# COMMENT ÉVALUER UN CHATBOT ASSISTANT DE COURS UTILISÉ EN FORMATION À DISTANCE ?

## Élaboration d'un questionnaire et évaluation d'ADIDBot

**MÉMOIRE RÉALISÉ EN VUE DE L'OBTENTION DE LA MAÎTRISE UNIVERSITAIRE EN SCIENCES ET TECHNOLOGIES DE L'APPRENTISSAGE ET DE LA FORMATION**

**PAR**
Aurélie Arnoldi

**CO-DIRECTEURS DU MÉMOIRE**
Pr. Gaëlle MOLINARI
et
Yannick Stéphane NLEME ZE, doctorant et assistant de recherche et d'enseignement

**JURY**
Estelle PRIOR

**Genève, le 4 février 2022**

**Université de Genève**
**Faculté de Psychologie et des Sciences de l'éducation**

UNIVERSITÉ DE GENÈVE
FACULTÉ DE PSYCHOLOGIE ET DES SCIENCES DE L'ÉDUCATION


## RÉSUMÉ

L'objectif de ce mémoire est d'élaborer un outil d'évaluation des chatbots utilisés en formation à distance comme assistants de cours et supports d'aide à la régulation de l'engagement et la motivation des étudiants. L'outil sera destiné à évaluer l'expérience des étudiants avec le chabot dans le cadre du cours ADID et pourra à terme être mis en application dans d'autres contextes de formation.

Pour le réaliser, nous avons adopté la méthodologie proposée par O'Brien & Toms (2010). Premièrement, nous avons défini les fondements théoriques, c'est-à-dire, explicité les différentes dimensions à évaluer au regard de la littérature. Deuxièmement, nous avons élaboré un questionnaire qui permet d'évaluer le plus de dimensions possibles relatives à la qualité du chatbot. Enfin, ce dernier a été soumis à un pré-test cognitif.

Le questionnaire créé mesure 9 critères (acceptabilité, comportement, efficacité, émotions, expérience utilisateur, motivation, présence, utilisabilité, utilité et impressions générales) à l'aide de 35 items d'évaluation qui permettent de déterminer l'appréciation des utilisateurs.

L'outil réalisé a su mettre en évidence les critères bien évalués ou non et ces derniers coïncident avec les propos exprimés lors des entretiens d'utilisation d'ADIDBot par les utilisateurs. En revanche, pour la collecte de données plus qualitatives, le questionnaire est limité, notamment pour rendre compte du ressenti des utilisateurs car les réponses fournies par écrit ne sont pas complètes et peuvent être sujettes à plusieurs interprétations. Pour pallier cette problématique, il convient de coupler l'utilisation du questionnaire avec la passation d'entretiens.



**Mots-clés**
Évaluation - Agent conversationnel - Questionnaire - Éducation - Formation à distance
Evaluation - Chatbot - Survey - Education - Distance learning

UNIVERSITÉ DE GENÈVE
FACULTÉ DE PSYCHOLOGIE ET DES SCIENCES DE L'ÉDUCATION

# *Déclaration sur l'honneur*

*Je déclare que les conditions de réalisation de ce travail de mémoire respectent la charte d'éthique et de déontologie de l'Université de Genève. Je suis bien l'auteure de ce texte et atteste que toute affirmation qu'il contient et qui n'est pas le fruit de ma réflexion personnelle est attribuée à sa source ; tout passage recopié d'une autre source est en outre placé entre guillemets.*

Genève, le 20 janvier 2022

Aurélie Arnoldi

Signature :

**Remerciements**

Je tiens à remercier toutes les personnes qui ont contribué de près ou de loin à la réalisation de ce mémoire et avant tout, mon co-directeur, Yannick Stéphane NLEME ZE, pour le temps consacré à me faire avancer dans la rédaction, pour sa patience et ses réponses à mes innombrables questions, pour son aide et ses précieux conseils qui m'ont grandement stimulée et guidée durant mes recherches.

Je remercie aussi tous les professeurs du TECFA de l'Université de Genève, qui ont fourni du contenu de qualité et les outils nécessaires à la réussite du MALTT. Ce master a été pour moi très enrichissant et rempli de belles rencontres.

De manière générale, je remercie tous les intervenants, amis, collègues et toutes les personnes qui m'ont apporté leur soutien moral et intellectuel tout au long de mes études. Je pense particulièrement à Martina SALEMMA mon binôme de travail durant ces deux années de master ; à Simon JÖHR, avec qui je partage ce sujet de mémoire, pour leur soutien dans ce travail ainsi qu'à ma maman, ma relectrice pour avoir lu, relu, fait des remarques en donnant un regard extérieur.

