## 3.000 0.000 1.000
## tli nnfi rfi
## 1.087 1.087 0.999
## nfi pnfi ifi
## 1.000 0.333 1.027
## rni logl unrestricted.logl
## 1.029 -95.478 -95.469
## aic bic ntotal
## 200.955 207.434 27.000
## bic2 rmsea rmsea.ci.lower
## 191.902 0.000 0.000
## rmsea.ci.upper rmsea.pvalue rmr

```

```

##          0.233          0.900          0.014
##      rmr_nomean      srmr      srmr_bentler
##          0.014          0.012          0.012
## srmr_bentler_nomean      crmr      crmr_nomean
##          0.012          0.007          0.007
##      srmr_mplus      srmr_mplus_nomean      cn_05
##          0.011          0.011          6147.974
##          cn_01          gfi          agfi
##      10617.940          1.000          0.998
##          pgfi          mfi          ecvi
##          0.167          1.018          0.371

##          lhs op          rhs      mi      epc sepc.lv sepc.all
## 9  LT3a_Confidence ~~ LT3b_Confidence 0.017 -0.025 -0.025 -0.171
## 10 LT3a_Confidence ~~ LT3d_Confidence 0.017  0.025  0.025  0.103
##      sepc.nox
## 9      -0.171
## 10      0.103

```

Table 3: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
lt3	LT3a_Confidence	0.931	0.164	5.660	0	0.959
lt3	LT3b_Confidence	0.706	0.144	4.892	0	0.802
lt3	LT3d_Confidence	0.706	0.144	4.892	0	0.629

```

## lavaan 0.6-5 ended normally after 19 iterations
##
##      Estimator                      ML
##      Optimization method          NLMINB
##      Number of free parameters      6
##      Number of equality constraints   1
##      Row rank of the constraints matrix 1
##
##                                Used      Total
##      Number of observations          28      84
##
## Model Test User Model:
##
##      Test statistic                  0.016
##      Degrees of freedom              1
##      P-value (Chi-square)            0.899
##
## Parameter Estimates:
##
##      Information                    Expected
##      Information saturated (h1) model Structured
##      Standard errors                Standard
##
## Latent Variables:
##      Estimate  Std.Err  z-value  P(>|z|)  Std.lv  Std.all
##      lt4 =~
##      LT4_Cnfdn      0.370   0.140   2.640   0.008   0.370   0.469
##      LT4b_Cnfd (aa)  1.077   0.150   7.191   0.000   1.077   1.012

```



```

##      LT4c_Cnfd (aa)      1.077      0.150      7.191      0.000      1.077      0.886
##
## Variances:
##              Estimate Std.Err z-value P(>|z|) Std.lv Std.all
##      .LT4a_Confidence  0.485   0.129   3.749   0.000   0.485   0.780
##      .LT4b_Confidence -0.027   0.095  -0.282   0.778  -0.027  -0.024
##      .LT4c_Confidence  0.316   0.127   2.486   0.013   0.316   0.214
##      lt4                1.000
##
##              npar              fmin              chisq
##              5.000              0.000              0.016
##              df              pvalue      baseline.chisq
##              1.000              0.899              52.579
##      baseline.df      baseline.pvalue              cfi
##              3.000              0.000              1.000
##              tli              nnfi              rfi
##              1.060              1.060              0.999
##              nfi              pnfi              ifi
##              1.000              0.333              1.019
##              rni              logl      unrestricted.logl
##              1.020             -93.343             -93.335
##              aic              bic              ntotal
##              196.686             203.347              28.000
##              bic2              rmsea      rmsea.ci.lower
##              187.802              0.000              0.000
##      rmsea.ci.upper      rmsea.pvalue              rmr
##              0.225              0.902              0.008
##              rmr_nomean              srmr      srmr_bentler
##              0.008              0.007              0.007
##      srmr_bentler_nomean              crmr      crmr_nomean
##              0.007              0.006              0.006
##              srmr_mplus      srmr_mplus_nomean              cn_05
##              0.006              0.006             6648.164
##              cn_01              gfi              agfi
##              11481.859              1.000              0.998
##              pgfi              mfi              ecvi
##              0.167              1.018              0.358
##
##              lhs op              rhs      mi      epc sepc.lv sepc.all
## 9  LT4a_Confidence ~~ LT4b_Confidence 0.016  0.01   0.01   0.089
## 10 LT4a_Confidence ~~ LT4c_Confidence 0.016 -0.01  -0.01  -0.026
##      sepc.nox
## 9      0.089
## 10     -0.026

```

Table 4: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
lt4	LT4a_Confidence	0.370	0.14	2.640	0.008	0.469
lt4	LT4b_Confidence	1.077	0.15	7.191	0.000	1.012
lt4	LT4c_Confidence	1.077	0.15	7.191	0.000	0.886

