



UNIVERSITÉ DE FRIBOURG
UNIVERSITÄT FREIBURG

Université de Fribourg Suisse
Service de didactique universitaire et compétences numériques

Travail de fin d'étude en vue de l'obtention du
Diplôme en Enseignement Supérieur et Technologie de l'Education (DAS)

Je dois utiliser quel test ?...

Description et analyse
d'un quizz auto-évaluatif en ligne
de compétences statistiques

Sous la direction du Prof. Philippe Genoud

Alain Chavaillaz
Département de Psychologie
Université de Fribourg

Septembre 2024

Déclaration sur l'honneur

Je déclare sur mon honneur que mon travail de fin d'étude est une œuvre personnelle, composée sans concours extérieur non autorisé.

Table des matières

1. Introduction.....	1
2. Description du projet.....	2
2.1. Contexte.....	2
2.2. Objectifs du projet.....	7
3. Mise en œuvre du projet.....	10
3.1. Acteurs et actrices.....	10
3.2. Comment résoudre le problème ?.....	10
3.3. Conception des items.....	12
3.4. Construction du cours Moodle et des quizz.....	14
3.5. Analyses des données des quizz.....	16
4. Evaluation du projet par les étudiant·e·s.....	19
4.1. Description.....	19
4.2. Résultats.....	21
4.3. Analyses des commentaires et suggestions.....	22
5. Discussion.....	24
5.1. Evaluation de l'outil.....	24
5.2. Expérience d'apprenant.....	25
5.3. Concepts et outils pertinents utilisés dans ce travail.....	26
5.4. Mes apprentissages durant ce travail.....	27
6. Conclusion et perspectives futures.....	29
References.....	31
Annexes.....	34
Annexe A.....	34
Annexe B.....	42

Table des figures et tableaux

<i>Figure 1</i> Modèle des Attitudes des Etudiant-e-s envers la Statistique.	4
<i>Figure 2</i> Illustration de la Taxonomie SOLO des Connaissances selon Biggs and Tang (2011).	9
<i>Figure 3</i> Arbre Décisionnel pour le Choix du Test Approprié.	12
<i>Figure 4</i> Première Section du Cours Moodle.....	14
<i>Figure 5</i> Illustration d'une Question Moodle.....	15
<i>Figure 6</i> Illustration du Feedback à une Question Moodle.....	16
<i>Figure 7</i> Section sur l'Evaluation des Quizz.	19
<i>Figure 8</i> Boîtes à Moustache pour les Items B1, B2 et B3 du Questionnaire d'Evaluation. ..	21
<i>Figure 9</i> Boîtes à Moustache pour les Items C1 à C6 du Questionnaire d'Evaluation	22
<i>Tableau 1</i> Source des Situations Expérimentales et Analyse Correspondante.....	13
<i>Tableau 2</i> Effectifs en Fonction du Type de Quizz et de l'Interaction avec la Page Moodle..	17
<i>Tableau 3</i> Questions (et Format de Réponse) du Formulaire d'Evaluation du Projet en Fonction des Différentes Parties.....	20

« Celui qui ose enseigner ne doit jamais cesser d'apprendre »

Richard Henry Dann

Lorsque j'ai commencé ma recherche de littérature, je me doutais bien de ne pas être le seul à m'intéresser et à me questionner sur l'enseignement de la statistique. Mais quelle ne fût pas ma surprise de découvrir qu'il n'existait non pas un, ni deux, ni trois, mais quatre journaux dédiés uniquement à ce sujet (« Teaching Statistics », « Journal of Statistics Education », « Statistical Education Research Journal » et « Statistique et Enseignement »).

1. Introduction

Lors de mes études en psychologie, j'ai toujours eu un fort intérêt pour la méthodologie de la recherche et, contrairement à nombre de mes camarades, plus particulièrement la statistique. Peut-être, est-ce le fait que je considère toujours cette matière comme une sorte de jeu d'enquête, dans lequel il faut trouver la meilleure solution à un problème en fonction des indices à disposition. Cet intérêt m'a valu d'être engagé comme tuteur de statistique dans les différentes unités d'enseignement proposées à l'époque par le département de Psychologie. J'ai beaucoup apprécié cette période qui m'a permis, je l'espère, d'aider mes camarades à mieux comprendre cette matière et à se préparer pour leurs évaluations. Le grand avantage du tutorat par rapport à un cours plus ou moins frontal est d'avoir le temps d'évaluer brièvement le niveau de connaissances individuel des étudiant·e·s par rapport à une thématique et de les aider à mieux surmonter leurs difficultés à résoudre les exercices proposés.

J'ai commencé ma formation en didactique universitaire lorsque j'ai commencé ma période de docteur-assistant de recherche car je voulais me préparer au mieux pour mes futurs enseignements que je souhaitais voir porter en partie sur les méthodes.

Par ma propre expérience en tant qu'étudiant et les nombreuses heures passées à aider mes camarades durant mes études et ma thèse, il me semblait nécessaire d'apprendre les bases de l'enseignement. Un doctorat financé par le Fond National Suisse pour la Recherche forme exclusivement à la recherche. Le superviseur ou la superviseuse de thèse peut (mais ne doit pas) proposer de participer (plus ou moins activement) à des activités d'enseignement, mais apprendre à enseigner n'est pas prévu. Ce projet a donc été conduit avec l'idée de proposer un outil aux étudiant·e·s de psychologie qui les aiderait à mieux comprendre certains ressorts de ce « jeu d'enquête ». De plus, il a été réalisé avant que je n'aie la chance d'être responsable de l'enseignement de la psychométrie, puis la statistique pour les premières années (en allemand).

“Tout comme certaines sciences occultes, les statistiques possèdent leur propre jargon, volontairement mis au point pour dérouter les non-initiés.”

G. O. Ashley

2. Description du projet

L'idée de ce projet est venue d'une constatation faite chaque année par les enseignant·e·s en statistique et les superviseurs et superviseuses de travaux de Bachelor du département de Psychologie de l'Université de Fribourg. Les étudiant·e·s ont souvent bien des difficultés à pouvoir extraire le design d'une expérience à partir d'un texte descriptif ou d'une explication orale pour pouvoir déterminer quel test statistique est le plus pertinent pour répondre à une question de recherche donnée. Cette compétence complexe est très importante car elle permet de pouvoir comprendre les recherches décrites dans la littérature et d'en évaluer la qualité. De plus, elle permet aux psychologues d'analyser leurs propres données de la manière la plus pertinente pour répondre à leurs questions de recherche.

2.1. Contexte

Bien que la statistique soit une discipline distincte des mathématiques,

La statistique est considérée comme une discipline scientifique à part entière qui fait le lien entre les mathématiques et les sciences (Ben-Zvi & Garfield, 2008). Les mathématiques traitent des nombres alors que la statistique s'intéresse aux données, i.e. des nombres dans un contexte. Sans contexte, les données perdent en effet leur signification et ne peuvent pas être analysées de façon pertinente (Cobb & Moore, 1997). Contrairement à l'enseignements des mathématiques où l'application des formules appropriées représentent une finalité en soit, l'enseignement actuel de la statistique porte principalement sur le choix et l'interprétation de la procédure permettant de modéliser au mieux un ensemble de données, les calculs sous-jacents étant souvent laissés à des programmes spécialisés (Garfield et al., 2008). En effet, certains manuels de statistique promeuvent le fait qu'il est possible de comprendre cette matière sans se frotter à aucune formule (e.g. « Statistiques sans maths pour psychologues », Dancey et al., 2016).

elle n'est pas bien perçue par les étudiant·e·s

La statistique est souvent un passage obligatoire dans de nombreux programmes d'études (Onwuegbuzie & Wilson, 2003; Zanakis & Valenzi, 1997). Au début de leur cursus en

psychologie, les étudiant·e·s ne s'imaginent pas devoir apprendre de manière si approfondie les méthodes de recherche, dont la statistique est l'outil principal, car ils/elles ont tendance à en sous-estimer la pertinence (Ruggeri, Dempster, et al., 2008). Nombre de personnes estiment que les cours de statistique sont les plus difficiles et les moins agréables qu'elles aient suivis durant leur cursus (Berk & Nanda, 1998; Gordon, 2004), ce qui peut les amener à procrastiner et / ou à en repousser la validation (e.g., Onwuegbuzie, 1997). Une étude révèle que 73% des étudiant·e·s interrogé·e·s n'auraient pas validé de cours de statistique si ces derniers n'étaient pas obligatoires (Gordon, 2004).

car elle est généralement associée à de mauvaises expériences en mathématiques

Différentes raisons permettent d'expliquer ce rejet de la statistique, comme l'affect négatif que suscite le sujet, son désintérêt, et la perception d'une absence de pertinence (e.g., Gordon, 2004). Contrairement aux autres cours du cursus de psychologie, la statistique ressemble aux mathématiques (Garfield et al., 2008). Les étudiant·e·s en sciences sociales sont en général peu versé·e·s dans ce domaine et se disent « pas mathématiques » (e.g. Murtonen & Lehtinen, 2003). De mauvais souvenirs liés à l'apprentissage des mathématiques peuvent provoquer une forte appréhension (ou même une profonde anxiété) envers la statistique (e.g. Birenbaum & Eylath, 1994; Schau et al., 1995).

et est considérée comme une matière difficile.

Murtonen et Lehtinen (2003) ont pu identifier un certain nombre de facteurs expliquant pourquoi les étudiant·e·s en sciences sociales éprouvent des difficultés avec les cours sur les méthodes quantitatives (y compris la statistique). En effet, les étudiant·e·s considèrent cette matière comme plus abstraite (par exemple se représenter le concept d'incertitude n'est pas chose facile) et donc plus difficile en comparaison avec les cours de langue par exemple. De plus, il semble probable que les personnes opposent les compétences en langue et en méthodes, de telle sorte qu'il n'est pas possible d'avoir de bonnes compétences en méthodes si elles sont fortes en langue. En outre, la quantité de matière présentée est aussi souvent considérée comme trop importante par rapport à la durée du cours, ce qui donne l'impression que l'enseignant·e n'aborde la matière que de manière superficielle. De plus, un reproche souvent fait aux enseignant·e·s de statistique est le manque de lien entre la théorie et la pratique ainsi que la difficulté à faire des liens entre les différents concepts méthodologiques pour que les étudiant·e·s puissent se créer une image globale de la matière. Par conséquent, les étudiant·e·s n'arrivent souvent pas à percevoir la pertinence que peut avoir la statistique dans la vie

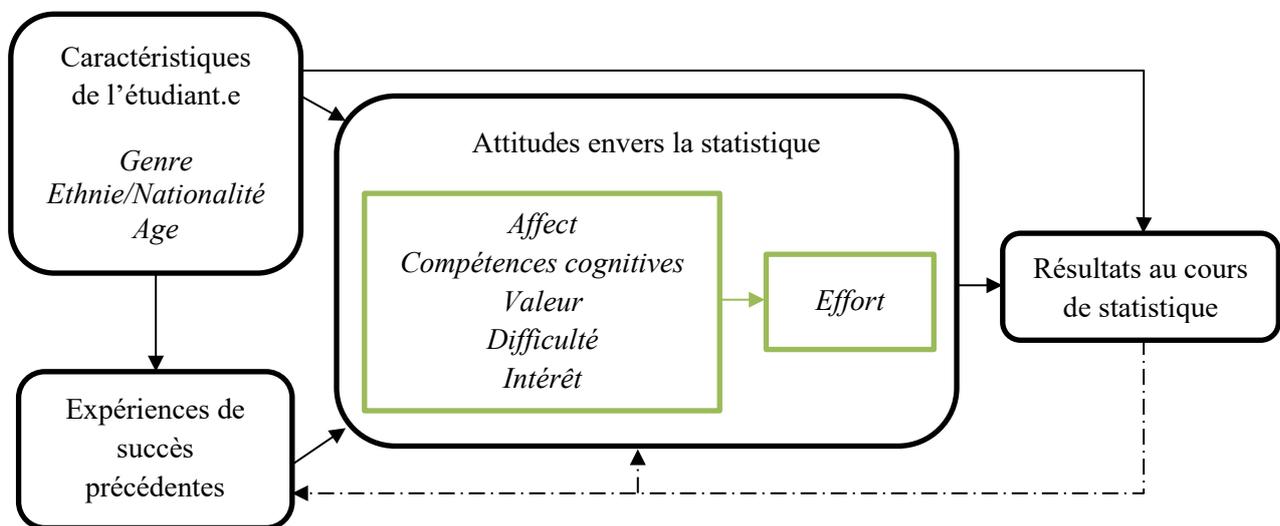
quotidienne : « La terminologie mathématique avec ses abréviations est autant pertinente pour un scientifique humain qu’un vélo tout terrain pour un poisson rouge (Murtonen & Lehtinen, 2003, p. 180) ». A contrario, Ben-Zvi et Garfield (2004) soulignent la difficulté qu’ont les étudiant·e·s à abstraire les concepts statistiques du contexte dans lequel ils ont été appris. Le fait de changer le contexte d’application (comme par exemple lors de la validation du cours) peut faire perdre ses repères à un·e étudiant·e, l’empêchant de résoudre un exercice alors même que le concept est maîtrisé dans son contexte d’apprentissage.