J'adresse également mes sincères remerciements à mon employeur, Gerofinance-Dunand SA, la directrice des Ressources Humaines, Céline THOMAS, mon responsable, Ruben ZAPICO et mon collègue, Raphaël GRITTI, pour avoir rendu possible ma participation à cette formation.

Et enfin, je remercie mon amoureux et mon bébé pour leur soutien, leur présence et leurs encouragements.

## Table des matières


1 Introduction et problématique 7

2 Cadre théorique 13

2.1 Qu'est-ce qu'un chatbot ? 13

2.2 Les chatbots dans le domaine éducatif 13

2.3 Comment évaluer un chatbot ? 15

2.3.1 Les méthodes d'évaluation des chatbots 15

2.3.2 Les outils d'évaluation des chatbots 16

2.3.3 Choix de l'outil pour évaluer les chatbots assistants utilisés en formation à distance 18

2.4 Définition des critères et des items d'évaluation 18

2.4.1 Expérience utilisateur 19

2.4.2 Critères d'évaluation des chatbots éducatifs selon Hobert (2019) 21

2.4.2.1 Acceptation, adoption et acceptabilité 22

2.4.2.2 Réussite de l'apprentissage 24

2.4.2.3 Motivation 24

2.4.2.4 Autres effets bénéfiques sur les processus d'apprentissage : comportement et humanité 25

2.4.2.5 Utilisabilité et utilité 27

2.4.2.6 Justesse technique : efficacité et performances 29

2.4.2.7 Autres facteurs psychologiques : émotions, sentiment de présence et engagement 29

3 Objectif de la recherche 33

4 Méthodologie 34

4.1 Une méthodologie inspirée par O'Brien & Toms (2010) 34

4.1.1 1ère étape : Définition des critères d'évaluation du Chatbot et des items y afférents 35

4.1.2 2ème étape : Réduction et revue par les experts 48

4.1.3 3ème étape : Validation des items par les utilisateurs 48

4.1.4 4ème étape : Réduction et revue par les utilisateurs puis construction du questionnaire 50

4.1.5 Test du questionnaire en contexte réel pour l'évaluation d'ADIDBot 50

4.1.5.1 Collecte des données 50

4.1.5.2 Analyse des données 51

5 Résultats 54

5.1 Des feedbacks des experts 54

5.2 Des feedbacks des entretiens couplés au tri de cartes 57

5.3 De la version finale du questionnaire 59

5.4 De l'évaluation du questionnaire 62

5.5 Évaluation d'ADIDBot 64

6 Discussion 68

6.1 La construction de l'outil au regard de la littérature 68
6.2 La revue des critères et des items par l'expert VS par les utilisateurs 69
6.3 L'évaluation du chatbot à l'aide du questionnaire et les dimensions à améliorer 70
6.4 De l'évaluation d'ADIDBot 71
6.5 Pistes d'amélioration pour ADIDBot 72
6.6 Limites et points d'amélioration 73
7 Conclusion 75
8 Références 77
9 Annexes 86
Annexe 1 - Tableau récapitulatif des thèmes, méta critères, critères, définitions et exemples de questions associées au modèle d'évaluation proposé par Hu, Trigano & Crozat (2001) 86
Annexe 2 - Items de la version française de l'AttrakDiff selon Lallemand et al. (2015) 92
Annexe 3 - Version française de l'UEQ+ de Schrepp & Thomaschewski (2019) 93
Annexe 4 - Items du modèle UTAUT et leur traduction (Venkatesh et al., 2003) 99
Annexe 5 - Échelle utilisée avec les apprenants (Caron & Heutte, 2017) 101
Annexe 6 - Échelle de motivation en formation des adultes (EMFA) de Fenouillet, Heutte & Vallerand (2015) 102
Annexe 7 - Items originaux du SUS de Brooke (1996) et leur traduction française (Gronier & Baudet, 2021) 104
Annexe 8 - Items originaux du TAM (Davis, 1989) et leur traduction par Huu Binh (2014) 105
Annexe 9 - Critères et items du questionnaire USE (Lund, 2001) et leur traduction par (Monfort et al., 2018) 106
Annexe 10 - Échelle SAM de Bradley & Lang (1994) 107
Annexe 11 - Computer Emotion Scale (Kay & Loverock, 2008) 108
Annexe 12 - Questionnaire sur l'état de présence (Witmer & Singer, 1998) 109
Annexe 13 - Échelle UES (O'Brien, Cairns & Hall, 2018) 111
Annexe 14 - Révision des items par les experts 112
Annexe 15 - Résultat du tri de cartes des items sous formes de phrases 127
Annexe 16 - Résultats du tri de cartes des items sous forme de paires de mots 129
Annexe 17 - Réponses au questionnaire 131
Annexe 18 - Formulaire de consentement 134
Annexe 19 - Guide d'entretien 138
Annexe 20 - Retranscriptions entretiens 148