```

## lavaan 0.6-5 ended normally after 13 iterations
##

```

```

## Estimator ML
## Optimization method NLMINB
## Number of free parameters 6
## Number of equality constraints 1
## Row rank of the constraints matrix 1
##
## Used Total
## Number of observations 26 84
##
## Model Test User Model:
##
## Test statistic 0.774
## Degrees of freedom 1
## P-value (Chi-square) 0.379
##
## Parameter Estimates:
##
## Information Expected
## Information saturated (h1) model Structured
## Standard errors Standard
##
## Latent Variables:
## Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## lt5 =~
## LT5_Cnfdn (aa) 0.398 0.124 3.205 0.001 0.398 0.618
## LT5b_Cnfd 0.620 0.186 3.331 0.001 0.620 0.844
## LT5c_Cnfd (aa) 0.398 0.124 3.205 0.001 0.398 0.532
##
## Variances:
## Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## .LT5a_Confidenc 0.257 0.101 2.544 0.011 0.257 0.618
## .LT5b_Confidenc 0.155 0.186 0.832 0.405 0.155 0.287
## .LT5c_Confidenc 0.402 0.132 3.034 0.002 0.402 0.717
## lt5 1.000 1.000 1.000
##
## npar fmin chisq
## 5.000 0.015 0.774
## df pvalue baseline.chisq
## 1.000 0.379 15.415
## baseline.df baseline.pvalue cfi
## 3.000 0.001 1.000
## tli nnfi rfi
## 1.055 1.055 0.849
## nfi pnfi ifi
## 0.950 0.317 1.016
## rni logl unrestricted.logl
## 1.018 -76.441 -76.054
## aic bic ntotal
## 162.882 169.172 26.000
## bic2 rmsea rmsea.ci.lower
## 153.653 0.000 0.000
## rmsea.ci.upper rmsea.pvalue rmr
## 0.494 0.394 0.041
## rmr_nomean srmer srmer_bentler

```

```

##          0.041          0.079          0.079
## srmr_bentler_nomean          crmr          crmr_nomean
##          0.079          0.063          0.063
##          srmr_mplus          srmr_mplus_nomean          cn_05
##          0.070          0.070          130.124
##          cn_01          gfi          agfi
##          224.020          0.981          0.885
##          pgfi          mfi          ecvi
##          0.163          1.004          0.414

##          lhs op          rhs          mi          epc sepc.lv sepc.all
## 9  LT5a_Confidence ~~ LT5b_Confidence 0.762 -0.102 -0.102 -0.512
## 11 LT5b_Confidence ~~ LT5c_Confidence 0.762 0.102 0.102 0.409
##          sepc.nox
## 9          -0.512
## 11          0.409

```

Table 5: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
lt5	LT5a_Confidence	0.398	0.124	3.205	0.001	0.618
lt5	LT5b_Confidence	0.620	0.186	3.331	0.001	0.844
lt5	LT5c_Confidence	0.398	0.124	3.205	0.001	0.532

PER & FEEDBACK Domain Analysis

For PER&FEEDBACK table in TELL Statements, I numeric character answers of PF 1a~5c Confidence, and NA values stay as same as NA that will not count in. First, I made CFA models for each subdomain whose variables should greater than 2 (ex: PF1 has 5 variables: PF1a_Confidence, PF1b_Confidence, PF1c_Confidence, PF1d_Confidence and PF1e_Confidence), or the P-value of that model will become NA. And we get an exception in PF table: PF4 only has 2 variables, so I combine PF4 with PF5 to one CFA model so that we have an available P-value. Second, we find factor loadings of each variables in each subdomain and record them. Third, we compare P-value of each subdomain to 0.05, if P-value > 0.05, our null hypothesis retained, and we do not need to make any further change on that subdomain; if P-value < 0.05, it means our null hypothesis is rejected, and we need to remodel by dropping the variable with lowest factor loadings in that subdomain and check its P-value again. Following are detailed results

First subdomain:

```

## lavaan 0.6-5 ended normally after 21 iterations
##
## Estimator          ML
## Optimization method          NLMINB
## Number of free parameters          10
##
##          Used          Total
## Number of observations          27          84
##
## Model Test User Model:
##
## Test statistic          15.646
## Degrees of freedom          5
## P-value (Chi-square)          0.008
##
## Parameter Estimates:

```