Les différences de perception des difficultés d’apprentissage peuvent être expliquées par différents facteurs, tels que

Ces différents facteurs expliquant les difficultés ressenties par les étudiant·e·s se retrouvent dans le « modèle des attitudes des étudiant·e·s envers la statistique » (ou « *Student’s Attitudes Towards Statistics – Model* », Ramirez et al., 2012). Ce modèle a été développé sur la base des études utilisant le « questionnaire des attitudes envers la statistique » (« *Survey of Attitudes Towards Statistics* » ou *SATS*, voir Schau, 2003; Schau et al., 1995) et vise à prédire les résultats à un cours de statistique sur la base de deux facteurs, à savoir les caractéristiques de l’individu et ses attitudes vis-à-vis de la statistique (voir Figure 1).

Figure 1

Modèle des Attitudes des Etudiant.e.s envers la Statistique.



Note. Adapté de Ramirez et al. (2012, p. 62). Les flèches plaines indiquent des relations observées alors que les flèches pointillées correspondent à des relations supposées.

les caractéristiques individuelles

Ces attitudes sont influencées par les caractéristiques individuelles, telles que le genre, l'ethnie ou la nationalité et l'âge. Concernant le genre, les étudiantes rapportent des attitudes en moyenne moins positives que les étudiants pour l'affect, les compétences cognitives et la difficulté mais des attitudes plus positives en ce qui concerne l'effort (e.g., Rejón-Guardia et al., 2019; van Es & Weaver, 2018). Bien que les résultats soient moins tranchés, des études suggèrent des différences au niveau des attitudes entre les nationalités (e.g., Ruggeri, Díaz, et al., 2008). Enfin, il semble que plus l'âge d'entrée à l'université est élevé, plus les personnes disent y trouver un intérêt personnel et y investir leur effort (Milic et al., 2016) mais les résultats des études à ce sujet (e.g. Coetzee & van der Merwe, 2010; Verhoeven, 2011) sont difficilement généralisables en raison de nombreux facteurs contextuels, comme la nationalité par exemple.

et les expériences antérieures avec la matière.

Comme suggéré plus haut déjà, les expériences précédentes en lien avec la statistique et / ou les mathématiques ont un impact sur les attitudes envers la statistique. La littérature montre que les expériences de succès conduisent à des attitudes positives (en particulier pour l'affect, les compétences cognitives et l'effort) alors que l'inverse se produit avec des expériences négatives (e.g., Dempster & McCorry, 2009).

L'ethnie peut influencer les attentes de performance.

Van Es et Weaver (2018) ont observé que bien que n'influençant pas les attitudes, l'ethnie a un impact sur les attentes et les notes obtenues. Les personnes d'ethnie asiatiques ont obtenu les meilleurs résultats qu'elles pensaient effectivement faire alors que les personnes d'ethnie afro-américaine et hispanique ont reçu des notes plus basses que ce qu'elles attendaient.

Une attitude négative péjore l'apprentissage de la matière

De manière générale, une attitude négative envers une matière tend à péjorer la qualité de son apprentissage (Waters et al., 1988). Des études suggèrent que les étudiant·e·s ont a priori des attitudes neutres voir positives par rapport à la statistique avant de commencer à étudier cette matière mais que ces dernières se péjorent au fil du temps (Carillo et al., 2016; Kerby & Wroughton, 2021; Schau & Emmioğlu, 2011).

et la performance à l'évaluation.

De plus, un grand nombre de recherches montrent un lien entre les attitudes envers la statistique et la performance à l'évaluation, à savoir que plus les attitudes envers la statistique sont

positives, plus la performance à l'évaluation est haute (e.g. Carlson & Winqvist, 2011). Ce lien semble particulièrement fort en ce qui concerne les compétences cognitives (e.g. Bandura, 1997; Gist & Mitchell, 1992).

Il est important de promouvoir une attitude plus positive des étudiant·e·s envers la statistique

La littérature suggère des objectifs pour améliorer l'attitude des étudiant·e·s vis-à-vis de la statistique (e.g. Emmioğlu & Capa-Aydin, 2021; Ramirez et al., 2012), à savoir les amener :

- A développer leur confiance en soi qu'ils / elles peuvent comprendre et utiliser la statistique
- A prendre conscience que la statistique est utile tant dans la vie professionnelle que privée
- A reconnaître que cette matière peut être intéressante
- A fournir l'effort nécessaire pour apprendre la façon de réfléchir et les compétences statistiques
- A réaliser que la statistique n'est ni trop simple, mais surtout ni trop difficile à apprendre

en renforçant leur sentiment d'efficacité

L'un des moyens pour parvenir à ces objectifs est d'augmenter le sentiment d'efficacité personnelle, à savoir le niveau de conviction qu'a une personne d'être capable d'atteindre les objectifs que elle s'est fixés (Bandura, 1997). Le sentiment d'efficacité personnelle comporte deux aspects, à savoir le sentiment de compétence (*efficacy expectation*) et le sentiment de la valeur du résultat (*outcome expectation*). Le sentiment de compétence reflète le niveau de conviction d'une personne quant à sa capacité à réaliser une tâche ou une action dans un contexte spécifique, alors que le sentiment de la valeur du résultat correspond au niveau de conviction d'une personne que les tâches ou actions qu'elle a réalisées vont aboutir aux résultats escomptés.

pour améliorer leur performance lors des évaluations.

Selon la revue de littérature de Usher et Parajes (2008), ce sont les expériences précédentes de réussite ou d'échec à une unité d'enseignement (de statistique) ou à une matière similaire qui façonnent la perception de l'efficacité personnelle de l'étudiant·e. Cette perception va ensuite influencer son comportement (en bien ou en mal) vis-à-vis de cette matière et son résultat à l'évaluation.

2.2. Objectifs du projet

Le premier objectif de ce projet est de permettre aux étudiant·e·s d'augmenter leur sentiment d'efficacité personnelle vis-à-vis de la résolution de problèmes statistiques à l'aide de quatre quizz formatifs en ligne. En effet, la littérature montre qu'un haut sentiment d'efficacité personnelle influence positivement l'attitude envers une matière (comme la quantité d'effort déployé, e.g. Bandura, 1997; Schunk et al., 2014). La réussite des quizz pourrait augmenter le nombre d'expériences de succès, l'un des facteurs participant à la génération d'une attitude positive envers la matière (voir le model SATS, Figure 1). Selon la revue de littérature de Usher et Pajares (2008), les expériences précédentes représentent le plus grand déterminant de l'efficacité personnelle, et plus particulier le sentiment de la valeur du résultat (Bandura, 1997) dans le cas présent.

Le second objectif de ce projet vise à fournir aux étudiant·e·s des alternatives aux moyens de transmission en présence (i.e., le cours et les séminaires) avec lesquels ils/elles puissent s'entraîner le plus possible à développer la séquence d'analyse nécessaire pour pouvoir choisir le test approprié dans le but de répondre à une question de recherche. Les TIC ont été utilisées pour que les étudiant·e·s puissent faire les quizz en dehors des heures de cours. Les quizz dans ce projet ont été réalisés sur Moodle 2, la plate-forme d'apprentissage en ligne principale utilisée par l'Université de Fribourg.

Au travers des quizz, l'étudiant·e est amené·e à disséquer une question de recherche pour pouvoir déduire le test statistique le plus approprié en fonction du type de variables qui sont présentées dans l'énoncé d'une situation expérimentale. Ce type de résolution de problème constitue une question classique d'examen de statistique et représente l'une des tâches principales du chercheur et de la chercheuse en psychologie expérimentale. Dans ce projet, je me suis focalisé sur les analyses comparant des conditions expérimentales entre elles, ce qui correspond en grand partie à la matière enseignée en première et deuxième année de psychologie à l'Université de Fribourg.

Le premier quizz porte sur l'analyse de l'échelle d'une variable (quizz nommé « Type d'échelle » dans Moodle). Le deuxième permet aux étudiant·e·s de déterminer si la variable est une variable indépendante ou dépendante (quizz « Nature de la variable »). Le troisième concerne la sélection du test approprié (quizz « Type de test »). Finalement, le dernier quizz permet d'exercer ces trois sous-compétences simultanément (quizz « Etude de cas ») et donc de tester la compétence principale dans sa totalité (i.e. la déduction de l'analyse statistique

appropriée en fonction d'une situation expérimentale donnée). Un retour est donné automatiquement pour chaque question de chaque quizz et chaque participant·e peut faire chaque quizz autant de fois qu'il/elle le souhaite.

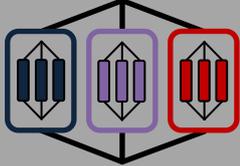
Selon la taxonomie d'Anderson et Krathwohl (2001), l'objectif de performance final visé par les quizz est l'analyse. En effet, les participant·e·s doivent être capables de distinguer la nature (dépendante vs indépendante) ainsi que le type d'échelle (nominale, ordinale ou quantitative) d'une variable et le type d'analyse statistique à employer sur la base d'une description d'une situation expérimentale.

Cet objectif de performance correspond au niveau relationnel de la taxonomie SOLO (*Structure of the Observed Learning Outcome*) de Biggs et collaborateurs (1982; 2011 ; voir Figure 2). Ce modèle centré sur l'apprentissage est composé de cinq niveaux. Les trois premiers niveaux (préstructurel, unistructurel et multistructurel) évaluent le nombre de connaissances acquises par l'apprenant·e (phase quantitative), alors que les deux derniers niveaux (relationnel et abstrait étendu) le niveau de compréhension de ses connaissances. Contrairement à la taxonomie d'Anderson et Krathwohl (2001), les tâches et leurs résultats peuvent se situer à des niveaux différents.

Selon mon évaluation, le premiers quizz « Type d'échelle » et « Nature de la variable » testent des connaissances de niveau unistructurel car l'étudiant·e doit seulement être capable d'identifier correctement une des deux caractéristiques d'une variable. Le troisième quizz « Type de test » évalue l'objectif de performance final, qui, comme susmentionné, est de niveau relationnel. Le dernier quizz « Etude de cas » est aussi du niveau relationnel mais inclut les deux premiers quizz de niveau unistructurel pour forcer les étudiant·e·s à prendre en compte l'ensemble de la situation expérimentale pour les guider vers le test approprié.

Figure 2

Illustration de la Taxonomie SOLO des Connaissances selon Biggs and Tang (2011).

	Phase quantitative			Phase qualitative	
					
Niveau	<i>Préstructurel</i> Pas de connaissance	<i>Unistrukturel</i> Un élément pertinent	<i>Multistructurel</i> Plusieurs éléments indépendants	<i>Relationnel</i> Relations entre les éléments	<i>Abstrait étendu</i> Généralisation à un autre domaine
Déclaratives	Echouer Manquer	Mémoriser Identifier Réciter	Décrire Classer	Comparer Contraster Expliquer Argumenter Analyser	Théoriser Prédire Généraliser Classer
Fonctionnelles		Compter Associer Ordonner	Calculer Illustrer	Appliquer Construire Traduire Prédire	Réfléchir et améliorer Inventer, créer Résoudre un problème nouveau Extrapoler

« T'échappes à la police, pas aux statistiques. »

Jean-Jacques Goldman

3. Mise en œuvre du projet

3.1. Acteurs et actrices

La page Moodle contenant les quizz est en ligne depuis 2016. 214 étudiant·e·s de psychologie de deuxième année des années académiques 2016-2017 et 2017-2018 ont reçu le lien vers cette page dans le cadre du cours « Statistique et analyse de données II et tutorat ». 34 personnes (16% du public visé) ont visité les différents quizz, alors que 27 personnes (13 % du public visé) y ont effectivement pris part. Finalement, 25 participant·e·s (24 avec la psychologie en domaine principal, un·e en sciences de la communication et média) ont répondu au questionnaire d'évaluation.