# 1 Introduction et problématique

De nos jours, on note une évolution constante du domaine de la formation et du quotidien des étudiants, en lien avec la digitalisation, les nouvelles technologies et le besoin de se former tout au long de la vie (Barbier, 2016 ; Pultrone, 2020). On trouve alors des dispositifs de formation à distance qui tentent de s'adapter à la vie privée et professionnelle des apprenants. Néanmoins, Sauvé (2020) constate de forts taux d'abandon dans ces formations qui oscillent entre 20 à 40%. La question de l'encadrement des apprenants est alors fondamentale car selon Brassard & Teutsch (2014), « l'absence de contacts directs [...] peut créer un sentiment d'éloignement et d'isolement qui, s'il est trop prononcé, peut aboutir à un abandon de la formation » (p.1).

Nissen (2009) remarque que les fonctions d'encadrement d'un enseignant dans une formation hybride ou à distance sont nombreuses car, en plus de son rôle d'encadrement pédagogique, ce dernier doit également soutenir les apprenants notamment dans l'utilisation des outils technologiques. Ces fonctions sont décrites par Gérin-Lajoie (2018) et présentées dans le tableau 1, ci-dessous.

Tableau 1 - Les types d'encadrements en formation à distance selon Gérin-Lajoie (2018) et exemples d'application selon Nissen (2009)

| Type d'encadrement et définition | | Exemples d'application |
|---|---|---|
| **Pédagogique** | **Cognitif – « Faciliter l'apprentissage »**<br><br>➔ Comprendre pourquoi l'apprenant ne comprend pas ou anticiper ce qu'il pourrait ne pas comprendre. | ● Répondre aux questions des apprenants<br>● Suggérer des pistes pour aller plus loin<br>● Aide pédagogique<br>● Stimulation cognitive<br>● Facilitation des apprentissages<br>● Explications<br>● Corrections et commentaires des travaux<br>● Aides individuelles en fonction des difficultés rencontrées<br>● Recherches d'informations<br>● Réalisation de tâches |
| | **Méthodologique – « Outiller l'apprentissage »**<br><br>➔ Comment l'apprenant peut ou doit s'organiser pour faire ce qui lui est demandé. | ● Proposer des stratégies d'apprentissage<br>● Proposer des stratégies pour gérer son temps<br>● Proposer des outils pour le travail collaboratif |
| | **Métacognitif – « Comprendre l'apprentissage »**<br><br>➔ Comment l'apprenant fait pour prendre du recul sur son parcours d'apprentissage et devient autonome. | ● Aider les apprenants à gérer leur temps<br>● Promouvoir une réflexion à propos de leur manière d'apprendre |

| Type d'encadrement et définition | | Exemples d'application |
|---|---|---|
| | **Motivationnel – « Inciter à l'apprentissage »**<br><br>➔ Comment l'apprenant est invité à s'engager dans l'action et à persister. | ● Maintenir un contact régulier<br>● Encouragements et stimulations pour impliquer les apprenants et maintenir leur motivation<br>● Fournir des feedbacks rapides<br>● Valorisation de l'apprentissage<br>● Rassurer les apprenants |
| | **Socio-affectif – « Vivre l'apprentissage »**<br><br>➔ Comment l'apprenant ressent et vit sa formation. | ● Création d'un bon climat social et un environnement d'apprentissage chaleureux<br>● Favorisation du sentiment d'appartenance<br>● Incitation à la communication dans le groupe<br>● Gérer les confrontations d'idées entre les apprenants |
| | **Organisationnel – « Accompagnement dans l'apprentissage »**<br><br>➔ Comment guider les étudiants et les inciter à suivre un rythme de formation. | ● Apprendre à apprendre en fournissant des aides aux apprenants pour la gestion du travail (comprendre, raisonner et s'organiser)<br>● Rappel des délais et des dates limites<br>● Synthèses des modules ou des séances |
| **Administratif** | ➔ Comment gérer les études. | ● Règles, règlements, procédures, etc. |
| **Technologique** | ➔ Comment aider les apprenants à maîtriser les outils utilisés dans la formation et développer leurs compétences technologiques. | ● Aider l'apprenant avec le système pour que ce dernier puisse se concentrer sur ses tâches d'apprentissage<br>● Donner des explications sur l'utilisation et le fonctionnement des outils<br>● Assister les apprenants en cas de difficultés dans leur utilisation<br>● Aider à comprendre l'institution de formation |