```

##
## Information Expected
## Information saturated (h1) model Structured
## Standard errors Standard
##
## Latent Variables:
## Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## PF1 =~
## PF1a_Confide nc 0.690 0.202 3.421 0.001 0.690 0.609
## PF1b_Confide nc 0.879 0.168 5.229 0.000 0.879 0.830
## PF1c_Confide nc 0.828 0.128 6.471 0.000 0.828 0.946
## PF1d_Confide nc 0.823 0.135 6.110 0.000 0.823 0.915
## PF1e_Confide nc 0.584 0.178 3.275 0.001 0.584 0.587
##
## Variances:
## Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## .PF1a_Confide nc 0.808 0.228 3.545 0.000 0.808 0.630
## .PF1b_Confide nc 0.349 0.110 3.160 0.002 0.349 0.311
## .PF1c_Confide nc 0.080 0.049 1.628 0.104 0.080 0.105
## .PF1d_Confide nc 0.132 0.057 2.306 0.021 0.132 0.163
## .PF1e_Confide nc 0.647 0.182 3.559 0.000 0.647 0.655
## PF1 1.000 1.000

```

Table 6: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
PF1	PF1a_Confidence	0.690	0.202	3.421	0.001	0.609
PF1	PF1b_Confidence	0.879	0.168	5.229	0.000	0.830
PF1	PF1c_Confidence	0.828	0.128	6.471	0.000	0.946
PF1	PF1d_Confidence	0.823	0.135	6.110	0.000	0.915
PF1	PF1e_Confidence	0.584	0.178	3.275	0.001	0.587

Since p-value of first subdomain is $0.008 < 0.05$, and the factor loadings of “PF1e_Confidence” is lowest, thus, we try to drop it from the first subdomain:

```

## lavaan 0.6-5 ended normally after 31 iterations
##
## Estimator ML
## Optimization method NLMINB
## Number of free parameters 8
##
## Used Total
## Number of observations 28 84
##
## Model Test User Model:
##
## Test statistic 0.068
## Degrees of freedom 2
## P-value (Chi-square) 0.967
##
## Parameter Estimates:
##
## Information Expected

```

```

## Information saturated (h1) model      Structured
## Standard errors                      Standard
##
## Latent Variables:
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## PF1 =~
## PF1a_Confidenc 0.660 0.196 3.360 0.001 0.660 0.593
## PF1b_Confidenc 0.835 0.173 4.830 0.000 0.835 0.780
## PF1c_Confidenc 0.796 0.130 6.135 0.000 0.796 0.914
## PF1d_Confidenc 0.831 0.129 6.432 0.000 0.831 0.940
##
## Variances:
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## .PF1a_Confidenc 0.804 0.223 3.601 0.000 0.804 0.649
## .PF1b_Confidenc 0.449 0.135 3.321 0.001 0.449 0.392
## .PF1c_Confidenc 0.125 0.061 2.048 0.041 0.125 0.165
## .PF1d_Confidenc 0.091 0.061 1.498 0.134 0.091 0.116
## PF1 1.000 1.000 1.000 1.000

```

P-value = 0.967 > 0.05, thus we do not need to change any more on the first subdomain.

Second subdomain:

```

## lavaan 0.6-5 ended normally after 18 iterations
##
## Estimator ML
## Optimization method NLMINB
## Number of free parameters 10
##
## Used Total
## Number of observations 27 84
##
## Model Test User Model:
##
## Test statistic 14.489
## Degrees of freedom 5
## P-value (Chi-square) 0.013
##
## Parameter Estimates:
##
## Information Expected
## Information saturated (h1) model Structured
## Standard errors Standard
##
## Latent Variables:
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## PF2 =~
## PF2a_Confidenc 0.561 0.172 3.255 0.001 0.561 0.587
## PF2b_Confidenc 0.948 0.159 5.981 0.000 0.948 0.905
## PF2c_Confidenc 0.575 0.188 3.060 0.002 0.575 0.558
## PF2d_Confidenc 0.896 0.151 5.941 0.000 0.896 0.901
## PF2e_Confidenc 1.016 0.173 5.880 0.000 1.016 0.896
##
## Variances:

```

##		Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
##	.PF2a_Confidenc	0.599	0.169	3.534	0.000	0.599	0.655
##	.PF2b_Confidenc	0.199	0.083	2.384	0.017	0.199	0.181
##	.PF2c_Confidenc	0.731	0.206	3.554	0.000	0.731	0.689
##	.PF2d_Confidenc	0.185	0.076	2.439	0.015	0.185	0.188
##	.PF2e_Confidenc	0.255	0.101	2.520	0.012	0.255	0.198
##	PF2	1.000				1.000	1.000

Table 7: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
PF2	PF2a_Confidence	0.561	0.172	3.255	0.001	0.587
PF2	PF2b_Confidence	0.948	0.159	5.981	0.000	0.905
PF2	PF2c_Confidence	0.575	0.188	3.060	0.002	0.558
PF2	PF2d_Confidence	0.896	0.151	5.941	0.000	0.901
PF2	PF2e_Confidence	1.016	0.173	5.880	0.000	0.896

Since p-value of first subdomain is $0.013 < 0.05$, and the factor loadings of “PF2c_Confidence” is lowest, thus, we try to drop it from the second subdomain:

