

# Questionnaire relatif à l'utilisation de t-prof

## Introduction

Nous avons mené un test utilisateur sur trois groupes d'étudiants-stagiaires en ESPE comportant un total de 62 répondants sur les sites de Gennevilliers et de Antony les 6 et 14 juin 2019. Le groupe 1 (22 personnes) était issu d'une formation en alternance organisée en 3 jours de cours en ESPE et 3 jours d'enseignement devant la classe (la classe étant partagée avec un professeur des écoles titulaire). Les groupes 2 et 3 (respectivement 28 et 12 personnes) étaient issus d'une formation en alternance organisée en 3 semaines de cours en ESPE et 3 semaines de d'enseignement devant la classe (la classe étant partagée avec un autre étudiant-stagiaire). Lors des testes utilisateur nous avons cependant noté les différences suivantes entre les trois groupes : Le groupe n°1 (GU) est arrivé motivé aux tests et l'implication des participants dans la séance est resté constante jusqu'à la fin des 3h. Le groupe n°2 (GD) a commencé la séance moins motivé et certains participants ont souhaité partir plus tôt, tout en remplissant le formulaire. Le groupe n°3 (GT), le plus petit des trois est arrivé lui aussi motivé aux tests et l'implication des étudiants-stagiaires est allé grandissante au fur et à mesure de la séance. A la fin, plus de 6 étudiant-stagiaires (sur 12) se sont déclarés intéressés pour faire un second test utilisateur dans les mois suivant l'expérimentation.

Les trois groupes de testeurs constituaient 62 répondants à 82% féminin et 17% masculin. L'âge moyen était de 31 ans, l'âge minimum de 22 ans et l'âge maximum de 56 ans. 60% des répondants étaient en Master 2 (c'est-à-dire issus d'un Master 1 MEEF) et 40% étaient issus d'un parcours adapté.

## Descriptif du test utilisateur

Chaque test utilisateur a été effectué selon le mode opératoire suivant : Pour commencer, les étudiants-stagiaires découvraient individuellement l'application, sur tablette (iPad) ou ordinateur de bureau. Ils avaient pour consigne de se familiariser avec l'application et de noter les points marquants. Après une mise en commun des premières impressions (mise en commun filmé) sous la modération des formatrices, les étudiants-stagiaires devaient parcourir la situation : « Cycle 2 - CE2 Dictée réflexive ». Ils avaient pour objectif d'arriver jusqu'à la fin de la situation et de noter les points du bilan qui étaient soit en accord, soit en désaccord de leur propre appréciation du parcours. Ces points faisaient alors l'objet d'une mise en commun (filmée) et modérée par les formatrices. Par la suite, les étudiants-stagiaires devaient à nouveau parcourir la situation « Cycle 2 - CE2 Dictée réflexive » soit en améliorant leur parcours, soit en testant d'autres possibilités. Pour finir, les formatrices effectuaient un bilan sur l'utilisation de l'application et les étudiants stagiaires remplissaient le questionnaire en ligne (voir annexe). L'ensemble de l'expérimentation durait 3 heures avec une pause de 15 minutes.

## Le questionnaire

### Acceptation Instrumentale appliquée au TICE et Intérêt situationnel pour les serious games

Le questionnaire, administré sur **Google Form**, a été mis à disposition des testeurs après la fin de chaque session d'expérimentation. Le questionnaire est basé majoritairement sur l'échelle d'Acceptation Instrumentale appliquée au TICE (TAM\_INJ) (Caron, Heutte, 2017) pour les dimensions de l'injonction, de l'utilisabilité et de l'utilité. Pour la mesure de l'intérêt situationnel, le questionnaire se base sur l'échelle de l'intérêt situationnel et individuel pour les Serious games (IS2G, Chainon, Fenouillet & Heutte, 2014).

#### Injonction

Ici, l'injonction est définie comme étant l'ordre ou l'impératif donné par l'institution aux étudiants-stagiaires pour tester l'application. L'injonction est calculée sur la base des questions suivantes :

- Je teste t-prof uniquement parce que l'on me le demande
- Je teste t-prof parce que mes collègues le testent aussi
- Je teste t-prof parce que j'espère de son usage une certaine reconnaissance au sein de mon ESPE.

#### Utilisabilité

On définit l'utilisabilité comme le degré selon lequel un produit peut être utilisé, par des utilisateurs identifiés, pour atteindre des buts définis avec efficacité, efficience et satisfaction, dans un contexte d'utilisation spécifié (norme ISO 9241-11). Plus l'utilisabilité est élevée, plus on considèrera que l'application est aboutie et agréable à utiliser par les utilisateur. L'utilisabilité de l'application est obtenue par la moyenne des trois questions suivantes :

- Je trouve t-prof facile à utiliser ;
- C'est facile d'apprendre à utiliser t-prof ;
- Lorsque j'utilise t-prof, j'arrive facilement à atteindre l'objectif que je me suis fixé ou qu'on m'a donné.

#### Utilité pour les stagiaire-étudiants

Premier éléments essentiel de la mesure : l'utilité pour les étudiant-stagiaire ; Nous souhaitons mesurer quelle est la perception de l'utilité de l'application dans un cadre de cours en ESPE. L'utilité pour les étudiants-stagiaires est mesurée par les questions suivantes :

- Je pense qu'utiliser T-Prof peut m'aider à mieux enseigner ;
- Je pense que T-Prof peut être utile pour enseigner.

## Intérêt situationnel pour la pédagogie

Nous voulons dans un premier temps nous assurer que cette application n'est pas uniquement accessible à un public de professionnels expérimentés. Pour ce faire nous mesurons l'intérêt situationnel pour la pédagogie et nous le calculons à partir des questions suivantes :

- J'ai toujours voulu être enseignant ;
- En matière de pédagogie, je me sens à l'aise ;
- Le métier auquel je me destine me convient.