3.2. Comment résoudre le problème ?

Décider quelle analyse statistique effectuer dépend de la question de recherche posée ainsi que de la situation expérimentale. Voici un exemple :

« Un groupe de chercheuses et de chercheurs souhaitent observer s'il existe une différence entre les hommes et les femmes quant à leur capacité de raisonnement spatial. Pour ce faire, 100 participants (50 hommes et 50 femmes) passent un test de rotation mentale contenant 30 questions. Chaque item est composé de deux images représentant une forme abstraite composée de cubes. Chaque forme est présentée sous un angle de vue différent. Le but de chaque participant·e est d'indiquer si les deux images représentent ou non le même objet. Les deux formes sont identiques après rotation mentale pour la moitié des items alors que ce n'est pas le cas pour l'autre moitié. Le score total de chaque participant·e correspond au pourcentage de réponses correctes. La distribution des scores totaux est normale dans chaque groupe. »

Pour résoudre ce problème, il faut identifier les variables pertinentes par rapport à la question de recherche car c'est leur nature qui va déterminer le choix de l'analyse. En statistique humaine, une variable correspond à une information spécifique appartenant à un·e participant·e.

Dans notre exemple, il y en a deux, à savoir (1) le genre de participant / de la participante et (2) son score au test de rotation mentale (TRM).

La première étape consiste à déterminer la nature de la variable. En effet, dans une expérience, on cherche à déterminer un lien causal (ou du moins envisagé comme tel) entre, ad minima, deux variables. La variable qui est sensée être la cause de l'effet est appelée variable indépendante VI (car elle est supposée ne pas être influencée par l'autre variable) alors que la variable qui est supposée être influencée par l'autre dans cette relation est nommée variable dépendante VD (car sa caractéristique ou sa valeur devrait « dépendre » de la variable indépendante). Dans notre exemple, le genre est la variable indépendante et le score au TRM la variable dépendante (donc théoriquement influencé par le genre de la personne qui répond).

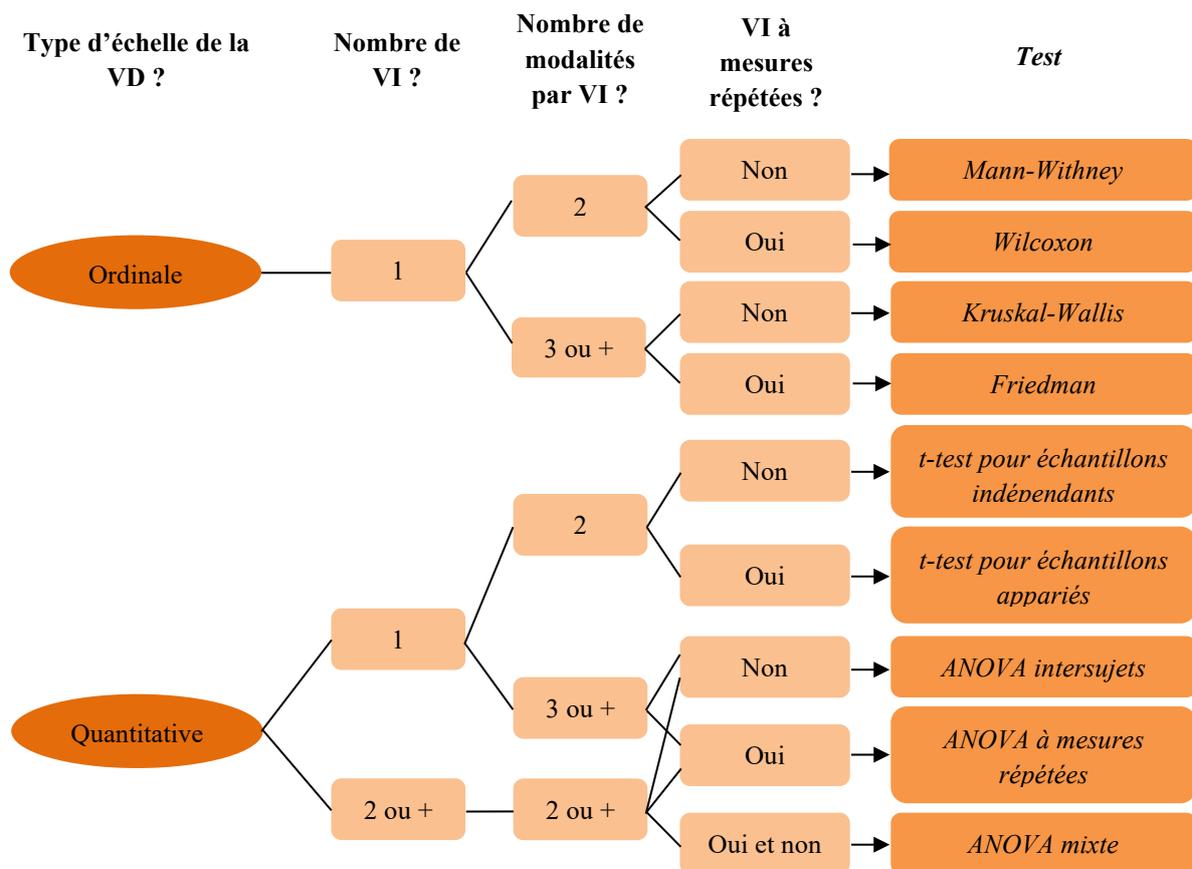
La deuxième étape consiste à déterminer quelle est l'échelle de chaque variable. L'échelle représente le type de liens qu'il existe entre les différentes modalités d'une variable. Il existe quatre types d'échelles. (1) Une échelle nominale est faite de catégories. Il n'y a aucune relation d'ordre entre les modalités de la variable (par exemple, la couleur des yeux). (2) Une échelle ordinale implique un ordre ou une hiérarchie entre les modalités d'une variable mais les écarts entre elles ne sont pas identiques (par ex. les grades militaires). (3) Une échelle d'intervalles, quant à elle, implique un intervalle identique entre chaque valeur et peut avoir des valeurs négatives (e.g. les degrés Celsius, le quotient intellectuel). Enfin, (4) une échelle de rapport a, en plus des qualités d'une échelle d'intervalle, un zéro absolu, à savoir que les valeurs ne peuvent pas aller en dessous de zéro (e.g. la taille, la vitesse), ce qui permet ainsi d'avoir des scores proportionnels. Les échelles d'intervalles et de rapport sont statistiquement traitées de la même manière et sont généralement regroupées sous le terme « échelles quantitatives ». Dans notre exemple, le genre est une variable nominale alors que le score au TRM est une variable quantitative (plus précisément de rapport).

La dernière étape consiste à déterminer le type de test le plus approprié pour répondre à la question de recherche. En plus de l'analyse effectuée lors deux étapes précédentes, trois informations supplémentaires concernant la (les) variable(s) indépendante(s) sont nécessaires pour y parvenir, à savoir (1) leur nombre (une ou plus d'une), (2) le nombre de modalités les composant (pour les échelles nominales) et (3) la dépendance des mesures, c'est-à-dire si les mesures sont répétées chez les mêmes sujets (analyse inter-sujets) ou effectuées dans des groupes différents, selon les modalités de la variable indépendante (analyse intra-sujets). La Figure 3 (inspirée de Field et al., 2012) présente toutes les analyses comparant des groupes ainsi que leurs caractéristiques propres. Dans notre exemple, il y a *une* seule variable indépendante

ayant *deux modalités* (i.e. femme – homme) et un·e participant·e ne peut se retrouver que dans une catégorie (analyse inter-sujets). Enfin, la variable dépendante est *quantitative*. Toutes ces informations permettent d'utiliser de façon correcte la Figure 3 pour trouver le chemin menant à la meilleure analyse. En suivant les indications de cette figure, l'analyse la plus pertinente est donc le *T de Student pour groupes indépendants*.

Figure 3

Arbre Décisionnel pour le Choix du Test Approprié.



Note. Adapté de Field et al. (2012) ; VD = Variable Dépendante ; VI = Variable Indépendante.

3.3. Conception des items

Deux options se sont présentées lors de la génération des situations expérimentales : utiliser des exemples fictifs ou des articles scientifiques contenant les analyses souhaitées. Bien que la première aurait été plus simple et plus rapide à mettre en place, j'ai choisi la seconde pour une question de validité écologique. Elle permet de mieux ancrer les situations expérimentales dans un contexte plus « authentique » que si elles avaient été inventées de toutes pièces, ce qui

devrait permettre aux participant·e·s de consolider le lien entre la théorie vue en classe et la pratique scientifique. Le renforcement de ce lien est vu comme central dans l'établissement d'une attitude positive envers la statistique (American Statistical Association, 2016).

Mon objectif était de trouver au moins un article pour chacune des neuf analyses entraînées (voir Figure 3) en fonction des deux critères suivants : la situation expérimentale ne devait pas être trop complexe (pour ne pas démotiver les participant·e·s), et les questions de recherche devaient avoir un lien avec la psychologie et provenir de domaines variés (e.g. psycholinguistique, psychologie clinique, psychologie cognitive, etc.). A cause de ces contraintes, il ne m'a été possible de trouver que huit articles après de longues heures de recherche. Je n'ai pas pu trouver des articles contenant des analyses non-paramétriques comparant plus de deux groupes (i.e. tests de Friedman et Kruskal-Wallis) ainsi que des ANOVA purement inter-sujets ou à mesures répétées (voir Tableau 1). Toutes les situations expérimentales ainsi que les questions posées se trouvent dans l'Annexe A.

Tableau 1

Source des Situations Expérimentales et Analyse Correspondante.

Analyse	Article
Test de Mann-Whitney	Loftus (1977) Medalia et al. (1998)
Test du signe de Wilcoxon	Nosarti et al. (2002)
Test de Friedman	-
Test de Kruskal-Wallis	-
T-test pour échantillons indépendants	Barceló and Knight (2002) Gotlib et al. (2004)
T-test pour échantillons appariés	Wagner et al. (2004)
ANOVA intersujets	-
ANOVA à mesures répétées	-
ANOVA mixte	Plaisted et al. (1998) Sato et al. (2013)

L'enseignante de statistique avec laquelle il était prévu de faire le contrôle de qualité de la description des situations expérimentales ainsi que des retours n'a malheureusement pas pu

remplir cette tâche. Cependant, le fait que certain·e·s participant·e·s obtiennent de très bonnes performances (voir point 3.5 *Analyses des données des quizz*) et que les réponses du questionnaire d'évaluation soient très positives (voir point 4 *Evaluation du projet par les étudiant·e·s*) suggère un (très) bon niveau de compréhension des quizz.

3.4. Construction du cours Moodle et des quizz

La première section du cours Moodle contient une introduction concernant les quizz ainsi que leurs modalités d'évaluation (voir Figure 4). A la fin de cette section, une incitation à évaluer les quizz ainsi qu'un lien vers les références des articles utilisés dans le projet sont présentés. Les quatre sections suivantes décrivent brièvement chaque quizz séparément et proposent le lien menant vers le quizz (voir Annexe B). Enfin, la dernière section explique l'importance de leur retour sur ce projet et leur fournit un lien vers un questionnaire (voir point 4 *Evaluation du projet par les étudiant·e·s*).

Figure 4

Première Section du Cours Moodle

Bienvenu-e-s dans le Quizz Stat!

Ce module vous permettra d'évaluer et d'exercer vos connaissances de base en statistiques. Il est composé de trois tests portant chacun sur un contenu spécifique. Un feedback vous est donné après chaque question.

- Le premier vous permet de tester vos connaissances sur le **type d'échelle** d'une variable (nominale, ordinale ou quantitative)
- Le deuxième porte sur la **nature des variables** (variables indépendantes ou dépendantes)
- Le troisième se focalise sur le **type de test** à employer en fonction des différentes variables à disposition

Enfin, un dernier test vous propose de tester l'ensemble de ces connaissances sur 3 situations de recherche.

Ces tests sont formatifs! N'hésitez pas à les refaire autant de fois que vous le souhaitez. L'ordre et les questions sont différents à chaque fois (une banque de 9 questions a été créée pour chaque test dont 3 sont choisies aléatoirement).

Chaque question provient de l'une des neuf recherches publiées dans une revue scientifique. Vous en trouverez la liste en cliquant sur le lien ci-dessous.

N'oubliez pas de remplir le questionnaire d'évaluation du quizz une fois que vous avez essayé tous les tests! Le lien vers le questionnaire se trouve dans la dernière section du quizz.