```
## lavaan 0.6-5 ended normally after 18 iterations
##
## Estimator ML
## Optimization method NLMINB
## Number of free parameters 8
##
## Used Total
## Number of observations 28 84
##
## Model Test User Model:
##
## Test statistic 1.559
## Degrees of freedom 2
## P-value (Chi-square) 0.459
##
## Parameter Estimates:
##
## Information Expected
## Information saturated (h1) model Structured
## Standard errors Standard
##
## Latent Variables:
## Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## PF2 =~
## PF2a_Confidenc 0.553 0.169 3.275 0.001 0.553 0.583
## PF2b_Confidenc 0.910 0.158 5.765 0.000 0.910 0.877
## PF2d_Confidenc 0.884 0.147 6.012 0.000 0.884 0.901
## PF2e_Confidenc 1.011 0.167 6.057 0.000 1.011 0.905
##
## Variances:
## Estimate Std.Err z-value P(>|z|) Std.lv Std.all
## .PF2a_Confidenc 0.596 0.166 3.588 0.000 0.596 0.660
## .PF2b_Confidenc 0.248 0.093 2.675 0.007 0.248 0.231
```

```
##      .PF2d_Confidenc    0.182    0.077    2.349    0.019    0.182    0.189
##      .PF2e_Confidenc    0.227    0.099    2.280    0.023    0.227    0.181
##      PF2                1.000                                1.000    1.000
```

P-value = 0.459 > 0.05, thus we can stay here for the second subdomain.

Third subdomain:

```
## lavaan 0.6-5 ended normally after 15 iterations
##
##      Estimator                      ML
##      Optimization method          NLMINB
##      Number of free parameters      10
##
##                                Used      Total
##      Number of observations         28        84
##
## Model Test User Model:
##
##      Test statistic                2.920
##      Degrees of freedom              5
##      P-value (Chi-square)           0.712
##
## Parameter Estimates:
##
##      Information                      Expected
##      Information saturated (h1) model  Structured
##      Standard errors                  Standard
##
## Latent Variables:
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all
##      PF3 =~
##      PF3a_Confidenc    0.466    0.184    2.533    0.011    0.466    0.485
##      PF3b_Confidenc    0.869    0.173    5.032    0.000    0.869    0.838
##      PF3c_Confidenc    0.628    0.146    4.300    0.000    0.628    0.746
##      PF3d_Confidenc    0.753    0.169    4.468    0.000    0.753    0.767
##      PF3e_Confidenc    0.543    0.171    3.179    0.001    0.543    0.588
##
## Variances:
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all
##      .PF3a_Confidenc    0.706    0.199    3.543    0.000    0.706    0.765
##      .PF3b_Confidenc    0.320    0.149    2.153    0.031    0.320    0.298
##      .PF3c_Confidenc    0.315    0.109    2.893    0.004    0.315    0.444
##      .PF3d_Confidenc    0.396    0.143    2.759    0.006    0.396    0.411
##      .PF3e_Confidenc    0.557    0.164    3.396    0.001    0.557    0.654
##      PF3                1.000                                1.000    1.000
```

Table 8: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
PF3	PF3a_Confidence	0.466	0.184	2.533	0.011	0.485
PF3	PF3b_Confidence	0.869	0.173	5.032	0.000	0.838
PF3	PF3c_Confidence	0.628	0.146	4.300	0.000	0.746
PF3	PF3d_Confidence	0.753	0.169	4.468	0.000	0.767

Latent Factor	Indicator	B	SE	Z	p-value	loading
PF3	PF3e_Confidence	0.543	0.171	3.179	0.001	0.588
Since p-value >	0.05, the third subdomain is ok, no longer to remodel it.					

Fourth subdomain:

PF4 only has 2 variables, so I combine PF4 with PF5 to one CFA model so that we can get an available P-value.