## Dimensions additionnelles

T-prof étant décrit comme un simulateur comportemental, c'est la dimension du transfert qui nous intéresse également. Comme l'indiquait Rogalski et Colon (2018), la question du transfert a toujours été un point central de l'évaluation des simulateurs et ce, depuis les origines de ces derniers. Pour un serious game, le transfert est défini comme le degré avec lequel les apprenants appliquent les compétences acquises depuis le jeu vers une situation professionnelle (Holton, Bates, Seyler, Garvalho, 1997, cité par Michel, Kreziak et Héraud, 2009, p.6). La simulation comportementale comporte un autre point central : l'analyse réflexive (HAS, 2012). C'est par la confrontation à l'expérience vécue au sein du simulateur que l'apprenant peut enrichir sa pratique professionnelle.

- En accord avec les compétences que nous souhaitons mettre en place au sein de l'application t-prof, nous avons développé trois dimensions supplémentaires relatives au transfert et à l'analyse réflexive induits par l'application. Pour rappel, les compétences visées sont les suivantes :
- Connaître les élèves et les processus d'apprentissage ;
- Prendre en compte la diversité des élèves ;
- Accompagner les élèves dans leur parcours de formation ;
- Agir en éducateur responsable et selon des principes éthiques ;
- S'engager dans une démarche individuelle et collective de développement professionnel.

## Transfert déclaré

Nous proposons donc la première dimension additionnelle, intitulée « Transfert déclaré » et basée sur les questions suivantes :

Je pense qu'un usage approfondi de cette application en début de formation permet de :

- Anticiper le déroulement de la classe réelle ;
- Formuler des consignes et gérer l'enrôlement des élèves en début de séance ;
- Percevoir l'enchaînement des interactions qui conduisent un enseignant à se retrouver dans une impasse ;

- Anticiper la gestion du groupe ou de certains élèves ;
- Réfléchir aux conséquences des choix de l'enseignant sur l'implication des élèves.

## Analyse réflexive projetée

Nos interrogeons ensuite l'analyse réflexive que les étudiant-stagiaires ont effectué au cours de la séance et développons la dimension additionnelle « analyse réflexive projetée », mesurée par les questions suivantes :

- Les points soulevés dans le bilan me donnent de nouvelles idées pour gérer un groupe d'élèves ;
- Je trouve que T-Prof est une application pertinente pour réfléchir aux gestes professionnels à adopter face à un groupe d'élèves ;
- J'ai apprécié de pouvoir expérimenter des choses que je n'ai pas encore faites en classe ;
- Certains choix m'ont amené à réfléchir à la posture à adopter face aux élèves.

## Recommandation

Enfin, nous faisons l'hypothèse que si les étudiants ont trouvé l'application utile pour leur formation, ils recommanderont cette application à leurs proches ou collègues. Nous développons alors la dimension additionnelle « Recommandation ». Nous mesurons cette dimension par les questions suivantes :

- J'aurais aimé avoir cet outil pendant ma formation ;
- Je pense recommander T-Prof à un ou des amis ;
- T-Prof est un bon outil de formation professionnelle.

Nous procédons à l'analyse statistique des différentes dimensions décrites.

Les différentes dimensions questionnées sont les suivantes:

```
[1] "is" "inj" "utb" "utf" "arp" "td" "rx" "Age" "SEXE" "GRP"
[11] "PAR"
```

IS : Intérêt situationnel pour la pédagogie ; INJ : Injonction à l'utilisation de l'application par l'institution ; UTB : Utilisabilité de l'application ; UTF : Utilité pour les étudiants-stagiaire ; ARP : Analyse réflexive projetées des étudiants sur l'application ; TD : Transfert déclaré ; RX : Recommandation du dispositif ;

## Statistiques descriptives élémentaires

### Éléments centraux et quantiles

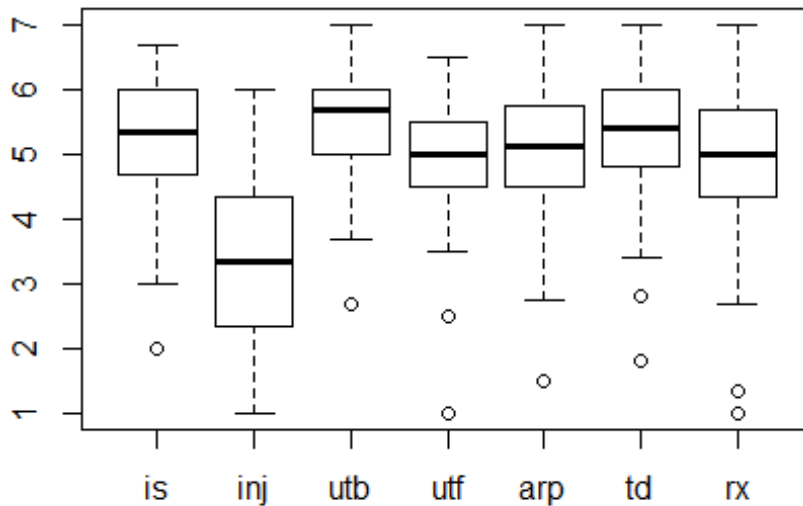
Une première analyse descriptive s'appuie sur les indicateurs de position et de dispersion. Les réponses sont exprimées sous forme de Likert à 7 points :

- 1: Pas du tout d'accord
- 2: très peu d'accord
- 3: un peu d'accord
- 4: moyennement d'accord
- 5: assez d'accord
- 6: fortement d'accord
- 7: tout à fait d'accord

Pour traiter nos données nous utilisons dans la suite de ce rapport le logiciel **R**

is		inj		utb		utf	
Min.	:2.000	Min.	:1.000	Min.	:2.670	Min.	:1.000
1st Qu.	:4.670	1st Qu.	:2.330	1st Qu.	:5.000	1st Qu.	:4.500
Median	:5.330	Median	:3.330	Median	:5.670	Median	:5.000
Mean	:5.167	Mean	:3.392	Mean	:5.473	Mean	:4.774
3rd Qu.	:5.918	3rd Qu.	:4.247	3rd Qu.	:6.000	3rd Qu.	:5.500
Max.	:6.670	Max.	:6.000	Max.	:7.000	Max.	:6.500
arp		td		rx		Age	
Min.	:1.500	Min.	:1.800	Min.	:1.000	Min.	:22.00
1st Qu.	:4.500	1st Qu.	:4.800	1st Qu.	:4.330	1st Qu.	:24.00
Median	:5.125	Median	:5.400	Median	:5.000	Median	:30.00
Mean	:5.016	Mean	:5.271	Mean	:4.865	Mean	:31.79
3rd Qu.	:5.750	3rd Qu.	:6.000	3rd Qu.	:5.670	3rd Qu.	:38.00
Max.	:7.000	Max.	:7.000	Max.	:7.000	Max.	:56.00
SEXE	GRP	PAR					
Femme:51	GD:28	M2:37					
Homme:11	GT:12	PA:25					
	GU:22						