Au vu du grand nombre de fonctions requises dans le cadre des formations hybrides et à distance, il apparaît nécessaire d'assister les enseignants afin de s'assurer de l'efficacité de ce type de formations et ainsi réduire leur taux d'abandon (Dussarps, 2015). Afin d'optimiser les méthodes d'enseignement et de faciliter l'accès à la formation, Karsenti (2003) recommande l'utilisation des outils numériques (p.225).

Force est de constater que suite à la croissance fulgurante des messageries instantanées, les chatbots deviennent de plus en plus populaires mais sont cependant très récents dans le domaine de la formation (Smutny & Schreiberova, 2020).

Un « chatbot » est un terme anglais qui est composé de deux aspects : le dialogue (« chat » pour la discussion) et l'automatisation (« bot » pour robot). Il désigne un programme conçu pour dialoguer avec des utilisateurs via une plateforme ou une messagerie. En français, on emploie également le terme « Agent conversationnel ». Son but est d'apporter des réponses automatisées (sans intervention humaine) aux demandes des utilisateurs et donc, d'échanger avec les utilisateurs par le biais de programmes de messageries instantanées, en se substituant à un individu et en simulant une conversation réelle. Pour ce faire, ils utilisent l'intelligence artificielle afin que l'utilisateur semble avoir une conversation la plus naturelle possible (Ward, 2005).

Les chatbots éducatifs présentent un grand intérêt dans le cadre des formations à distance dans le sens où ils peuvent être une alternative aux forums de discussion qui sont peu utilisés. L'idée est de proposer aux étudiants un chatbot (une sorte de Wiki interactif et de professeur virtuel) qui joue le rôle d'assistant pour les enseignants en répondant aux questions des apprenants. Ainsi, le chatbot permettrait d'améliorer et de personnaliser l'apprentissage, et de soutenir l'engagement des apprenants, car plus disponible que les enseignants, grâce à sa capacité de répondre automatiquement, et à tout moment, aux différentes demandes qui lui sont formulées (Smutny & Schreiberova, 2020). Mieux, des chatbots comme Duolingo[1], peuvent aider à satisfaire les conditions d'efficacité des formations à distance énoncées par Karsenti (2003) : (1) l'attractivité, la convivialité et la facilité d'accès à la formation; (2) les interactions entre les apprenants et les formateurs; (3) l'enseignement d'un contenu certifié par des experts du domaine éducatif et aussi riche qu'un cours donné en salle de classe; (4) la coopération et la collaboration entre apprenants; (5) la personnalisation de l'apprentissage et la diversification des approches pédagogiques; (6) le soutien technique et pédagogique et (7) l'évaluation du dispositif de la formation (p.231). En particulier, le tableau 2, illustre comment Duolingo, l'un des chatbots utilisés pour l'apprentissage des langues, satisfait les conditions d'efficacité des formations à distance (Karsenti, 2003).

[1] https ://www.duolingo.com

Tableau 2 - Les conditions d'efficacité d'une formation à distance et leur application dans Duolingo

| Conditions d'efficacité des formations à distance (Karsenti, 2003) | Leur application dans Duolingo |
|---|---|
| **Accès/Attrait** | ● Accessible à tous : gratuit<br>● Facile d'accès : disponible sur plusieurs plateformes<br>● Attractif grâce à l'avatar |
| **Interaction** | ● Apprentissage ludique grâce aux fonctionnalités, au système de points et aux défis proposés<br>● Apprentissage principalement en ligne mais des rencontres ou d'autres formes de contenus (podcast) sont également disponibles |
| **Contenu** | ● Le contenu respecte la norme internationale pour la maîtrise des langues<br>● Leçons courtes et interactives<br>● Leçons organisées par thématiques d'apprentissage<br>● Tests de niveaux après chaque leçon pour valider les connaissances |
| **Approche pédagogique** | ● Apprentissage organisé par niveaux de compétences<br>● Fixation d'objectifs personnels et défis ludiques pour augmenter la motivation |
| **Ressources** | ● Enseignement personnalisé et adapté au niveau<br>● Possibilité de répéter ou de passer des niveaux<br>● Exercices pratiques en lien avec des situations de la vie réelle<br>● Plusieurs types d'exercices : sélections de mots, prononciation, traduction de phrases complètes et transcription de ce qu'on entend |
| **Soutien** | ● Système d'aide grâce au forum |
| **Pérennité et aspects éthiques** | ● Contrôle permanent de l'efficacité de la méthode d'enseignement à l'aide des quiz de point de contrôle |