```
## lavaan 0.6-5 ended normally after 22 iterations
##
##      Estimator                      ML
##      Optimization method          NLMINB
##      Number of free parameters      11
##
##                               Used      Total
##      Number of observations          24        84
##
## Model Test User Model:
##
##      Test statistic                12.824
##      Degrees of freedom              4
##      P-value (Chi-square)           0.012
##
## Parameter Estimates:
##
##      Information                      Expected
##      Information saturated (h1) model Structured
##      Standard errors                  Standard
##
## Latent Variables:
##              Estimate Std.Err z-value P(>|z|) Std.lv Std.all
##      PF4 =~
##      PF4a_Confidence 0.754 0.148 5.099 0.000 0.754 0.888
##      PF4b_Confidence 0.681 0.145 4.699 0.000 0.681 0.835
##      PF5 =~
##      PF5a_Confidence 0.690 0.173 3.988 0.000 0.690 0.757
##      PF5b_Confidence 0.632 0.158 4.003 0.000 0.632 0.759
##      PF5c_Confidence 0.535 0.192 2.782 0.005 0.535 0.567
##
## Covariances:
##              Estimate Std.Err z-value P(>|z|) Std.lv Std.all
##      PF4 ~~
##      PF5              0.877 0.113 7.732 0.000 0.877 0.877
##
## Variances:
##              Estimate Std.Err z-value P(>|z|) Std.lv Std.all
##      .PF4a_Confidence 0.153 0.101 1.511 0.131 0.153 0.212
##      .PF4b_Confidence 0.201 0.094 2.131 0.033 0.201 0.302
##      .PF5a_Confidence 0.355 0.143 2.484 0.013 0.355 0.427
##      .PF5b_Confidence 0.294 0.119 2.470 0.014 0.294 0.424
##      .PF5c_Confidence 0.603 0.192 3.135 0.002 0.603 0.678
##      PF4              1.000                1.000 1.000
```



```
##      PF5                1.000                1.000      1.000
```

Table 9: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
PF4	PF4a_Confidence	0.754	0.148	5.099	0.000	0.888
PF4	PF4b_Confidence	0.681	0.145	4.699	0.000	0.835
PF5	PF5a_Confidence	0.690	0.173	3.988	0.000	0.757
PF5	PF5b_Confidence	0.632	0.158	4.003	0.000	0.759
PF5	PF5c_Confidence	0.535	0.192	2.782	0.005	0.567

Since P-value is $0.012 < 0.05$, and the lowest factor loading is “PF5c_Confidence”, thus we try to drop it from the subdomain:

```
## lavaan 0.6-5 ended normally after 21 iterations
##
##      Estimator                ML
##      Optimization method      NLMINB
##      Number of free parameters      9
##
##                                Used      Total
##      Number of observations        24        84
##
## Model Test User Model:
##
##      Test statistic                0.832
##      Degrees of freedom              1
##      P-value (Chi-square)           0.362
##
## Parameter Estimates:
##
##      Information                Expected
##      Information saturated (h1) model      Structured
##      Standard errors              Standard
##
## Latent Variables:
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all
##      PF4 =~
##      PF4a_Confidenc      0.724      0.155      4.671      0.000      0.724      0.851
##      PF4b_Confidenc      0.710      0.148      4.808      0.000      0.710      0.871
##      PF5 =~
##      PF5a_Confidenc      0.812      0.171      4.742      0.000      0.812      0.890
##      PF5b_Confidenc      0.635      0.160      3.970      0.000      0.635      0.763
##
## Covariances:
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all
##      PF4 ~~
##      PF5                0.755      0.137      5.507      0.000      0.755      0.755
##
## Variances:
##      Estimate Std.Err z-value P(>|z|) Std.lv Std.all
##      .PF4a_Confidenc      0.199      0.116      1.720      0.085      0.199      0.275
##      .PF4b_Confidenc      0.160      0.107      1.497      0.134      0.160      0.241
```

##	.PF5a_Confidenc	0.173	0.157	1.100	0.271	0.173	0.207
##	.PF5b_Confidenc	0.289	0.124	2.342	0.019	0.289	0.418
##	PF4	1.000				1.000	1.000
##	PF5	1.000				1.000	1.000

P-value is 0.362 > 0.05, thus no longer remodel this subdomain.

Learning Experience Domain Analysis

For learning experience table in TELL Statements, we numeric character answers of LE 1a~6d Confidence, and NA values stay as same as NA that will not count in. First, I made CFA models for each subdomain (ex: LE1 has 5 variables: LE1a_Confidence, LE1b_Confidence, LE1c_Confidence, LE1d_Confidence and LE1e_Confidence). Then we have an available P-value for each subdomain and we find factor loadings of each variables in each subdomain. Third, we compare P-value of each subdomain to 0.05, if P-value > 0.05, our null hypothesis retained, and we do not need to make any further change on that subdomain; if P-value < 0.05, it means our null hypothesis is rejected, and we need to remodel by dropping the variable with lowest factor loadings in that subdomain and check its P-value again. Following are detailed results

First subdomian