## Affichage des boîtes à moustache

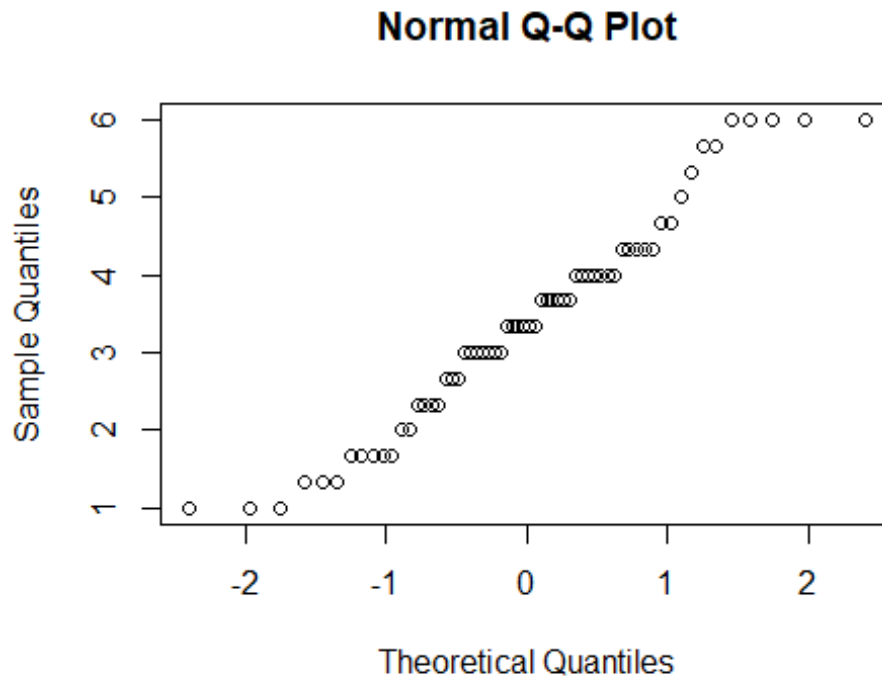


Les diagrammes en boîtes à moustaches montrent que l'injonction est la seule dimension dont la majorité de la variance est en dessous de 4. On remarque un intérêt situationnel relativement élevé et surtout un bon niveau d'utilisabilité (premier quartile à 5, troisième quartile à 6). De même, l'utilité de l'application pour les étudiant-stagiaires varie majoritairement entre 4.5 et 5.5. On peut donc dans un premier temps considérer l'utilité comme validée. Nous cherchons à présent à trouver quels sont les dimensions qui seraient à même de renforcer l'ancrage de la mesure de l'utilité pour les étudiant-stagiaires.

## Normalité de la distribution

Dans un premier temps nous testons la normalité de l'échantillon avec la méthode Q-Q plot ou «quantile-quantile plot».

## 1/ Méthode graphique



Les points n'étant pas alignés, nous vérifions par un calcul statistique que l'échantillon ne suit pas une loi normale.

## 2/ Méthode par test

Nous nous proposons d'utiliser le test de Shapiro

```
Shapiro-Wilk normality test
```

```
data: questionnaire_moy$inj  
W = 0.96182, valeur p = 0.05118
```

Hypothèse nulle : l'échantillon suit une loi normale. Pas conséquent si la valeur p est significative, l'échantillon ne suit pas une loi normale.

L'échantillon ne suit pas une loi normale car la valeur p n'est pas significative (valeur p = 0.05).

## Recherche de corrélation

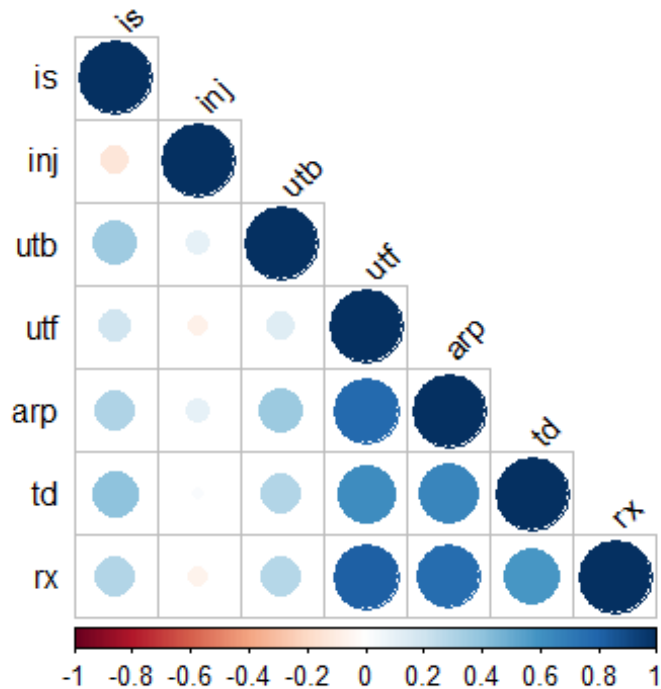
Nous cherchons à présent quelles sont les corrélations entre les différentes dimensions.