 [Références quizz stat](#)

Chaque quizz comporte trois situations expérimentales (i.e. trois questions Moodle). Chaque question Moodle commence par la description de la situation expérimentale, suivie d'un nombre variable de QCM en fonction de la nature de la situation expérimentale. Dans les quizz « Type d'échelles » et « Nature de la variable », il y avait autant de QCM que de variables impliquées dans la situation expérimentale (i.e. deux ou trois) et il n'y avait qu'une QCM par situation pour le quizz « Type de tests ». Enfin, le nombre de QCM pour le quizz « Etude de

cas » correspond à la somme des QCM des trois autres quizz pour chaque situation expérimentale, à savoir cinq ou sept QCM.

L'intitulé de chaque QCM concernant le type de mesure est : « La variable *XXX* est : » et les participant·e·s doivent choisir parmi trois options, à savoir : *nominale*, *ordinaire* ou *quantitative*. Pour le quizz « Nature de la variable », l'intitulé est le même que pour la question précédente, mais il n'y a cette fois que deux possibilités de réponse, soit *indépendante* ou *dépendante*. Enfin, concernant le type de test, l'intitulé de chaque question est : « Quelle analyse *le groupe de recherche* doit-il utiliser pour vérifier leur hypothèse ? » avec neuf réponses à choix (*un test de Mann-Whitney, un test de Wilcoxon, un test de Friedman, un test de Kruskal-Wallis, un t-test à mesures indépendantes, un t-test à mesure appariées, une ANOVA intersujets, une ANOVA à mesures répétées ou une ANOVA mixte*).

Figure 5

Illustration d'une Question Moodle

Gotlib et al. (2004) souhaitent observer si les personnes souffrant d'un épisode dépressif majeur présentent des biais attentionnels plus importants au niveau du traitement de stimuli interpersonnels négatifs que les personnes souffrant d'anxiété généralisée et des personnes contrôle. Avant de faire passer leur expérience, ils veulent vérifier que les personnes souffrant d'un épisode dépressif majeur qu'ils ont recrutées ont un niveau de dépression plus élevées que celles des autres groupes. Pour ce faire, ils leur font passer le BDI (Beck Depression Index, scores de 0 à 63).

- La variable type de participant·e·s est-elle:
 - indépendante
 - dépendante
- La variable *dépression* (mesurée par le score au BDI) est-elle:
 - indépendante
 - dépendante

Merci de répondre à tous les éléments de la question.

Vérifier

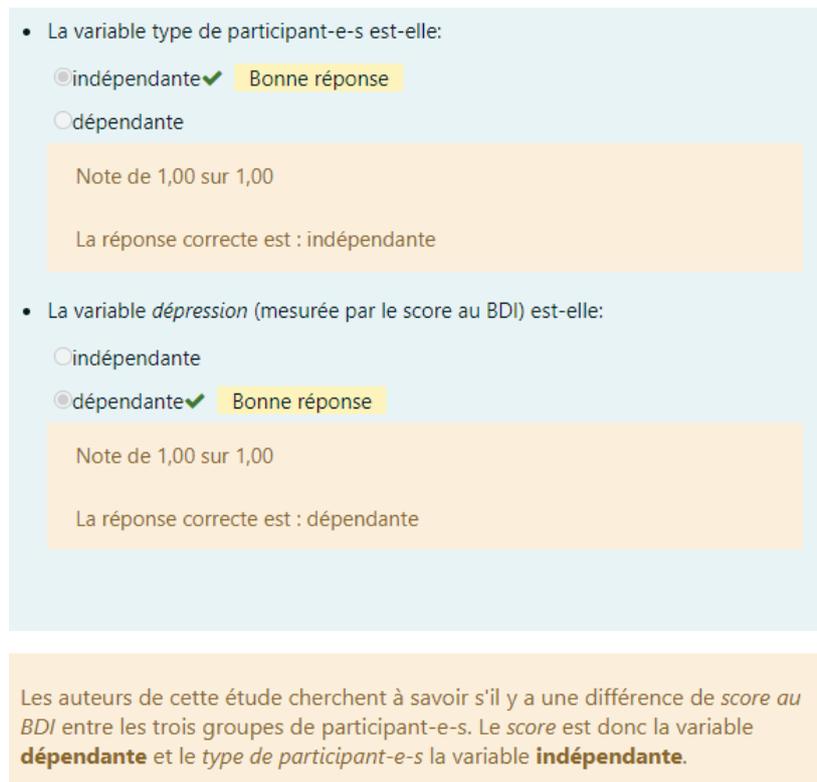
Note. Illustration provenant du quizz « Nature de la variable ».

Pour pouvoir vérifier les réponses (à l'aide du bouton « Vérifier »), les participant·e·s doivent répondre à chaque QCM liée à une situation expérimentale. Si ce n'est pas le cas, le message « Merci de répondre à tous les éléments de la question » apparaît en rouge (voir Figure 5). Deux types de retour sont fournis, l'un par rapport à la réponse donnée, et l'autre par rapport à la situation expérimentale en général (voir Figure 6). Le retour par rapport à la réponse donnée

indique si la réponse est juste ou fausse et la réponse correcte est donnée systématiquement en dessous chaque QCM. Le retour général, présenté dans un encadré à la fin de chaque question Moodle, permet de comprendre quelles sont les bonnes réponses.

Figure 6

Illustration du Retour à une Question Moodle.



The screenshot shows two quiz questions. The first question asks 'La variable type de participant-e-s est-elle:' with options 'indépendante' (selected and marked correct) and 'dépendante'. It shows a score of 1.00/1.00 and the correct answer 'indépendante'. The second question asks 'La variable *dépression* (mesurée par le score au BDI) est-elle:' with options 'indépendante' and 'dépendante' (selected and marked correct). It also shows a score of 1.00/1.00 and the correct answer 'dépendante'. Below the questions, a summary box states: 'Les auteurs de cette étude cherchent à savoir s'il y a une différence de *score au BDI* entre les trois groupes de participant-e-s. Le *score* est donc la variable **dépendante** et le *type de participant-e-s* la variable **indépendante**.'

Note. Illustration provenant du quizz « Nature de la variable ».

3.5. Analyses des données des quizz

Le Tableau 2 présente la fréquentation des différents quizz. De plus, cinq personnes ont refait un quizz au moins une fois. Les performances à un quizz correspondent au pourcentage moyen de réponses correctes sur toutes les tentatives complètes d'une personne. Une tentative est considérée comme complète lorsqu'une personne a donné une réponse à chaque question posée. Ainsi, nous observons un taux moyen de réponses correctes de 78% pour le quizz « Type d'échelle », de 92% pour le quizz « Nature de la variable », de 35% pour le quizz « Type de test » et de 86% pour le quizz « Etude de cas ».

Ces analyses montrent que les étudiant·e·s ayant participé aux quizz sont plus nombreux· et nombreuses à déterminer la nature de la variable et de l'échelle de cette dernière qu'à choisir

l'analyse correcte. De plus, les performances sont nettement plus basses pour le quizz « Type de test » que pour « Nature de la variable » et « Type d'échelle ». Cela pourrait suggérer que les participant·e·s de ces quizz sont capables d'analyser correctement certains aspects spécifiques du problème mais n'auraient pas forcément encore acquis un niveau de connaissances suffisant leur permettant d'analyser correctement la situation expérimentale dans son ensemble. Ce constat est renforcé par le fait que le taux de réponses correctes au quizz « Cas d'études » est plus élevé que pour le quizz « Type de test ». En effet, cette différence suggère que les étudiant·e·s bénéficient de la présence explicite des questions sur la nature de la variable et le type d'échelle pour pouvoir choisir le test approprié. Sans ces questions, présentes dans le quizz « Cas d'études », il leur est plus difficile d'analyser correctement la situation expérimentale.

Selon le modèle SOLO (Biggs & Collis, 1982), les étudiant·e·s qui ont pris part à ce projet se situeraient principalement au niveau multistructurel car ils / elles sont capables d'identifier les caractéristiques pertinentes des variables de façon individuelle mais n'arrivent pas à les combiner ces informations pour choisir le test pertinent de manière spontanée. Cependant, s'ils / elles sont soutenu·e·s dans leur prise de décision, ils / elles arrivent à atteindre le niveau relationnel. Cette différence entre les quizz « Type de test » et « Etudes de cas » suggère qu'entraîner les étudiant·e·s à systématiquement prendre en considération le type d'échelle et la nature de la variable les aiderait à atteindre un niveau de connaissance relationnel.

Tableau 2

Effectifs en Fonction du Type de Quizz et de l'Interaction avec la Page Moodle

Quizz	Type d'échelle	Nature de la variable	Type de test	Cas d'étude
Visite seule	7	3	8	10
Participation	24	26	16	11
Total	31	29	24	21

Les implications tirées de l'analyse des données objectives de ce projet sont à considérer avec prudence. En effet, seulement 13% du public visé a effectivement pris part aux quizz, ce qui réduit la généralisabilité de ces implications. La participation aux quizz étant volontaire et

n'apportant pas de bénéfice direct (comme une augmentation de la note finale au cours), les étudiant·e·s n'ont peut-être pas jugé pertinent d'y consacrer du temps. De plus, la présentation de ces quizz comme soutien à la réussite de l'évaluation n'a peut-être pas été suffisamment marquée pour susciter l'intérêt. Enfin, la page Moodle a été ouverte au semestre d'automne 2016. Il est possible que les étudiant·e·s n'aient pas perçu la pertinence de ces quizz auto-évaluatifs pour l'évaluation finale sur le moment et en ont oublié lors de leurs révisions. Il aurait été donc plus pertinent de mieux insérer les quizz dans le cadre du cours « Statistique et analyse des données et tutorat II » pour en augmenter la pertinence et d'en rappeler l'existence quelques temps avant l'évaluation.

Il est aussi intéressant de noter qu'un quart des personnes ayant visité la page Moodle n'ont pas fini un seul quizz. Cela peut suggérer que ces tests formatifs n'ont pas convaincu au premier coup d'œil. Peut-être l'utilisation du format quizz n'a pas été appréciée comme soutien à l'apprentissage. Il est aussi possible que l'utilisation de situations expérimentales réelles rendent les exercices trop compliqués, ce qui a pu décourager certains.

De plus, bien que le descriptif général des quizz stipule qu'il est possible de les repasser plusieurs fois, l'analyse du nombre de tentatives par quizz montre que seulement cinq personnes sur 27 ont fait au moins deux fois le même quizz. Cela suggère que la grande majorité des participant·e·s n'a pas pleinement profité des quizz. Il aurait peut-être été pertinent d'ajouter une page à la fin de chaque quizz rappelant la possibilité de refaire le quizz, en rappelant sa pertinence pour l'évaluation finale.

« La seule véritable erreur est celle dont on ne retire aucun enseignement. »

John Powell

4. Evaluation du projet par les étudiant·e·s

4.1. Description

Une fois le(s) quizz fini(s), les étudiant·e·s étaient invité·e·s dans la dernière section du cours Moodle à évaluer le projet à l'aide d'un questionnaire anonyme programmé avec Google Forms (voir Figure 7). Cette évaluation contenait quatre parties (voir Tableau 3). La première concernait l'affiliation du/de la participant·e. La seconde portait sur la pertinence d'un tel support en ligne (quatre questions). La troisième reprenait six questions reprise du questionnaire pour les étudiant·e·s développé dans le cadre du projet HY-SUP visant à évaluer l'impact des cours hybrides sur l'apprentissage (voir section 20 du questionnaire ; Deschryver & Charlier, 2012). Enfin, la dernière partie permettait aux participant·e·s de faire part de leurs commentaires additionnels de manière libre.

Figure 7

Section sur l'Evaluation des Quizz.

Questionnaire

Dans le but de mesurer et d'améliorer l'utilité de ce quizz, pourriez-vous remplir le questionnaire suivant? [Lien vers le questionnaire](#)

Ce questionnaire est anonyme et vos données ne seront pas traitées à un niveau individuel.

Merci d'avance pour votre participation!

Tableau 3

Questions (et Format de Réponse) du Formulaire d'Evaluation du Projet en Fonction des Différentes Parties.