```
## lavaan 0.6-5 ended normally after 28 iterations
##
##      Estimator                      ML
##      Optimization method          NLMINB
##      Number of free parameters      10
##
##                                     Used      Total
##      Number of observations          29         84
##
## Model Test User Model:
##
##      Test statistic                  2.594
##      Degrees of freedom              5
##      P-value (Chi-square)            0.762
##
## Parameter Estimates:
##
##      Information                    Expected
##      Information saturated (h1) model Structured
##      Standard errors                Standard
##
## Latent Variables:
##
##              Estimate Std.Err z-value P(>|z|)
##      LE1 =~
##      LE1a_Confidenc    1.000
##      LE1b_Confidenc    0.601    0.230    2.619    0.009
##      LE1c_Confidenc    0.837    0.298    2.812    0.005
##      LE1d_Confidenc    0.589    0.236    2.495    0.013
##      LE1e_Confidenc    0.351    0.215    1.632    0.103
##
## Variances:
##
##              Estimate Std.Err z-value P(>|z|)
##      .LE1a_Confidenc    0.690    0.699    0.987    0.324
##      .LE1b_Confidenc    2.142    0.632    3.391    0.001
##      .LE1c_Confidenc    3.173    0.985    3.221    0.001
```

```
##      .LE1d_Confidenc    2.411    0.696    3.465    0.001
##      .LE1e_Confidenc    2.590    0.698    3.711    0.000
##      LE1                2.646    1.091    2.424    0.015
```

Table 10: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
LE1	LE1a_Confidence	1.000	0.000	NA	NA	0.891
LE1	LE1b_Confidence	0.601	0.230	2.619	0.009	0.556
LE1	LE1c_Confidence	0.837	0.298	2.812	0.005	0.607
LE1	LE1d_Confidence	0.589	0.236	2.495	0.013	0.525
LE1	LE1e_Confidence	0.351	0.215	1.632	0.103	0.334

The p-value of this subdomian is 0.762, so we will keep all the questions in this subdomian.

Second Subdomain

```
## lavaan 0.6-5 ended normally after 32 iterations
##
##      Estimator                      ML
##      Optimization method          NLMINB
##      Number of free parameters      12
##
##                                Used      Total
##      Number of observations          28        84
##
## Model Test User Model:
##
##      Test statistic                  18.696
##      Degrees of freedom                9
##      P-value (Chi-square)              0.028
##
## Parameter Estimates:
##
##      Information                      Expected
##      Information saturated (h1) model  Structured
##      Standard errors                   Standard
##
## Latent Variables:
##
##      Estimate  Std.Err  z-value  P(>|z|)
##      LE2 =~
##      LE2a_Confidenc    1.000
##      LE2b_Confidenc    1.036    0.486    2.131    0.033
##      LE2c_Confidenc    1.428    0.539    2.647    0.008
##      LE2d_Confidenc    1.164    0.493    2.360    0.018
##      LE2e_Confidenc    0.855    0.407    2.101    0.036
##      LE2f_Confidenc    1.419    0.597    2.378    0.017
##
## Variances:
##
##      Estimate  Std.Err  z-value  P(>|z|)
##      .LE2a_Confidenc    1.221    0.403    3.033    0.002
##      .LE2b_Confidenc    2.313    0.691    3.345    0.001
##      .LE2c_Confidenc    1.539    0.593    2.595    0.009
```

```
##      .LE2d_Confidenc    1.975    0.627    3.151    0.002
##      .LE2e_Confidenc    1.656    0.492    3.364    0.001
##      .LE2f_Confidenc    2.840    0.907    3.131    0.002
##      LE2                0.769    0.482    1.596    0.110
```

Table 11: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
LE2	LE2a_Confidence	1.000	0.000	NA	NA	0.622
LE2	LE2b_Confidence	1.036	0.486	2.131	0.033	0.513
LE2	LE2c_Confidence	1.428	0.539	2.647	0.008	0.710
LE2	LE2d_Confidence	1.164	0.493	2.360	0.018	0.587
LE2	LE2e_Confidence	0.855	0.407	2.101	0.036	0.503
LE2	LE2f_Confidence	1.419	0.597	2.378	0.017	0.594

In the second subdomain, the p-value is $0.028 < 0.05$, so we will drop the question LE2a to see how the model will be.

```
## lavaan 0.6-5 ended normally after 30 iterations
##
##      Estimator                      ML
##      Optimization method          NLMINB
##      Number of free parameters      10
##
##                                     Used      Total
##      Number of observations         29        84
##
## Model Test User Model:
##
##      Test statistic                  1.583
##      Degrees of freedom              5
##      P-value (Chi-square)           0.903
##
## Parameter Estimates:
##
##      Information                    Expected
##      Information saturated (h1) model Structured
##      Standard errors                 Standard
##
## Latent Variables:
##      Estimate Std.Err z-value P(>|z|)
##      LE2 =~
##      LE2b_Confidenc    1.000
##      LE2c_Confidenc    1.065    0.423    2.520    0.012
##      LE2d_Confidenc    0.791    0.363    2.179    0.029
##      LE2e_Confidenc    0.640    0.305    2.098    0.036
##      LE2f_Confidenc    0.871    0.424    2.053    0.040
##
## Variances:
##      Estimate Std.Err z-value P(>|z|)
##      .LE2b_Confidenc    1.704    0.642    2.652    0.008
##      .LE2c_Confidenc    1.676    0.675    2.482    0.013
##      .LE2d_Confidenc    2.153    0.662    3.250    0.001
```