## Affichage de la matrice de corrélation pour les différentes variables

```
      is      inj      utb      utf      arp      td
is  1.0000000 -0.13707542 0.3519546  0.19500600 0.3013357 0.40348757
inj -0.1370754  1.00000000 0.1025807 -0.07513464 0.1050285 0.02273201
utb 0.3519546  0.10258067 1.0000000  0.13768091 0.3623164 0.29975658
utf 0.1950060  -0.07513464 0.1376809  1.00000000 0.7772882 0.62835728
arp 0.3013357  0.10502854 0.3623164  0.77728822 1.0000000 0.65731364
td  0.4034876  0.02273201 0.2997566  0.62835728 0.6573136 1.00000000
rx  0.2942584 -0.06833366 0.2891308  0.81475840 0.7677162 0.58935088
      rx
is  0.29425837
inj -0.06833366
utb 0.28913085
utf 0.81475840
arp 0.76771622
td 0.58935088
rx 1.00000000
```

Warning: package 'corrplot' was built under R version 3.5.3

corrplot 0.84 loaded



Nous voyons que sur le correlogramme, les dimensions utilisabilité (utf), analyse réflexive projeté (arp), transfert déclaré (td) et recommandation (rx) sont corrélées avec des coefficients de corrélation variant entre 0.5 et 0.8. Pour plus de précision, nous proposons de calculer la matrice de corrélation avec ses taux de significativité.



## Matrice et taux de significativité

	is	inj	utb	utf	arp	td
is						
inj	-0.17					
utb	0.22	0.11				
utf	0.20	-0.07	0.09			
arp	0.28*	0.14	0.38**	0.70****		
td	0.36**	-0.02	0.20	0.46**	0.52****	
rx	0.29*	-0.04	0.24	0.78****	0.67****	0.44**

Cette matrice nous indique que nous avons les corrélations fortes suivantes :

- Utilité pour les étudiant-stagiaires <-> analyse réflexive projetée
- Utilité pour les étudiant-stagiaires <-> recommandation

Et la corrélation moyenne suivante :

- Utilité pour les étudiant-stagiaires <-> transfert déclaré

Pour tous ces résultats, on peut écarter le hasard (l'hypothèse H0) au risque de 1/10 000. valeur ps(0.0001, 0.001, 0.01, 0.05, 0.1, 1) <=> symboles(“\*\*\*\*”, “\*\*\*”, “\*\*”, “\*”, “.”, “ ”)

Nous validons ainsi plusieurs de nos hypothèses de départ. Premièrement, plus un dispositif est jugé utile, plus il sera recommandé. Dans notre cas, la dimension recommandation étant fortement corrélée à l'utilité, nous en déduisons une utilité élevée de l'application t-prof.

De même, un dispositif permettant une analyse réflexive, même projetée, renforce son utilité propre. Dans notre cas, l'analyse réflexive projetée est aussi fortement corrélée avec l'utilité du dispositif.

Enfin, le transfert déclaré est moyennement corrélé à l'analyse réflexive projetée (très significatif) et à l'utilité pour les étudiant-stagiaires (significatif).

Nous avons donc une application dont l'utilité mesurée est élevée. De plus, cette dimension est, d'une part renforcé par l'analyse réflexive sur la pratique professionnelle des étudiant-stagiaires et d'autre part, validé par un haut niveau de recommandation, lui aussi corrélée à l'utilité. D'autre part l'application permet un transfert, renforçant ainsi son utilité et l'analyse réflexive qu'il en découle.

Par ailleurs, nous constatons que l'intérêt situationnel est légèrement corrélé au transfert déclaré. L'intérêt situationnel vient donc renforcer cette dimension sans pour autant être un verrou d'accessibilité au dispositif.

## Remarques annexes

Nous calculons la corrélation entre l'âge des répondants et la dimension injonction selon la méthode de Spearman.

Spearman's rank correlation rho

data: questionnaire\_moy\_2\$inj and questionnaire\_moy\_2\$Age  
S = 54690, valeur p = 0.00251

```
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
-0.377208
```

Nous obtenons une valeur de l'ordre de  $-0.37$ , c'est à dire une corrélation faible mais significative. Nous en concluons que la sensibilité à l'injonction est inversement proportionnel à l'âge des participants. Nous utiliserons ce résultat lorsque que nous analyserons la différence entre les groupes et les parcours.

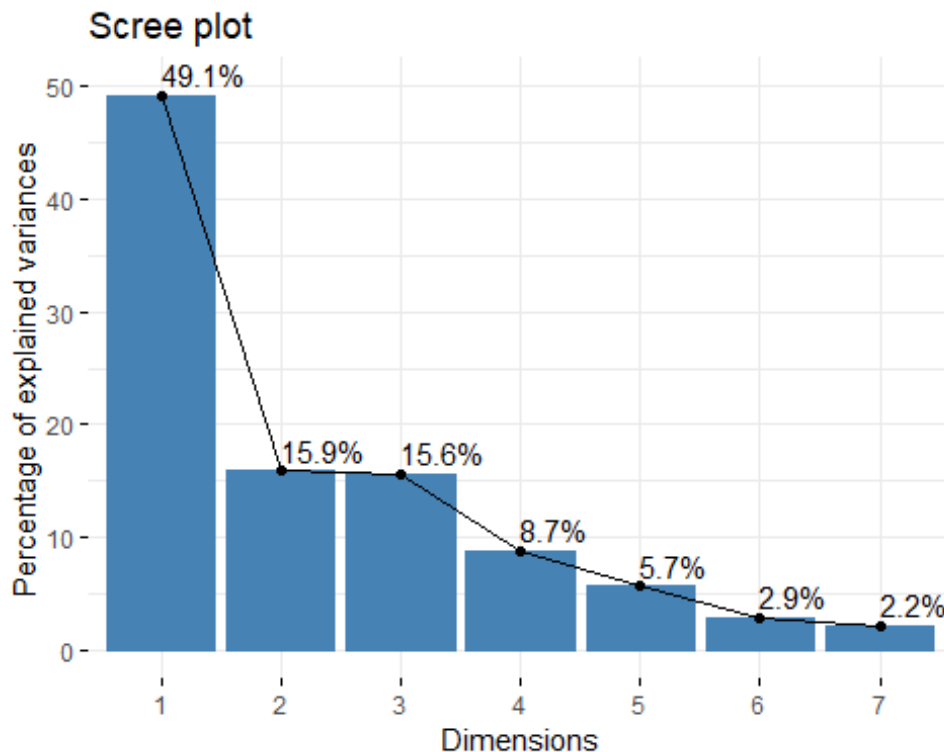
## Réalisation d'une analyse en composante principale

L'analyse en composante principale (ACP) a pour but d'identifier les dimensions qui rendent le mieux compte de la variabilité des résultats. Cette analyse structure les perceptions des utilisateurs selon deux axes.