Néanmoins, l'émergence de chatbots (Smutny & Schreiberova, 2020) et l'abondance de produits à visée éducative dans le monde des technologies de l'apprentissage et de la formation (Crozat, Hû & Trigano, 1999) génèrent une véritable concurrence entre les différents outils présents sur le marché. On observe alors que, de manière générale, de nombreux outils à visée éducative ne répondent pas forcément aux besoins et aux attentes des utilisateurs (Ahaji et al., 2008). Ces deux aspects devraient pourtant être pris en compte par les concepteurs afin que les utilisateurs choisissent d'utiliser leur produit plutôt qu'un autre.

L'évaluation de l'outil permet de rendre compte de ces aspects et est « une étape clé dans le processus de conception ». Elle « consiste à tester la qualité des solutions produites », sur plusieurs itérations jusqu'à l'obtention de la solution la plus optimale possible (Lallemand & Gronier, 2015, p.485).

Dans le cadre du développement d'un outil numérique, en plus du concepteur, Nogry et al. (2004) distinguent deux types d'acteurs : les enseignants et les apprenants.

L'enseignant « amène l'utilisation d'un EIAH (Environnements Informatiques pour l'Apprentissage Humain) » en mandatant un concepteur afin de concevoir un outil pour les apprenants (« les utilisateurs finaux ») (p.40). Celui-ci a pour mission de concevoir un outil qui réponde à la fois au besoin du commanditaire mais également à la demande et aux besoins de l'utilisateur.

Ainsi, lorsqu'on évalue un produit à visée éducative, on peut choisir de l'évaluer soit du point de vue exclusif du concepteur en s'intéressant à évaluer son développement (Tavanapour & Bittner, 2018 ; Radziwill & Benton, 2017) ou alors, on peut évaluer l'adéquation de l'outil avec les besoins des utilisateurs finaux (Peras, 2018 ; Hobert, 2019). L'évaluation permet de renseigner sur les perceptions, attentes, et observations des différents acteurs (apprenants, enseignants, responsables de formation, etc.) suite à leur expérience avec le système pour que ce dernier réponde au mieux à leurs besoins (Temperman et al., 2021). Évaluer présente ainsi plusieurs avantages pour les différents acteurs (Ahaji et al., 2008) :

- Chez les apprenants : (1) l'amélioration des apprentissages et de leurs connaissances ;
- Chez l'enseignant, (2) l'aide à la sélection d'outils à intégrer dans le scénario d'apprentissage pour l'amélioration de l'enseignement, (3) la visualisation de l'état de la qualité et la progression des performances des apprenants ainsi que (4) la validation des objectifs d'apprentissage et des contenus ;
- Chez les dirigeants : (5) le choix des logiciels à acquérir et mettre à disposition des enseignants et des apprenants.
- Chez les concepteurs : (6) la validation des choix de conception ou (7) l'obtention de pistes d'amélioration afin de parfaire ces solutions (Lallemand & Gronier, 2015, p.485).

Cependant, il manque des méthodes appropriées et uniformes pour évaluer ces logiciels (Ahaji et al., 2008). Ce constat est identique concernant l'évaluation spécifique des chatbots qui manque de mesures systématiques et unifiées, autres que celles qui se centrent sur les aspects techniques (Peras, 2018). Dans le domaine éducatif, Hobert (2019) constate que l'évaluation se concentre sur le succès de l'apprentissage. Cependant, il convient de prendre en compte d'autres critères afin de les évaluer de manière exhaustive (p. 260).