```
##      .LE2e_Confidenc    1.610    0.485    3.323    0.001
##      .LE2f_Confidenc    3.219    0.959    3.358    0.001
##      LE2                1.354    0.811    1.670    0.095
```

Table 12: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
LE2	LE2b_Confidence	1.000	0.000	NA	NA	0.665
LE2	LE2c_Confidence	1.065	0.423	2.520	0.012	0.692
LE2	LE2d_Confidence	0.791	0.363	2.179	0.029	0.532
LE2	LE2e_Confidence	0.640	0.305	2.098	0.036	0.506
LE2	LE2f_Confidence	0.871	0.424	2.053	0.040	0.492

After dropping the LE2a, we have a p value of $0.9 > 0.05$. So we will keep all the other questions.

Third Subdomain

```
## lavaan 0.6-5 ended normally after 33 iterations
##
##      Estimator                      ML
##      Optimization method          NLMINB
##      Number of free parameters      14
##
##                                Used      Total
##      Number of observations          29        84
##
## Model Test User Model:
##
##      Test statistic                  20.428
##      Degrees of freedom              14
##      P-value (Chi-square)            0.117
##
## Parameter Estimates:
##
##      Information                    Expected
##      Information saturated (h1) model Structured
##      Standard errors                Standard
##
## Latent Variables:
##
##      Estimate Std.Err z-value P(>|z|)
##      LE3 =~
##      LE3a_Confidenc    1.000
##      LE3b_Confidenc    0.977    0.377    2.593    0.010
##      LE3c_Confidenc    0.273    0.219    1.242    0.214
##      LE3d_Confidenc    0.563    0.326    1.729    0.084
##      LE3e_Confidenc    0.653    0.306    2.132    0.033
##      LE3f_Confidenc    0.679    0.308    2.205    0.027
##      LE3g_Confidenc    0.302    0.246    1.229    0.219
##
## Variances:
##
##      Estimate Std.Err z-value P(>|z|)
##      .LE3a_Confidenc    2.782    1.002    2.778    0.005
##      .LE3b_Confidenc    1.836    0.785    2.340    0.019
```

```
##      .LE3c_Confidenc    1.865    0.504    3.700    0.000
##      .LE3d_Confidenc    3.508    0.985    3.561    0.000
##      .LE3e_Confidenc    2.503    0.751    3.335    0.001
##      .LE3f_Confidenc    2.399    0.733    3.272    0.001
##      .LE3g_Confidenc    2.355    0.636    3.703    0.000
##      LE3                2.069    1.248    1.657    0.097
```

Table 13: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
LE3	LE3a_Confidence	1.000	0.000	NA	NA	0.653
LE3	LE3b_Confidence	0.977	0.377	2.593	0.010	0.720
LE3	LE3c_Confidence	0.273	0.219	1.242	0.214	0.276
LE3	LE3d_Confidence	0.563	0.326	1.729	0.084	0.397
LE3	LE3e_Confidence	0.653	0.306	2.132	0.033	0.511
LE3	LE3f_Confidence	0.679	0.308	2.205	0.027	0.534
LE3	LE3g_Confidence	0.302	0.246	1.229	0.219	0.273

In the third subdomain, we have a p value of $0.117 > 0.05$, so we will keep all the questions.

Fourth Subdomain

```
## lavaan 0.6-5 ended normally after 30 iterations
##
##      Estimator                      ML
##      Optimization method          NLMINB
##      Number of free parameters      10
##
##                                     Used      Total
##      Number of observations          29         84
##
## Model Test User Model:
##
##      Test statistic                  8.065
##      Degrees of freedom                5
##      P-value (Chi-square)             0.153
##
## Parameter Estimates:
##
##      Information                    Expected
##      Information saturated (h1) model Structured
##      Standard errors                 Standard
##
## Latent Variables:
##      Estimate Std.Err z-value P(>|z|)
##      LE4 =~
##      LE4a_Confidenc    1.000
##      LE4b_Confidenc    0.570    0.215    2.654    0.008
##      LE4c_Confidenc    0.593    0.171    3.466    0.001
##      LE4d_Confidenc    0.869    0.224    3.872    0.000
##      LE4e_Confidenc    0.522    0.229    2.285    0.022
##
## Variances:
```