### Inertie et contribution

#### Inertie des axes factoriels

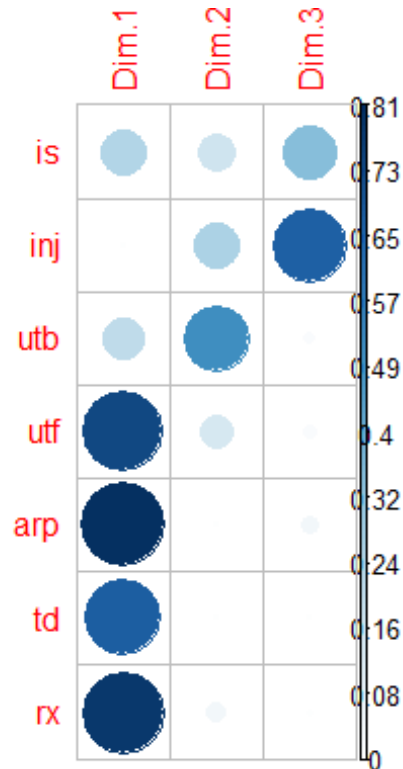
Dans un premier temps nous calculons l'inertie des axes factoriels. L'inertie des axes factoriels indique d'une part si les variables sont structurées et suggère d'autre part le nombre judicieux de composantes principales à étudier



Nous considérerons dans le présent rapport les dimensions 1, 2 et 3. Ensemble, elles représentent 80% de la variance.

### Contribution de chaque dimension

Dans un second temps, nous calculons la contribution de chaque dimension aux axes proposés par l'acp

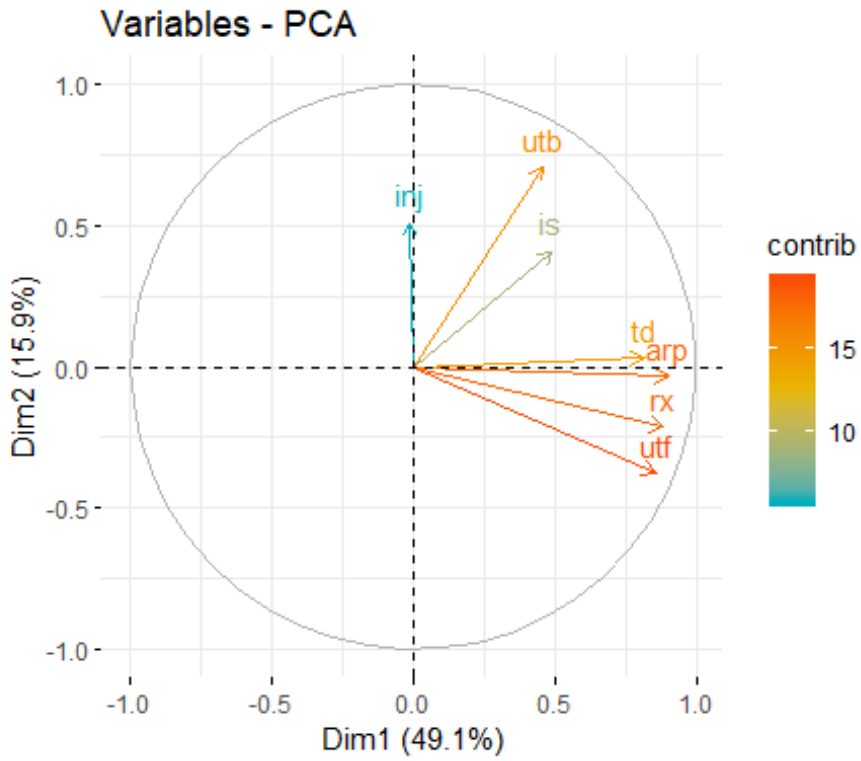


Nous observons donc une première dimension, constituée de l'utilité, de l'analyse réflexive projetée, du transfert déclaré et de la recommandation. Cette dimension représente 49% de la variance. La dimension 2 est principalement constituée de l'utilisabilité et plus faiblement de l'injonction. Enfin, la dimension 3 est constituée de l'injonction et de l'intérêt situationnel, dans une plus faible mesure.

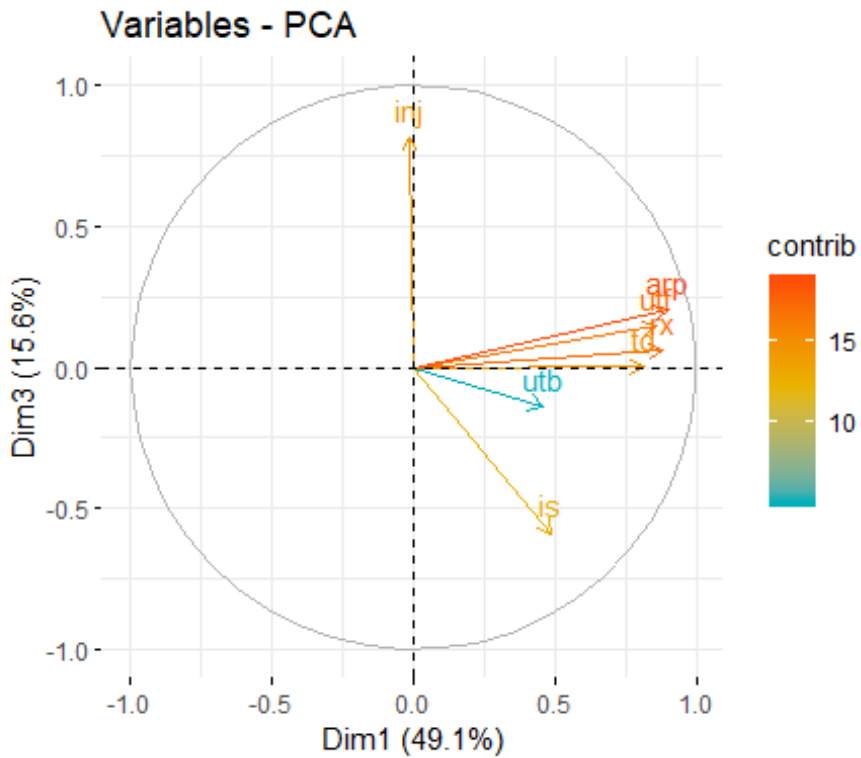
### Réalisation de l'Analyse en composante principale