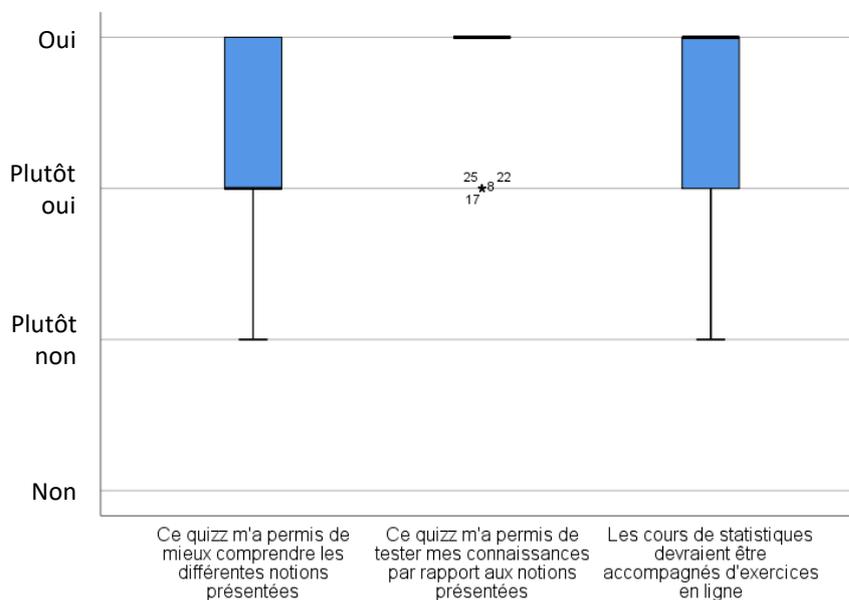
Partie	Question	Format de réponse
A. Affiliation	A.1. Affiliation / domaine d'études principal	Échelle nominale à quatre catégories (Unifr / Psychologie, Unifr / Sciences de l'éducation, HES / Haute école de santé, Autre)
B. Pertinence	B.1. Ce quizz m'a permis de mieux comprendre les différentes notions présentées	Échelle de Likert à quatre niveaux (Non, plutôt non, plutôt oui, oui)
	B.2. Ce quizz m'a permis de tester mes connaissances par rapports aux notions présentées	
	B.3. Les cours de statistiques devraient être accompagnés d'exercices en lignes	
	B.4. Le nombre de questions par quizz est suffisant	Échelle dichotomique (oui/non)
	B.5. Si vous avez répondu non à la question précédente, veuillez indiquer combien de questions vous aimeriez avoir par test ?	Réponse ouverte brève
C.	<i>Grâce à ce quizz...</i>	Échelle de Likert à quatre niveaux (Non, plutôt non, plutôt oui, oui)
	C.1. Je suis d'avantage actif	
	C.2. J'apprends davantage (en quantité)	
	C.3. Mes apprentissages sont de meilleures qualité	
	C.4. Les activités proposées sont plus proches de celles de la vie professionnelle	
	C.5. Je suis davantage en démarche scientifique de recherche	
C.6. Le cours m'amène davantage à réfléchir sur ma manière d'apprendre		
D. Commentaire	D.1. Tous vos commentaires et remarques sont les bienvenus!	Réponse ouverte brève

4.2. Résultats

Les résultats de la partie B du questionnaire d'évaluation (voir Figure 8) montrent que les étudiant·e·s pensent que les quizz les ont aidé·e·s à mieux comprendre les notions exercées (médian de 3 sur 4 pour la question B.1). De plus, ils / elles pensent que les quizz permettent de tester leurs connaissances au sujet des notions proposées (congruence des exercices avec la matière ; médian de 4 sur 4 pour la question B.2). Les étudiant·e·s pensent aussi que l'hybridation des cours de statistiques est importante (médian de 4 sur 4 pour la question B.3). Ils semblent apprécier le fait de pouvoir tester leurs connaissances à distance et recevoir un retour instantané. Ces assomptions doivent cependant être confirmées à l'aide d'un questionnaire plus précis. Enfin, la question B.4 concernant le nombre de questions présente des résultats moins tranchés. Un peu plus de la moitié des participant·e·s (13 sur 25) estiment qu'il faut plus de questions par quizz que les trois proposés. Celles et ceux qui y ont répondu souhaiteraient en avoir entre 5 et 20 (question B.5).

Figure 8

Boîtes à Moustaches pour les Items B1, B2 et B3 du Questionnaire d'Evaluation.



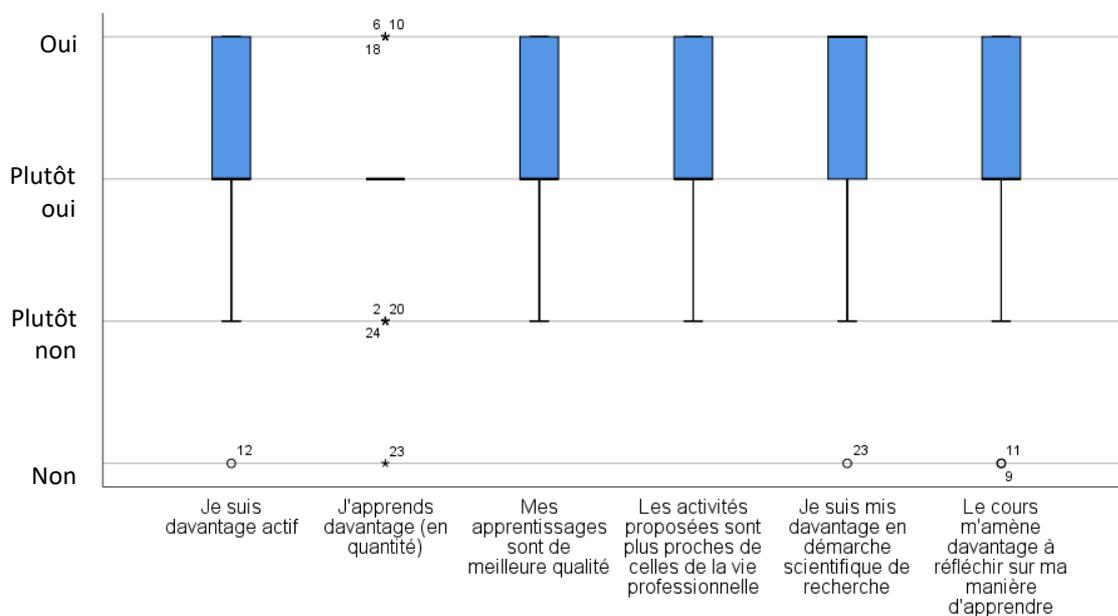
Les résultats obtenus à la partie B suggèrent que les quizz permettent de mieux comprendre les notions exercées et de les tester de manière satisfaisante. Il semble que les participant·e·s soient aussi intéressé·e·s par des outils d'apprentissage complémentaires à ceux utilisés en cours. Cela confirme l'intérêt et la pertinence de ces quizz formatifs en ligne pour les personnes y ayant pris part. Les commentaires fournis supportent l'hypothèse que les participant·e·s n'ont pas

compris qu'il est possible de faire plusieurs fois le même quizz mais avec des situations expérimentales différentes (voir point 3.5).

Les résultats de la partie C concernant l'impact d'un cours hybride sur l'apprentissage sont aussi encourageants car les quizz semblent avoir favorisé leur apprentissage des notions testées (voir Figure 9). Il semble que les participant·e·s soient davantage actifs /actives lorsqu'ils / elles apprennent (médian de 3 sur 4 pour la question C.1) et apprennent plus (médian de 3 sur 4 pour la question C.2) et mieux (médian de 3 sur 4 pour la question C.3). De plus, ils / elles affirment réfléchir davantage sur leur manière d'apprendre (médian de 3 sur 4 pour la question C.6). Enfin, il est intéressant de constater que les quizz ont permis aux participant·e·s d'entrer dans une démarche scientifique (médians de 3 et 4 sur 4 pour les questions C.4 et C.5 respectivement).

Figure 9

Boîtes à Moustaches pour les Items C1 à C6 du Questionnaire d'Evaluation



4.3. Analyses des commentaires et suggestions

Comme prévu, peu de personnes ayant répondu au questionnaire (5 sur 25) ont laissé des commentaires. Les commentaires sont plutôt positifs. Ils montrent que les exercices proposés sont porteurs de sens pour les participant·e·s et facilitent l'apprentissage grâce à la concrétude des situations expérimentales comme le montrent le commentaire suivant : « ... ça m'a bien aidée en général, notamment à comprendre les différences entre t-test et ANOVA ». Avoir la

possibilité d'exercer à volonté des notions de statistique semblent plaire : « Je trouve très positif d'avoir accès à un tel outils d'exercice et je pense que la pratique est absolument nécessaire à la compréhension de ce genre de matière ! ». Les commentaires soulignent aussi l'importance de faire des liens entre différentes notions, comme c'est le cas dans le dernier quizz : « Le dernier quizz liant les 3 notions est très instructif ! Il oblige à réfléchir aux liens entre ces 3 notions et permet de mieux les comprendre à travers des situations concrètes. », « J'ai trouvé bien le dernier quizz avec le choix du test statistique ».

Quelques personnes ont proposé des améliorations. Un commentaire accentue le manque d'exercices par quizz (déjà constaté grâce à la question B.4) : « Il manquerait encore un exercice de plus afin d'approfondir la réflexion. ». De plus, il semble que quelqu'un a remarqué que le test ne couvrait malheureusement pas tous les tests comparant des groupes : « C'était un très bon test, mais à mon avis, il manquerait juste une question qui ferait un exemple de variable ordinale. ». Il est possible que la personne n'ait fait qu'une seule fois les quizz et que la randomisation des questions ne lui ait pas proposé d'exercices avec des variables ordinales. De manière indirecte, ce commentaire souligne à nouveau le peu d'exercices proposés par quizz ainsi que leur manque de variété. Une personne aurait souhaité un rappel de la théorie (probablement au début du cours Moodle) « afin de se rafraîchir la mémoire ou de modifier de faux apprentissages ». De plus, elle souligne que « certaines explications sont plus complètes que d'autres ». Enfin, elle aimerait pouvoir recommencer les quizz « autant de fois que souhaité, [...] notamment lorsque l'on fait beaucoup d'erreurs ! Ou pour réviser avant l'examen... ».

« Les statistiques, ça vous fait penser à des choses qu'on n'imaginerait jamais autrement. »

Keith Ridgway

5. Discussion

5.1. Evaluation de l'outil

Le but de l'outil développé durant ce travail est d'entraîner la recherche d'informations pertinentes en vue de sélectionner le test statistique le plus adapté pour répondre à une question de recherche. Les données récoltées ont permis d'évaluer empiriquement le niveau de connaissances des participant·e·s en lien avec cette compétence. De plus, d'un point de vue qualitatif, il semble que les participant·e·s ont apprécié ce nouvel outil qu'ils / elles jugent pertinent dans l'apprentissage de cette compétence. Malgré le retour généralement positif par rapport à cet outil, trois aspects principaux pourraient être améliorés, à savoir la recherche des variables pertinentes, la possibilité de refaire chaque quizz avec de nouvelles situations et le retour proposé.

Généralement, le nombre de variables d'une base de données récoltées lors d'une étude dépasse celui du nombre de variables en lien avec une question de recherche. La première étape pour sélectionner le test le plus approprié consiste donc à trouver les variables pertinentes pour répondre à la question de recherche. Seul le quizz « type de test » permet de tester indirectement cette première étape. En effet, dans ce quizz, les participant·e·s n'ont aucun indice par rapport à quelles variables doivent être utilisées. Dans les autres quizz, les variables sont explicitement nommées dans la question, ce qui implique que les participant·e·s n'ont pas besoin de les trouver par eux / elles-mêmes. Cela leur donne aussi un indice sur le type d'analyse à utiliser dans le quizz « Etude de cas ». En effet, si trois variables doivent être traitées, le nombre d'analyses possibles se réduit à trois (ANOVA inter-sujets, à mesures répétées ou mixte) devrait être utilisée. Il serait intéressant de développer une série d'exercices pour cette première étape où les participant·e·s devraient sélectionner les variables directement dans le texte ou parmi un certain nombre d'alternatives (QCM) ou les indiquer dans un champ de texte. Cela permettrait aussi de vérifier que toutes les variables pertinentes ont été identifiées, par exemple les deux variables indépendantes et la variable dépendante dans le cas où une ANOVA à deux facteurs devrait être utilisée.

Comme déjà mentionné plus haut (cf. point 4.3 *Analyses des commentaires et suggestions*), bien que mentionné dans la première section du cours Moodle (cf. Figure 4), les participant·e·s

ne semblent pas avoir compris qu'ils / elles avaient la possibilité de refaire le même quizz plusieurs fois. Rajouter une page à la fin de chaque quizz mentionnant cette possibilité aurait probablement augmenté le nombre de tentatives par personne à chaque quizz.