```
##           Estimate Std.Err z-value P(>|z|)
## .LE4a_Confidenc  1.148   0.708   1.622   0.105
## .LE4b_Confidenc  3.175   0.893   3.555   0.000
## .LE4c_Confidenc  1.603   0.496   3.234   0.001
## .LE4d_Confidenc  2.159   0.766   2.820   0.005
## .LE4e_Confidenc  3.848   1.059   3.635   0.000
##    LE4           3.501   1.345   2.603   0.009
```

Table 14: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
LE4	LE4a_Confidence	1.000	0.000	NA	NA	0.868
LE4	LE4b_Confidence	0.570	0.215	2.654	0.008	0.513
LE4	LE4c_Confidence	0.593	0.171	3.466	0.001	0.659
LE4	LE4d_Confidence	0.869	0.224	3.872	0.000	0.742
LE4	LE4e_Confidence	0.522	0.229	2.285	0.022	0.446

In the fourth subdomain, we have a p value of 0.153. We will keep all the questions in this subdomain.

Fifth subdomain

```
## lavaan 0.6-5 ended normally after 26 iterations
##
## Estimator ML
## Optimization method NLMINB
## Number of free parameters 8
##
## Used Total
## Number of observations 29 84
##
## Model Test User Model:
##
## Test statistic 4.188
## Degrees of freedom 2
## P-value (Chi-square) 0.123
##
## Parameter Estimates:
##
## Information Expected
## Information saturated (h1) model Structured
## Standard errors Standard
##
## Latent Variables:
##           Estimate Std.Err z-value P(>|z|)
## LE5 =~
## LE5a_Confidenc  1.000
## LE5b_Confidenc  0.570   0.243   2.343   0.019
## LE5c_Confidenc  1.325   0.474   2.794   0.005
## LE5d_Confidenc  0.629   0.289   2.178   0.029
##
## Variances:
##           Estimate Std.Err z-value P(>|z|)
## .LE5a_Confidenc  1.564   0.580   2.696   0.007
```

```
##      .LE5b_Confidenc    1.328    0.381    3.487    0.000
##      .LE5c_Confidenc    0.689    0.723    0.952    0.341
##      .LE5d_Confidenc    1.985    0.558    3.557    0.000
##      LE5                1.351    0.765    1.766    0.077
```

Table 15: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
LE5	LE5a_Confidence	1.000	0.000	NA	NA	0.681
LE5	LE5b_Confidence	0.570	0.243	2.343	0.019	0.498
LE5	LE5c_Confidence	1.325	0.474	2.794	0.005	0.880
LE5	LE5d_Confidence	0.629	0.289	2.178	0.029	0.460

In the fifth subdomain, we have a p value of 0.123, so we will keep all the questions in this subdomain.

Sixth subdomain

```
## lavaan 0.6-5 ended normally after 37 iterations
##
##      Estimator                      ML
##      Optimization method          NLMINB
##      Number of free parameters      8
##
##                                     Used      Total
##      Number of observations          29         84
##
## Model Test User Model:
##
##      Test statistic                  0.832
##      Degrees of freedom                2
##      P-value (Chi-square)             0.660
##
## Parameter Estimates:
##
##      Information                      Expected
##      Information saturated (h1) model  Structured
##      Standard errors                  Standard
##
## Latent Variables:
##
##      Estimate  Std.Err  z-value  P(>|z|)
##      LE6 =~
##      LE6a_Confidenc    1.000
##      LE6b_Confidenc    2.725    2.229    1.223    0.222
##      LE6c_Confidenc    2.058    1.717    1.198    0.231
##      LE6d_Confidenc    3.160    2.585    1.222    0.222
##
## Variances:
##
##      Estimate  Std.Err  z-value  P(>|z|)
##      .LE6a_Confidenc    4.983    1.322    3.770    0.000
##      .LE6b_Confidenc    0.807    0.466    1.732    0.083
##      .LE6c_Confidenc    1.545    0.473    3.270    0.001
##      .LE6d_Confidenc    1.195    0.640    1.867    0.062
##      LE6                0.309    0.503    0.614    0.539
```


Table 16: Factor Loadings

Latent Factor	Indicator	B	SE	Z	p-value	loading
LE6	LE6a_Confidence	1.000	0.000	NA	NA	0.241
LE6	LE6b_Confidence	2.725	2.229	1.223	0.222	0.860
LE6	LE6c_Confidence	2.058	1.717	1.198	0.231	0.677
LE6	LE6d_Confidence	3.160	2.585	1.222	0.222	0.849

In the sixth subdomain, the p-value is $0.66 > 0.05$. We will not drop any question in this subdomain.

Conclusion / Discussion

For PER&FEEDBACK table, I dropped “PF1e_Confidence”, “PF2c_Confidence” and “PF5c_Confidence” so that P-value of all subdomains are greater than 0.05 finally.

For the Learning Experience table, we will only drop “LE2a_Confidence” and keep all the remaining questions in order to let the P-value of all subdomains are greater than 0.05 finally.

Appendix