#### Affichage de l'analyse

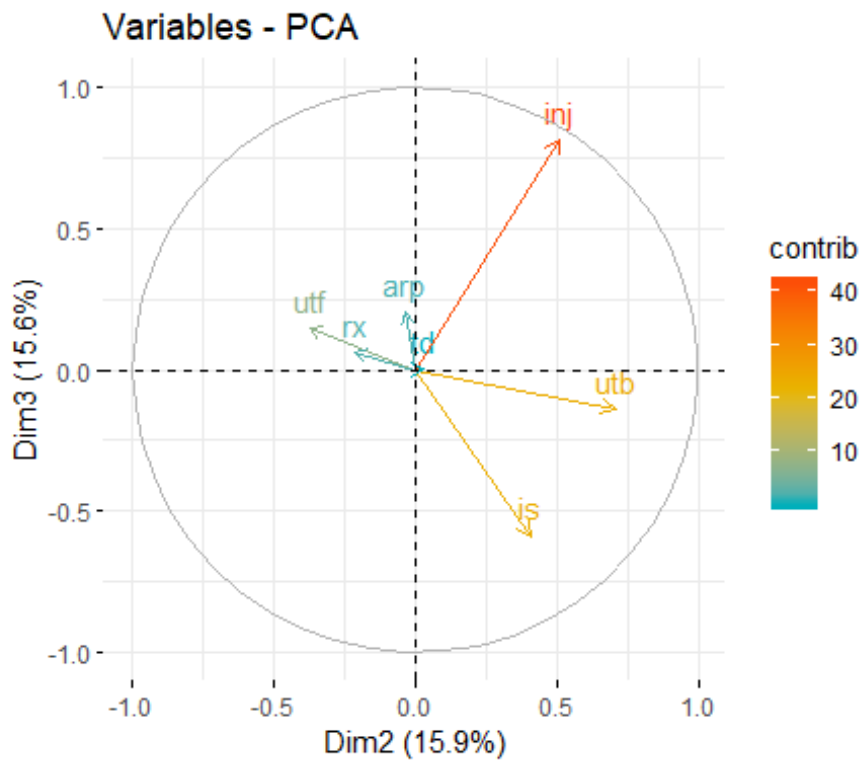
Nous affichons le graphique en deux dimensions de l'analyse en composante principale (acp), selon les axes des dimensions 1 et 2 :



Puis nous affichons le graphique de l'acp selon les axes 1 et 3.



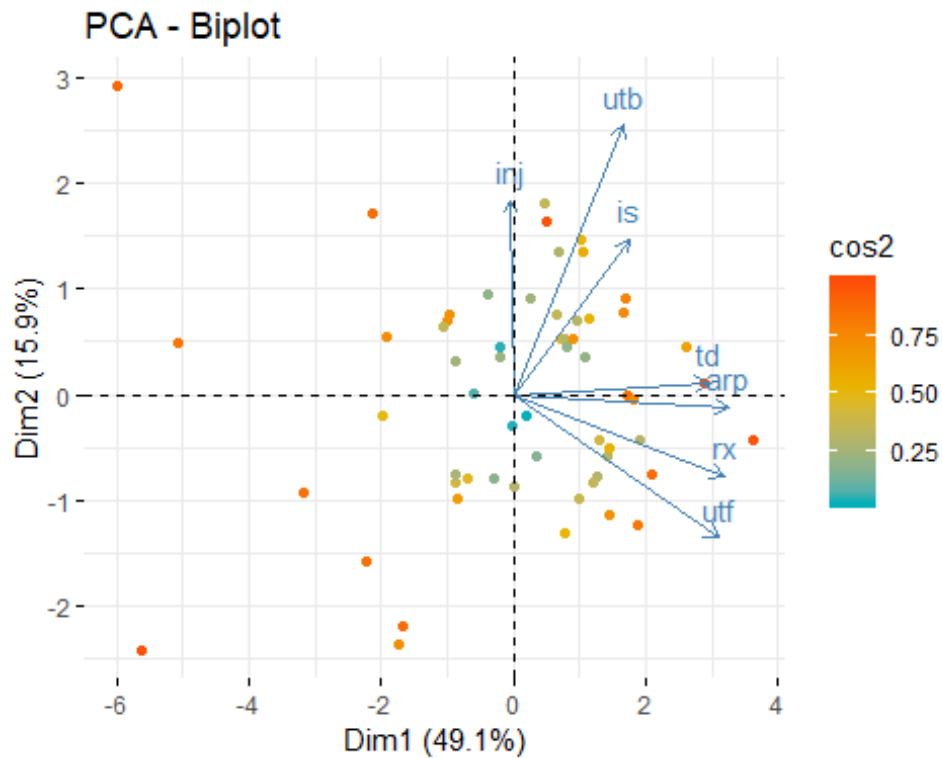
Et enfin selon les axes 2 et 3 :



Nous observons donc bien un bouquet de vecteurs constitués des dimensions utf, arp, rx et td. Normal à ces vecteurs, nous observons le vecteur de l'injonction. Enfin, le vecteur utilisabilité se positionne normalement au bouquet de vecteur et à l'injonction.