Enfin, le retour aux participant·e·s pourrait être améliorée. En effet, bien que généralement bonne, la qualité des différents retours varie selon les exercices (cf. point 4.3 *Analyses des commentaires et suggestions*). La relecture des retours par une seconde personne (ou un groupe) enseignant la statistique aurait permis d'en améliorer la qualité. De plus, le retour proposé fournit constamment la réponse correcte car les quizz sont programmés pour ne permettre qu'une seule tentative. Cette implémentation est pertinente dans le cas des quizz « Nature de la variable » et « Type d'échelle » car quelle que soit la réponse donnée, la réponse correcte permet à l'étudiant·e de comprendre son erreur. En revanche, le retour pour la sélection du test approprié dans les quizz « Type de tests » et « Etudes de cas » pourrait être plus formatif lors d'une réponse incorrecte. A la place de fournir directement la réponse correcte, le retour pourrait réorienter la réflexion de la personne en fonction de sa réponse. Par exemple, si elle a choisi le t-test à mesure indépendantes alors qu'une ANOVA inter-sujets est attendue, le retour pourrait être formulé de la manière suivante : « Le t-test pour échantillons indépendants permet de tester la différence entre deux groupes seulement (VI à deux modalités). Dans la situation présence, la VI a plus de deux modalités ». Dans ce cas, il serait pertinent de laisser à la personne autant de tentatives que nécessaire pour s'assurer qu'elle trouve la bonne réponse.

5.2. Expérience d'apprenant

Durant la première année de psychologie, je n'avais de loin pas tout compris en profondeur le cours de statistique. En effet, j'avais principalement acquis un savoir-faire divergent, à savoir quel test utilisé pour quels types de données dans quelles circonstances, le cœur de ce projet. En revanche, la théorie sous-jacente, notamment la théorie de l'échantillonnage, m'était très floue. Je savais ce que je devais faire mais sans être capable d'en donner la raison.

Deux années plus tard, je me retrouve tuteur pour les exercices liés à cette unité d'enseignement et mon supérieur me demande de réexpliquer cette théorie lors du séminaire idoine. C'est au moment où j'étais dans l'obligation de l'expliquer que j'ai finalement compris cette théorie (à force de lecture et de travail sur les notes de cours). Selon la taxonomie SOLO de Biggs et Collis (1982), je pense que j'étais en première année dans une phase quantitative multistructurale car j'étais capable d'utiliser les algorithmes de cette théorie vus en cours mais sans comprendre leurs liens entre eux et la théorie de tests en général. Par contre, en tant que

tuteur, je crois avoir réussi à passer dans la phase qualitative, au minimum relationnelle et au mieux abstraite étendue. Actuellement, donnant le cours de statistique de première année de psychologie en allemand, je pense avoir développé ma compréhension du sujet de manière plus profonde et espère être en grande partie dans la phase abstraite étendue. En effet, j'ai l'impression d'arriver à faire plus de liens (spontanés) entre les différentes théories et analyses et à généraliser les concepts. Cependant, j'ai l'impression que très peu d'étudiant·e·s arrivent (ou sont intéressé·e·s) à comprendre certains de ces liens lorsque je les montre en cours après seulement un ou deux semestres d'étude de cette matière.

5.3. Concepts et outils pertinents utilisés dans ce travail

Les connaissances acquises (ou rafraichies) au travers des différents axes de la formation en didactique m'ont servi tout au long de ce travail. L'axe A « Enseignement et apprentissage » de la formation en didactique m'a apporté de nombreuses connaissances sur l'apprentissage aux travers des différents modèles présentés (par ex., SOLO) et des outils efficaces pour élaborer des concepts pédagogiques adaptés aux étudiant·e·s. Cela m'a permis de mieux comprendre où se situent les problèmes rencontrés par les étudiant·e·s (i.e. la difficulté à combiner les informations pertinentes sans soutien) et comment les aider au mieux (i.e. « driller » la recherche des informations pertinentes). L'unité d'enseignement A6 « Usage des technologies pour l'enseignement et les apprentissages m'a permis de découvrir les nombreuses possibilités qu'offre Moodle et que je ne connaissais pas au moment de la réalisation de ce projet. De plus, cet enseignement m'a permis de créer relativement rapidement mes examens en ligne durant la période COVID.

L'axe B « Evaluation des apprentissages et des dispositifs » a été un plaisant rafraîchissement des connaissances de statistique, de psychométrie et de biais cognitifs acquises lors de mes études de licence en psychologie. Il m'a particulièrement été pratique dans la cadre de la rédaction des retours lorsque les étudiant·e·s ne répondaient pas correctement à une question. De plus, ce projet aurait pu s'intégrer dans un scénario pédagogique hybride de l'unité d'enseignement « Statistique et analyse de données II ». Une telle intégration aurait peut-être permis d'augmenter le nombre de situations expérimentales. En combinaison avec l'unité d'enseignement A6, l'unité d'enseignement B5 « Outils d'évaluation et de suivi des apprentissages on-line » m'a servi de bac à sable pour développer le présent projet et m'a permis de choisir le support logiciel le plus adapté pour le réaliser.

L'axe C « Développement professionnel et gestion du travail académique » m'a permis de renforcer mes compétences en littérature grâce à l'unité d'enseignement C8 « Recherche et gestion de l'information », ce qui m'a aidé à trouver la littérature qui a servi de base à ce projet. Les autres unités d'enseignement n'ont malheureusement pas pu être mises en œuvre dans le cadre de ce projet. Cependant, le module de base m'a cependant permis de réfléchir aux compétences transversales qui sont principalement acquises en début de carrière académique (telles que la gestion de projet, la supervision d'étudiant·e·s) qui ne sont pas forcément connues du monde non-académique. Il m'a aussi permis de mieux analyser certaines situations au niveau départemental et d'adapter mon comportement en fonction de la situation et des enjeux, surtout depuis que j'y travaille en tant que lecteur et conseiller aux études master. Enfin, en tant que conseiller aux études Master, les unités d'enseignement C4 « Gestion des conflits » et C5 « L'art d'argumenter » me sont couramment utiles lors de mes diverses interactions avec les différents corps de l'université, surtout les étudiant·e·s. En effet, le rôle de conseiller aux études nécessite parfois des compétences en gestion de conflit.

5.4. Mes apprentissages durant ce travail

Ce projet m'a apporté de nombreuses connaissances et m'a fait réfléchir sur trois niveaux principaux, celui de la matière en elle-même et de son impact sur les étudiant·e·s, celui de la conceptualisation des connaissances et leur développement, et enfin sur l'investissement nécessaire à un enseignement de qualité.

Tout d'abord, ce projet m'a permis de réaliser que l'enseignement de la statistique, matière qui n'a pas forcément bonne presse auprès des étudiant·e·s (e.g., Berk & Nanda, 1998; Ruggeri, Dempster, et al., 2008), tient une place importante au sein de la didactique universitaire. En effet, quatre périodes scientifiques y sont spécifiquement consacrés et il existe de nombreux modèles décrivant les attitudes des étudiant·e·s envers cette matière, dont le modèle SATS (Ramirez et al., 2012).

Ensuite, ce projet m'a forcé à prendre du recul par rapport au développement de mes connaissances en statistique et d'en évaluer les différents stades et de repérer les moments-clés qui sont à l'origine d'un passage d'un stade à un autre. Cette réflexion n'a été possible qu'à l'aide du modèle le plus adapté à ce genre de matière, à savoir le modèle SOLO (Biggs & Collis, 1982; Biggs & Tang, 2011). En plus des observations de terrain réalisées par mes collègues, cette prise en recul m'a aussi fait comprendre l'importance de connaître le niveau des

connaissances des étudiant·e·s pour pouvoir leur proposer des activités les soutenant dans la construction de leurs propres connaissances.

Enfin, bien qu'il ne pourrait représenter qu'une petite partie du développement d'une unité d'enseignement, ce projet a nécessité beaucoup de ressources pour être mené à bien et être, je l'espère, d'une qualité pédagogique suffisante. En effet, la formation en didactique m'a montré la complexité de l'élaboration d'une unité d'enseignement de haute qualité et l'investissement nécessaire pour y arriver. En plus de la construction du scénario pédagogique (esquissé durant certaines unités d'enseignement, puis finalisé lors de la réalisation de ce projet), le travail de recherche de littérature nécessaire pour générer des exercices pertinents basés sur des recherches publiées (qui sont, je l'espère, intéressants et motivants) a été beaucoup plus ardu que prévu. Il en a résulté le nombre restreint de situations expérimentales utilisables. En outre, bien qu'ayant certaines notions de programmation et ayant suivi l'unité d'enseignement A6, la programmation correcte des quizz a requis beaucoup de temps (Moodle n'étant pas toujours très ergonomique). Ce temps passé entre frustrations et satisfaction d'avoir réussi m'a grandement aidé dans l'utilisation plus poussée de Moodle. Cela m'a permis d'employer cette plateforme autrement que pour mettre des documents à disposition des étudiant·e·s.

“La pensée statistique sera un jour une qualification aussi nécessaire pour une citoyenneté efficace que la capacité de lire et d’écrire.”

H. G. Wells

6. Conclusion et perspectives futures

Ce projet est parti du constat que les étudiant·e·s en psychologie de première et deuxième année rencontrent des difficultés lorsqu’il faut choisir le test le plus approprié en fonction de la situation expérimentale à laquelle ils / elles sont confronté·e·s, que ce soit lors d’un examen de statistique, durant les travaux pratiques expérimentaux ou leur travail de Bachelor empirique. Le but de ce travail est de comprendre où se situe le problème et comment le surmonter. Pour ce faire, quatre quizz ont été développés. Les deux premiers quizz évaluent chacun une connaissance nécessaire (niveau de l’échelle de mesure et nature de la variable) pour effectuer le choix le plus pertinent en fonction d’une situation expérimentale spécifique. Les deux derniers testent la capacité des étudiant·e·s à choisir le test le plus appropriée, l’un forçant les étudiant·e·s à prendre en considération les deux connaissances testées par les deux premiers quizz (quizz avec soutien), l’autre non (quizz sans soutien). Les participant·e·s aux quizz ont pu évaluer la qualité et pertinence des quizz à l’aide d’un questionnaire en ligne.

L’analyse des données a révélé que les participant·e·s sont en possession des connaissances élémentaires, mais ne semblent pas arriver pas à les combiner de façon spontanée pour atteindre la compétence visée. Cependant, lorsqu’ils / elles sont poussé·e·s à prendre en considération ces connaissances, ils / elles sont capables de faire le choix le plus pertinent. Ces résultats suggèrent donc que les participant·e·s auraient besoin d’acquérir une routine incluant l’analyse du niveau de l’échelle et de la nature des variables. Cela pourrait être accompli à l’aide de l’élargissement de la base de données des situations expérimentales et en « drillant » la routine durant les différentes unités d’enseignement de statistique. Concernant le retour par rapport aux quizz, il est à souligner l’intérêt de tels exercices en ligne, mais le besoin de plus de situations expérimentales. Enfin, il est important de relever la relative faible participation aux quizz, qui semble confirmer un manque de motivation pour l’apprentissage de cette matière.

Pour favoriser l’utilisation des quizz, il faudrait en accroître la pertinence auprès des étudiant·e·s. Il serait judicieux de montrer (si possible de prouver) que s’entraîner à l’aide de ces quizz augmente les résultats à l’examen. Réussir les quizz renforcerait le sentiment d’efficacité personnelle (Bandura, 1997), améliorant à son tour l’attitude envers la statistique

(e.g. Ramirez et al., 2012) et conduisant potentiellement à un meilleur résultat lors de l'évaluation (Usher & Pajares, 2008). Pour rendre ce lien plus présent auprès des étudiant·e·s, il serait intéressant de le mentionner tout au long de l'année académique et plus particulièrement avant les sessions d'examen. Il serait aussi judicieux de conseiller les quizz aux personnes ayant échoué lors de leur première tentative à l'examen.

Pour tester l'impact réel de la mise en pratique de la routine à l'aide de ces quizz, il serait intéressant de proposer un plan expérimental en trois phases, (1) ligne de base, (2) intervention et (3) évaluation de l'impact de l'intervention, dans le cadre de l'unité d'enseignement « Statistique et analyses de données II ». Dans un premier temps, les étudiant·e·s participent à un quizz sans soutien sur le choix du test approprié et reçoivent uniquement comme retour le nombre de réponses correctes. Cette ligne de base permettrait de déterminer le niveau de performance de chaque participant·e. Dans un deuxième temps, les participant·e·s peuvent faire une autoévaluation formative de leurs connaissances à l'aide du quizz avec soutien. Cette intervention leur permettrait de connaître leur niveau de performance et leur fournit un retour en fonction de leurs erreurs. De plus, la possibilité de refaire plusieurs fois le quizz leur permettrait de mesurer la progression de leurs compétences et l'impact de l'entraînement sur leurs performances. Enfin, pour évaluer la pertinence de l'intervention, les étudiant·e·s participent à un dernier quizz sans soutien avec retour sur leurs performances. Les situations expérimentales devraient être différentes à chaque phase.