## Graphe des individus

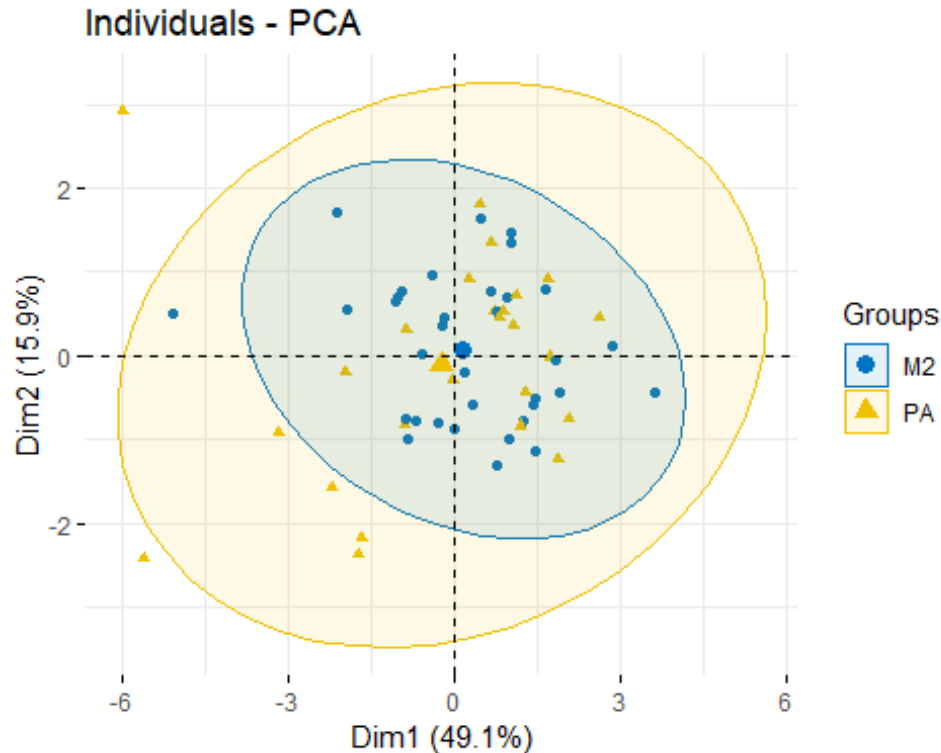
Nous affichons maintenant le graphe des individus avec discrimination en  $\text{COS}^2$ .



L'ensemble des individus se répartissent sur les axes horizontal et vertical. Il y a sur l'axe vertical une opposition entre, en bas, les individus peu sensibles à l'injonction à ceux, en haut, sensible à l'injonction. Sur l'axe horizontal une opposition entre, à droite, les individus déclarant une utilité et une analyse réflexive et, à gauche, les individus ne percevant pas ou peu d'utilité au dispositif.

## Coloration par parcours

Dans une première approche descriptive nous proposons de voir si une distinction par le parcours est pertinente pour décrire la variabilité des individus.



En première analyse, il semble que les parcours issus d'un master 1 MEEF soient plus sensible à l'injonction que les parcours adaptés. Comme les données ne sont pas normales, nous procédons à un calcul d'anova avec le test de Wilcoxon.

```
Warning in wilcox.test.default(x = c(4, 3.33, 3.67, 3, 5.67, 1.67, 4.67, :  
cannot compute exact valeur p with ties
```

Wilcoxon rank sum test with continuity correction

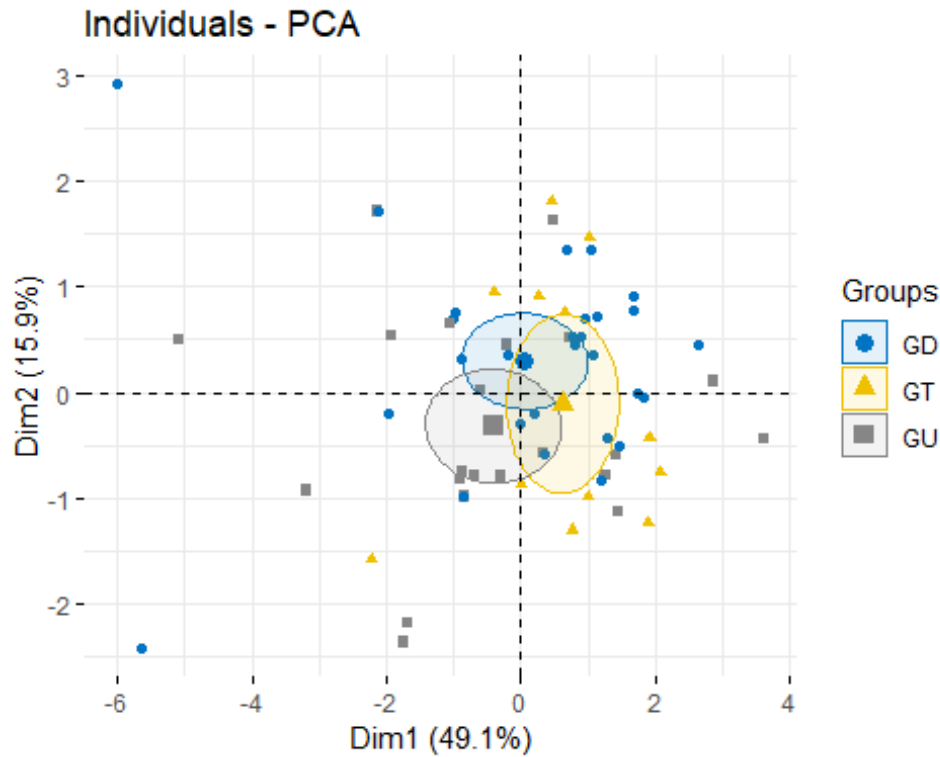
```
data: questionnaire_moy_2$inj by questionnaire_moy_2$PAR  
W = 634.5, valeur p = 0.01354  
alternative hypothesis: true location shift is not equal to 0
```

Le test se révèle positif, la valeur  $p$  étant de l'ordre de 0.01. Pour rappel, la valeur  $p$  est utilisée pour quantifier la significativité statistique d'un résultat dans le cadre d'une hypothèse nulle. Un résultat statistiquement significatif est un résultat qui serait improbable si l'hypothèse nulle (qui représente en général la norme) était vérifiée. Ici la valeur  $p$  est de 0.01 on peut dire que l'on a 1 chance sur 100 que le résultat obtenu soit du au hasard, on peut donc affirmer que selon le parcours l'injonction est différentes.

Cependant, comme nous l'avons calculé précédemment, l'injonction est inversement corrélée à l'âge des participants. Ils se trouvent que la moyenne d'âge des Master 2 et celles des parcours adaptés

sont respectivement de 28 ans et de 37 ans. La différence de sensibilité à l'injonction s'explique donc par la différence d'âge moyen entre les deux parcours.

### Coloration par groupes



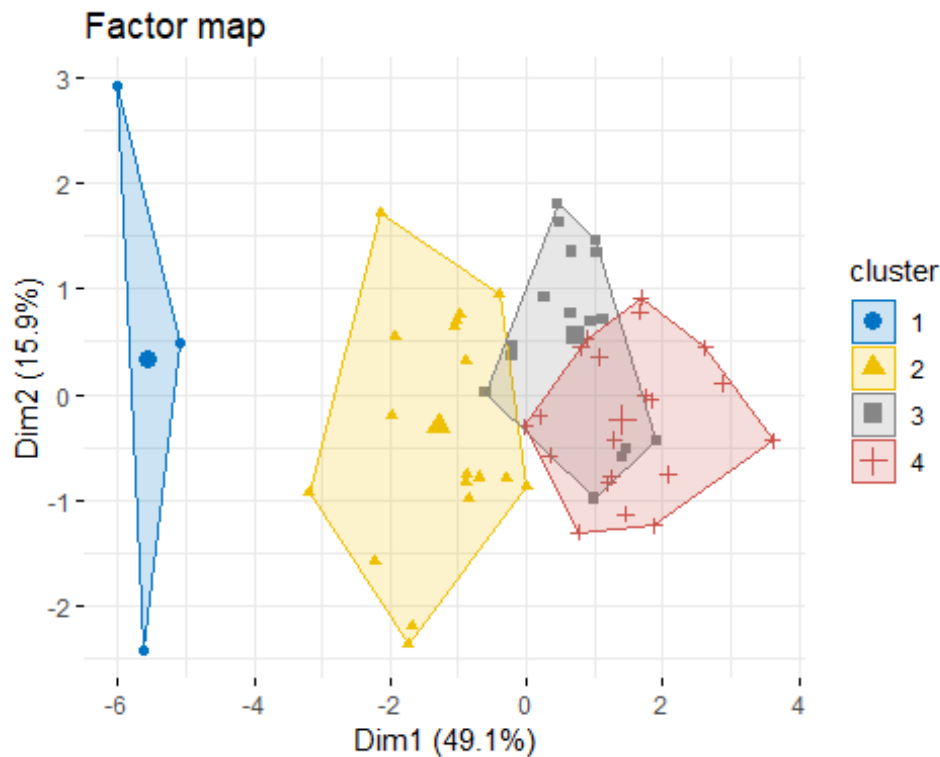
Nous observons ici les mêmes éléments que ceux décrit en introduction. Le groupe n°2, moins motivé, a été plus sensible à l'injonction. Le groupe n°3 plus réduit, c'est montré plus réceptif aux potentiels que l'application leur offrait. Enfin le groupe n°1, moins sensible à l'injonction est resté malgré tout moins réceptif à l'utilité du dispositif.

### Répartition des individus

Au regard des différenciations que nous venons de réaliser, nous en déduisons pour l'instant une répartition relativement homogènes des individus.



## Réalisation d'une cluterisation hierarchique



La classification hiérarchique que nous réalisons met ainsi en évidence 4 profils dont d'ordonnement suit globalement la dimension horizontale de l'utilité pour les étudiant-stagiaires. Le cluster 1 est délicat à décrire car il n'est composé que de 3 individus. Cependant nous notons des niveaux d'injonction extrême au sein de ce groupe et surtout, une absence marquée de la perception de l'utilité du dispositif. Le cluster 2, constitué de 21 individus, mu par l'intérêt situationnel pour la pédagogie, a perçu l'utilisabilité du dispositif ainsi qu'une part de l'utilité et des apports. Le cluster 3, constitué de 18 individus, sensible à l'injonction mais également au contenu de la séance de test, a perçu positivement les apports de l'application. Le cluster 4, constitué de 22 individus, s'étant détaché de l'injonction a pleinement perçu l'utilité pour les étudiant-stagiaires ainsi que les apports que permet l'application.

## Conclusion

Au regard de l'expérimentation menée nous observons que l'accessibilité à l'application n'est pas restreinte à des experts en pédagogie. L'intérêt situationnel ne se comporte pas comme un verrou mais peut éventuellement renforcer la perception de l'utilisabilité. En nous replaçant dans le contexte du modèle de l'acceptation instrumentale appliquée au TICE (TAM\_INJ) (Caron et Heutte, 2017) nous observons que grâce à un niveau d'injonction relativement bas, l'utilisabilité de l'application est bien perçue par les participants puisque au moins 3/4 des participants ont trouvé l'application « assez » utilisable et la moitié « fortement » utilisable ou plus. De même, l'utilité de l'application pour les étudiant-stagiaires est jugée par moitié des participants comme « assez » utile et fortement renforcée par ailleurs par les dimensions additionnelles que sont le transfert déclaré, l'analyse réflexive projetée et la recommandation.

En effectuant une analyse en composantes principales, nous observons quatre groupes ou clusters. Si nous excluons les membres du cluster 1 nous pouvons en déduire que 59 des 62 participants ont perçu une bonne utilisabilité à l'application ainsi qu'une utilité et des apports. En moyenne, en considérant les individus appartenant aux clusters 3 et 4, nous observons que 40 des 62 participants ont été à même de faire une analyse réflexive de leur gestes professionnels au sein de l'application. De plus ces 40 individus ont trouvé l'application utile (moyenne de 5,4/7, c'est à dire supérieur à « assez d'accord »).

Dans l'ensemble ce sont 40 individus sur 62 qui ont trouvé une bonne utilité à l'application, qui ont pu réaliser une analyse réflexive, qui ont déclaré percevoir un transfert et qui recommanderont l'application. Le transfert déclaré et l'analyse réflexive projetée, par leur corrélation avec l'utilité montrent les compétences visées sont bien abordées dans l'application.

Lors des prochaines expérimentations nous nous pencherons plus précisément sur l'analyse réflexive et nous mesurerons en quoi l'analyse de la pratique au sein de l'application permet d'acquérir les compétences visées.