Si le quizz avec soutien permet effectivement de « driller » de façon adéquate la routine d'analyse, il est attendu que, de manière générale, les performances s'améliorent significativement entre la ligne de base et l'évaluation de l'intervention. De plus, l'amélioration des performances devrait être plus importante pour les personnes ayant un niveau de connaissances bas que pour celles ayant un niveau de connaissances élevé avant l'intervention. Enfin, plus le nombre de répétitions du quizz avec soutien dans la phase d'intervention est important, plus la routine d'analyse devrait être intégrée, plus les performances en phase d'évaluation de l'intervention devraient être élevées.

References

- American Statistical Association. (2016). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report*.
http://www.amstat.org/asa/files/pdfs/GAISE/GaiseCollege_Full.pdf
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of educational objectives*. Longman.
- Bandura, A. (1997). *Self-efficacy: The exercise of control* (1st ed.). W.H. Freeman.
- Barceló, F., & Knight, R. T. (2002). Both random and perseverative errors underlie WCST deficits in prefrontal patients. *Neuropsychologia*, 40(3), 349–356. [https://doi.org/10.1016/S0028-3932\(01\)00110-5](https://doi.org/10.1016/S0028-3932(01)00110-5)
- Ben-Zvi, D., & Garfield, J. B. (2004). *The challenge of developing statistical literacy*. Springer Science + Business Media, Inc; Springer Netherlands; Imprint; Springer.
- Ben-Zvi, D., & Garfield, J. B. (2008). Introducing the emerging discipline of statistics education. *School Science and Mathematics*, 108(8), 355–361. <https://doi.org/10.1111/j.1949-8594.2008.tb17850.x>
- Berk, R. A., & Nanda, J. P. (1998). Effects of jocular instructional methods on attitudes, anxiety, and achievement in statistics courses. *Humor - International Journal of Humor Research*, 11(4). <https://doi.org/10.1515/humr.1998.11.4.383>
- Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy (structured of the observed learning outcome)*. Educational psychology. Academic Press.
- Biggs, J. B., & Tang, C. S. (2011). *Teaching for quality learning at university: What the student does* (4th Edition). SRHE and Open University Press Imprint. Open University Press.
<https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=405333>
- Birenbaum, M., & Eylath, S. (1994). Who is afraid of statistics? Correlates of statistics anxiety among students of educational sciences. *Educational Research*, 36(1), 93–98.
<https://doi.org/10.1080/0013188940360110>
- Carillo, K., Galy, N., Guthrie Cameron, & Vanhems, A. (2016). "J'aime pas les stats !" - Mesure et analyse de l'attitude à l'égard du cours de statistique dans une école de management. *Statistique Et Enseignement*, 7(1), 3–31.
- Carlson, K. A., & Winquist, J. R. (2011). Evaluating an active learning approach to teaching introductory statistics: A classroom workbook Approach. *Journal of Statistics Education*, 19(1). <https://doi.org/10.1080/10691898.2011.11889596>
- Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly*, 104(9), 801. <https://doi.org/10.2307/2975286>
- Coetzee, S., & van der Merwe, P. (2010). Industrial psychology students' attitudes towards statistics. *SA Journal of Industrial Psychology*, 36, 1–8.
http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S2071-07632010000100009&nrm=iso
- Dancey, C. P., Reidy, J., Gauvrit, N., & Baggio, S. (2016). *Statistiques sans maths pour psychologues*. De Boek Superior.
- Dempster, M., & McCorry, N. K. (2009). The role of previous experience and attitudes toward statistics in statistics assessment outcomes among undergraduate psychology students. *Journal of Statistics Education*, 17(2). <https://doi.org/10.1080/10691898.2009.11889515>
- Deschryver, N., & Charlier, B. (2012). *Dispositifs hybrides: Nouvelles perspectives pour une pédagogie renouvelée de l'enseignement supérieur*. Rapport final. <http://prac-hysup.univ-lyon1.fr/spiral-files/download?mode=inline&data=1757974>

- Emmioğlu, E., & Capa-Aydin, Y. (2021). ATTITUDES AND ACHIEVEMENT IN STATISTICS: A META-ANALYSIS STUDY. *Statistics Education Research Journal*, 11(2), 95–102. <https://doi.org/10.52041/serj.v11i2.332>
- Field, A. P., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage Publications.
- Garfield, J. B., Ben-Zvi, D., Chance, B., Medina, E., Roseth, C., & Zieffler, A. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Springer Science+Business Media B.V.
- Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *The Academy of Management Review*, 17(2), 183–211. <https://doi.org/10.5465/amr.1992.4279530>
- Gordon, S. (2004). Understanding students' experiences of statistics in a service course. *Statistics Education Research Journal*, 3(1), 40–59. [https://iase-web.org/documents/SERJ/SERJ3\(1\)_gordon.pdf](https://iase-web.org/documents/SERJ/SERJ3(1)_gordon.pdf)
- Gotlib, I. H., Krasnoperova, E., Yue, D. N., & Joormann, J. (2004). Attentional biases for negative interpersonal stimuli in clinical depression. *Journal of Abnormal Psychology*, 113(1), 121–135. <https://doi.org/10.1037/0021-843X.113.1.121>
- Kerby, A. T., & Wroughton, J. R. (2021). When do students' attitudes change? Investigating student attitudes at midterm. *Statistics Education Research Journal*, 16(2), 476–486. <https://doi.org/10.52041/serj.v16i2.202>
- Loftus, E. (1977). Shifting human color memory. *Memory & Cognition*, 5(6), 696–699.
- Medalia, A., Aluma, M., Tryon, W., & Merriam, A. E. (1998). Effectiveness of attention training in schizophrenia. *Schizophrenia Bulletin*, 24(1), 147–152. <https://doi.org/10.1093/oxfordjournals.schbul.a033306>
- Milic, N. M., Masic, S., Milin-Lazovic, J., Trajkovic, G., Bukumiric, Z., Savic, M., Milic, N. V., Cirkovic, A., Gajic, M., Kostic, M., Ilic, A., & Stanisavljevic, D. (2016). The importance of medical students' attitudes regarding cognitive competence for teaching applied statistics: Multi-site study and meta-analysis. *PloS One*, 11(10), e0164439. <https://doi.org/10.1371/journal.pone.0164439>
- Murtonen, M., & Lehtinen, E. (2003). Difficulties experienced by education and sociology students in quantitative methods courses. *Studies in Higher Education*, 28(2), 171–185. <https://doi.org/10.1080/0307507032000058064>
- Nosarti, C., Roberts, J. V., Crayford, T., McKenzie, K., & David, A. S. (2002). Early psychological adjustment in breast cancer patients: A prospective study. *Journal of Psychosomatic Research*, 53(6), 1123–1130. [https://doi.org/10.1016/S0022-3999\(02\)00350-1](https://doi.org/10.1016/S0022-3999(02)00350-1)
- Onwuegbuzie, A. J. (1997). Writing a research proposal: The role of library anxiety, statistics anxiety, and composition anxiety. *Library & Information Science Research*, 19(1), 5–33. [https://doi.org/10.1016/S0740-8188\(97\)90003-7](https://doi.org/10.1016/S0740-8188(97)90003-7)
- Onwuegbuzie, A. J., & Wilson, V. A. (2003). Statistics anxiety: Nature, etiology, antecedents, effects, and treatments--a comprehensive review of the literature. *Teaching in Higher Education*, 8(2), 195–209. <https://doi.org/10.1080/1356251032000052447>
- Plaisted, K., O'Riordan, M., & Baron-Cohen, S. (1998). Enhanced visual search for a conjunctive target in autism: A research note. *Journal of Child Psychology and Psychiatry*, 39(5), 777–783. <https://doi.org/10.1017/S0021963098002613>
- Ramirez, C., Schau, C., & Emmioğlu, E. (2012). The importance of attitudes in statistics education. *Statistics Education Research Journal*, 11(2), 57–71. <https://doi.org/10.52041/serj.v11i2.329>
- Rejón-Guardia, F., Vich-I-Martorell, G. A., Juaneda, C., & Claredas, M. (2019). Gender differences in attitudes towards statistics in social science degrees. In *11th International Conference on Education and New Learning Technologies*, Palma, Spain.

- Ruggeri, K., Dempster, M., Hanna, D., & Cleary, C. (2008). Experiences and expectations: The real reason nobody likes stats. *Psychology Teaching Review, 14*(2), 75–83. <https://doi.org/10.53841/bpsptr.2008.14.2.75>
- Ruggeri, K., Díaz, C., Kelley, K., Papousek, I., Dempster, M., & Hanna, D. (2008). International issues in education. *Psychology Teaching Review, 14*(2), 65–74. <https://doi.org/10.53841/bpsptr.2008.14.2.65>
- Sato, S., Gygax, P. M., & Gabriel, U. T. (2013). Gender inferences: Grammatical features and their impact on the representation of gender in bilinguals. *Bilingualism: Language and Cognition, 16*(4), 792–807. <https://doi.org/10.1017/S1366728912000739>
- Schau, C. (2003). Students' attitudes: The "other" important outcome in statistics education. In *Proceedings of the joint statistical meetings*, San Francisco, CA.
- Schau, C., & Emmioğlu, E. (2011). Changes in US students' attitudes toward statistics across introductory statistics courses. In *Proceedings of the 58th World Statistics Congress*.
- Schau, C., Stevens, J., Dauphinee, T. L., & Del Vecchio, A. (1995). The development and validation of the survey of attitudes toward statistics. *Educational and Psychological Measurement, 55*(5), 868–875. <https://doi.org/10.1177/0013164495055005022>
- Schunk, D. H., Meece, J., & Pintrich, P. (2014). *Motivation in education : Theory, research, and applications* (4th ed.). Pearson.
- Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78*(4), 751–796. <https://doi.org/10.3102/0034654308321456>
- van Es, C., & Weaver, M. M. (2018). Race, Sex, and their Influences on Introductory Statistics Education. *Journal of Statistics Education, 26*(1), 48–54. <https://doi.org/10.1080/10691898.2018.1434426>
- Verhoeven, P. S. (2011). Attitudes toward statistics: The effort of learning. In *58th World Statistics Congress of the International Statistical Institute*.
- Wagner, K. D., Robb, A. S., Findling, R. L., Jin, J., Gutierrez, M. M., & Heydorn, W. E. (2004). A randomized, placebo-controlled trial of citalopram for the treatment of major depression in children and adolescents. *The American Journal of Psychiatry, 161*(6), 1079–1083. <https://doi.org/10.1176/appi.ajp.161.6.1079>
- Waters, L. K., Martelli, T. A., Zakrajsek, T., & Popovich, P. M. (1988). Attitudes toward Statistics: An Evaluation of Multiple Measures. *Educational and Psychological Measurement, 48*(2), 513–516. <https://doi.org/10.1177/0013164488482026>
- Zanakis, S. H., & Valenzi, E. R. (1997). Student anxiety and attitudes in business statistics. *Journal of Education for Business, 73*(1), 10–16. <https://doi.org/10.1080/08832329709601608>

Annexes

Annexe A

Exercice 1

Gotlib et al. (2004) souhaitent observer si les personnes souffrant d'un épisode dépressif majeur présentent des biais attentionnels plus importants au niveau du traitement de stimuli interpersonnels négatifs que les personnes souffrant d'anxiété généralisée et des personnes contrôle. Avant de faire passer leur expérience, ils veulent vérifier que les personnes souffrant d'un épisode dépressif majeur qu'ils ont recrutées ont un niveau de dépression plus élevées que celles des autres groupes. Pour ce faire, ils leur font passer le BDI (Beck Depression Index, scores de 0 à 63).

1) Question échelle

- La variable *type de participant-e-s* est-elle : nominale, ordinale, quantitative
- La variable *dépression* (mesurée par le score au BDI) est-elle : nominale, ordinale, quantitative ?

2) Question (in)dépendance

- La variable *type de participant-e-s* est-elle une variable dépendante ou indépendante ?
- La variable *dépression* (mesurée par le score au BDI) est-elle une variable dépendante ou indépendante ?

3) Question design

Quelle analyse Gotlib et al. doivent-ils utiliser pour vérifier leur hypothèse ?

Exercice 2

Sato et al. (2013) souhaitent observer l'influence des règles d'orthographe sur la représentation du genre d'un groupe de personnes exerçant un métier stéréotypé. Un exemple pourrait être que les chirurgiens sont tous des hommes et le personnel soignant est exclusivement composé de femmes. Contrairement à l'anglais, le français est une langue employant des genres. Une des règles orthographiques veut qu'un groupe de personnes mixtes soit représenté par le masculin pluriel (par ex, les clients), même si la majorité des personnes du groupe sont des femmes. Si les règles orthographiques liées au genre influencent la représentation du genre, les personnes parlant le français devraient avoir une représentation plus masculine d'un groupe mixte, quel que soit le métier de ces personnes. Par contre, les personnes parlant l'anglais ne devraient être sensibles qu'au stéréotype du métier, étant donné qu'elles ne sont pas biaisées par l'orthographe. Pour tester cette hypothèse, les auteurs ont fait lire de courtes histoires de quelques lignes à 61 francophones et 66 anglophones. Chacune de ces histoires ont comme sujets un groupe de personnes dont le métier est stéréotypé soit masculin (les chirurgiens), soit féminin (les infirmières), soit n'est pas stéréotypé. La dernière phrase de chaque histoire donnait le genre du groupe (les hommes) et les participant-e-s devaient indiquer si cette dernière phrase pouvait représenter une continuation logique de l'histoire. Les auteurs ont mesuré le degré d'approbation en calculant le pourcentage de oui.

1) Question échelle

- La variable *langue des participant-e-s* est-elle : nominale, ordinale, quantitative
- La variable *pourcentage de oui* est-elle : nominale, ordinale, quantitative ?
- La variable *stéréotype du métier* est-elle : nominale, ordinale, quantitative ?

2) Question (in)dépendance

- La variable *pourcentage de oui* est-elle une variable dépendante ou indépendante ?
- La variable *langue des participant-e-s* est-elle une variable dépendante ou indépendante ?
- La variable est-elle *stéréotype du métier* une variable dépendante ou indépendante ?

3) Question design

Quelle analyse les auteurs doivent-ils utiliser pour vérifier leur hypothèse ?

Exercice 3

Plaisted et al. (1998) veulent vérifier si les performances de personnes souffrant d'autisme dans une tâche de détection difficile sont meilleures que celles de personnes contrôle. Ils font passer deux types de tâches de recherche visuelles (une facile et une difficile) à 8 personnes souffrant d'autisme et 8 personnes contrôle. Dans chacune des tâches, le/la participant·e doit indiquer si l'image qui est présenté contient ou non une cible. Dans la tâche de détection simple, la cible est un S parmi des T et des X, alors que dans la tâche de difficile, il faut trouver un X rouge parmi des X verts et des T rouges. Les hypothèses font l'hypothèse que les personnes souffrant d'autisme auront un temps de réaction plus rapide que les personnes contrôle, mais uniquement pour la tâche de détection difficile.

1) Question échelle

- La variable *groupe de participant-e-s* est-elle : nominale, ordinale, quantitative
- La variable *temps de réaction* est-elle : nominale, ordinale, quantitative ?
- La variable *type de tâche* est-elle : nominale, ordinale, quantitative ?

2) Question (in)dépendance

- La variable *type de tâche* est-elle une variable dépendante ou indépendante ?
- La variable *groupe de participant-e-s* est-elle une variable dépendante ou indépendante ?
- La variable est-elle *temps de réaction* une variable dépendante ou indépendante ?

3) Question design"

Quelle analyse les auteurs doivent-ils utiliser pour vérifier leur hypothèse ?

Exercice 4

Wagner et al. (2004) souhaitent observer si l'effet placebo peut modifier la perception de la douleur. Un groupe de participant-e-s reçoivent des chocs électriques douloureux au niveau de leur poignet droit et doivent évaluer leur douleur sur une échelle allant de 1 de 10. Dans un premier bloc d'essais, ils reçoivent les chocs sans traitement. Au début du second bloc, un expérimentateur leur applique au poignet une nouvelle crème analgésique qui est supposée réduire la douleur (placebo). Les auteurs font l'hypothèse que l'application de la crème placebo va réduire la douleur perçue par les participant-e-s.

1) Question échelle

- La variable *bloc d'essais* est-elle : nominale, ordinale, quantitative
- La variable *douleur perçue* est-elle : nominale, ordinale, quantitative ?

2) Question (in)dépendance

- La variable *bloc d'essais* est-elle une variable dépendante ou indépendante ?
- La variable *douleur perçue* est-elle une variable dépendante ou indépendante ?

3) Question design

Quelle analyse les auteurs doivent-ils utiliser pour vérifier leur hypothèse ?

Exercice 5

Un nombre important de femmes souffrant du cancer du sein font simultanément l'expérience de problèmes psychiatriques. Nosarti et al. (2002) ont voulu observer si ces femmes pouvaient s'adapter rapidement à leur nouvelle condition. Pour ce faire, ils ont fait passer le MAC (questionnaire d'ajustement mental au cancer – Mental Adjustment to Cancer) à 87 patientes âgées entre 40 et 70 ans lors d'un suivi effectué 8 semaines (t1) puis 9 mois (t2) après la détection du cancer. Ce questionnaire comporte cinq dimensions (esprit de combat, désespoir, préoccupation anxieuse, stoïcisme-fatalisme, et évitement-déni) de 8 items à coter sur une échelle allant de 1 (ne s'applique pas du tout à moi) à 4 (s'applique tout à fait à moi). Un score total (addition des scores aux items) est calculé pour chaque dimension. L'analyse descriptive des données révèlent que les variables ne sont pas distribuées normalement. L'analyse inférencielle montre que les patientes présentent un score au plus petit de préoccupation anxieuse en t2 qu'en t1.

1) Question échelle

- La variable *préoccupation anxieuse* est-elle : nominale, ordinale, quantitative
- La variable *suivi* est-elle : nominale, ordinale, quantitative ?

2) Question (in)dépendance

- La variable *préoccupation anxieuse* est-elle une variable dépendante ou indépendante ?
- La variable *suivi* est-elle une variable dépendante ou indépendante ?

3) Question design

Quelle analyse les auteurs doivent-ils utiliser pour vérifier leur hypothèse ?

Exercice 6

Medalia et al. (1998) souhaitent mesurer l'efficacité de l'entraînement de l'attention sur la schizophrénie. Deux groupes de 27 personnes souffrant de schizophrénie prennent part à cette expérience. Le groupe expérimental a participé à des séances individuelles d'entraînement attentionnel sur ordinateur pendant que le groupe contrôle regardent des documentaires vidéo. Puis, pour mesurer leurs capacités attentionnelles, les participant-e-s passaient une tâche de détection à choix forcé d'une durée de 10 minutes. Cette dernière était composée de 600 essais. A chacun d'eux, ils/elles devaient indiquer si la lettre qui était présentée à l'écran étaient un A. L'analyse descriptive des données révèlent que les variables ne sont pas distribuées normalement. L'analyse inférencielle montre que le groupe expérimental obtient un pourcentage de réponses correctes que le groupe contrôle.

1) Question échelle

- La variable *groupes* est-elle : nominale, ordinale, quantitative
- La variable *pourcentage de réponses correctes* est-elle : nominale, ordinale, quantitative ?

2) Question (in)dépendance

- La variable *pourcentage de réponses correctes* est-elle une variable dépendante ou indépendante ?
- La variable *groupes* est-elle une variable dépendante ou indépendante ?

3) Question design

Quelle analyse les auteurs doivent-ils utiliser pour vérifier leur hypothèse ?

Exercice 7

Barcelo et Knight (2002) font passer le Wisconsin card sorting test à un groupe de 8 patients préfrontaux et à un groupe contrôle de 6 personnes sans problèmes neurologiques. Dans cette tâche, quatre cartes sont présentées à la personne et vont lui servir de stimuli de base. Les cartes diffèrent en termes de type, du nombre et de la couleur des formes qui y sont dessinées. On demande ensuite à la personne d'associer une à une les 128 cartes avec l'une des quatre cartes de bases. L'examineur ou l'expérimentatrice décide de la règle d'association (classement par couleur, forme ou quantité) et indique uniquement à la personne si son choix est juste ou non. Cette tâche permet de mesurer la flexibilité cognitive, entre autres au travers du nombre d'erreurs de persévérance (i.e., la personne continue à appliquer la même stratégie de classement alors que cette dernière a changé) L'expérimentateur ou l'expérimentatrice change la règle (sans l'annoncer) après un certain nombre de réussites consécutives. Les auteurs ont observé que les patients préfrontaux faisaient plus d'erreurs de persévérance que les personnes sans problèmes neurologiques.

1) Question échelle

- La variable *groupes* est-elle : nominale, ordinale, quantitative
- La variable *nombre d'erreurs de persévérance* est-elle : nominale, ordinale, quantitative ?

2) Question (in)dépendance

- La variable *nombre d'erreurs de persévérance* est-elle une variable dépendante ou indépendante ?
- La variable *groupes* est-elle une variable dépendante ou indépendante ?

3) Question design

Quelle analyse les auteurs doivent-ils utiliser pour vérifier leur hypothèse ?

Exercice 8

En 1977, Loftus a voulu tester la sensibilité de la mémoire à de fausses informations chez les témoins oculaires. Pour ce faire, elle demanda à 95 étudiant-e-s de regarder une série de 30 diapositives (présentées 3 s chacune) illustrant un accident entre une voiture et un piéton. Une Datsun rouge roule sur une route secondaire, arrive à un carrefour, tourne à droite et renverse une personne qui traversait sur le passage piéton. Une voiture verte passe le carrefour sans s'arrêter (visible sur une seule diapositive). Les participant-e-s ont ensuite rempli un questionnaire de 12 items sur la présentation. L'induction d'un faux souvenir se faisait au travers de l'item 10 qui demandait si la voiture qui a traversé le carrefour sans s'arrêter avec un porte-skis. Dans le groupe expérimental (48 personnes), la question précisait que cette voiture était de couleur bleue alors que dans le groupe contrôle, aucune information sur la couleur de la voiture n'était fournie. Puis les participant-e-s devaient indiquer la couleur de différents objets vu durant la présentation à l'aide de 30 bandes de couleurs classées selon leur longueur d'ondes. Les bandes 1 à 5 correspondaient à des nuances de pourpre, les bandes 6 à 10 au bleu, 11 à 15 au vert, 16 à 20 au jaune, 21 à 25 à l'orange et 26 à 30 au rouge. Les résultats ont montré que le groupe expérimental indiquait que la voiture qui est passée au carrefour sans s'arrêter était de couleur bleu-vert (10.19) alors que le groupe contrôle choisissait une teinte verte (12.55).

1) Question échelle

- La variable *groupes* est-elle : nominale, ordinale, quantitative
- La variable *couleur* est-elle : nominale, ordinale, quantitative ?

2) Question (in)dépendance

- La variable *nombre d'erreurs de persévérance* est-elle une variable dépendante ou indépendante ?
- La variable *couleur* est-elle une variable dépendante ou indépendante ?

3) Question design

Quelle analyse les auteurs doivent-ils utiliser pour vérifier leur hypothèse ?

Description et Lien vers chaque Quizz.

Type d'échelle

Cette partie vous permet d'exercer vos compétences par rapport au type d'échelle d'une variable. Ai-je plutôt affaire à une variable nominale, ordinale ou quantitative? Cette notion est cruciale car elle permet de savoir quels indices rapportés et quel test statistique employé. Cliquez sur le lien suivant pour commencer le quizz!



Quizz type d'échelle

Nature de la variable

Cette partie vous permet d'exercer vos compétences à déterminer quelle est la nature de la variable. Ai-je plutôt affaire à une variable dépendante ou indépendante? Cette notion est cruciale car elle permet de savoir quel test statistique employé.

Cliquez sur le lien suivant pour commencer le quizz!



Quizz nature de la variable

Type de test

Cette partie vous permet d'exercer vos compétences à déterminer quel test employer en fonction de la situation. Dois-je plutôt utiliser un test de Mann-Whitney ou une ANOVA à un facteur intersujets? Utiliser le bon test est nécessaire pour une interprétation correcte des données.

Cliquez sur le lien suivant pour commencer le quizz!



quizz type de test

Etude de cas

Pour lier ces trois notions, le quizz suivant vous propose de déterminer la catégorie et la nature des variables ainsi que le test y correspondant pour chacune des trois situations proposées.

Cliquez sur le lien suivant pour commencer le quizz!



Etude de cas