Au vu de ce qui précède, il apparaît important de pouvoir s'assurer de la qualité et de l'adéquation de ces nouveaux produits avec les besoins des utilisateurs, en les évaluant de manière systématique et organisée, en amont, avant leur utilisation dans une séquence d'apprentissage (Ahaji et al., 2008). Toutefois, puisqu'il n'existe pas d'outil d'évaluation des chatbots (Meyer von Wolff, Hobert & Schumann, 2019, p.101 ; Peras, 2018), ce travail aura pour objectif de concevoir un outil qui permette d'évaluer un chatbot utilisé comme assistant dans le cadre d'une formation hybride ou à distance et de répondre à la question de recherche suivante : Comment évaluer un chatbot créé dans le but d'assister les enseignants dans le cadre d'une formation à distance ?

Dans le prochain chapitre, qui traite du cadre théorique, nous allons parcourir les approches théoriques qui nous permettront de répondre à la question de recherche. Ensuite, dans la partie consacrée à la méthodologie, nous expliquerons le processus d'élaboration de l'outil d'évaluation des chatbots assistants de cours ainsi que la technique de sélection des critères à prendre en compte. Enfin, dans la partie résultats, nous détaillerons les pistes d'amélioration obtenues grâce au test de l'outil réalisé, en contexte réel, pour évaluer un chatbot en cours de conception.

# 2 Cadre théorique

## 2.1 Qu'est-ce qu'un chatbot ?

Selon Smutny & Schreiberova (2020), les chatbots sont des assistants digitaux capables de fournir des informations, de répondre à des questions, de discuter de sujets spécifiques ou de réaliser une tâche à la demande des utilisateurs sans intervention humaine (p.1).

Les chatbots sont utilisés principalement dans quatre domaines : l'actualité, la restauration, les transports et le divertissement (Ronce, 2018), notamment pour réaliser des tâches nombreuses et variées telles que la réservation d'hôtels, la recherche et la commande de biens, l'information sur divers sujets et la mise en place d'automatismes afin de recevoir des alertes à propos d'un secteur d'activités.

Nieves (2018), cité par Garcia Brustenga et al. (2018), décrit 4 éléments qui constituent un chatbot. Le premier est l'intelligence artificielle conversationnelle qu'il identifie comme le principe de base d'un chatbot. Un chatbot primaire se focalise sur l'interprétation et la reconnaissance de schémas (*patterns*) et de règles prédéfinies. Pour les chatbots plus avancés, ceux-ci sont capables d'analyser l'entrée (*input*) et d'apprendre des conversations et ainsi générer la meilleure réponse possible au regard de ce qu'ils ont vu par le passé. Le deuxième élément est l'expérience utilisateur qui permet d'avoir une conversation naturelle, intelligente et cohérente. L'interface, qui constitue le troisième élément, permet à l'utilisateur de voir ou d'entendre la discussion avec le chatbot. Enfin, le dernier élément qui caractérise le chatbot, concerne le design conversationnel qui permet à l'interaction artificielle d'avoir une logique humaine.

## 2.2 Les chatbots dans le domaine éducatif

Garcia Brustenga et al. (2018) distinguent deux types de chatbots dans le domaine de l'éducation : (1) les chatbots sans intention éducative et (2) les chatbots avec intention éducative (voir Tableau 3).

Tableau 3 - Types de chatbots en éducation (Garcia Brustenga et al., 2018)

| Types de chatbots en éducation | |
|---|---|
| **Sans intention éducative** | **Chatbots utilisés pour les tâches administratives (orientation, assistance personnelle) et de support (FAQ)** |

| | | |
|---|---|---|
| **Avec intention éducative** | **2 types de chabots conçus pour favoriser l'enseignement et l'apprentissage :** | 1. Des assistants permettant de construire le processus d'apprentissage<br>- Adapter le contenu pédagogique en fonction des besoins et du rythme de l'étudiant.<br>- Favoriser la réflexion<br><br>2. Un programme d'exercices pratiques pour l'acquisition de compétences<br>- Question - Réponse<br>- Évaluation immédiate de la réponse par le chatbot |

Les chatbots (1) sans intention éducative peuvent prendre en charge des tâches administratives et de gestion pour favoriser l'intégration et la productivité personnelle en fournissant une assistance personnelle aux étudiants comme la gestion du calendrier des tâches, la date limite de soumission d'un devoir ou les rappels d'évaluations. Les chatbots peuvent également s'occuper de la foire aux questions afin de fournir rapidement une réponse aux questions fréquentes concernant par exemple les aspects administratifs (admissions, inscriptions, services financiers, problèmes techniques, etc.).

Pour les chatbots (2) avec intention éducative, ceux-ci sont divisés en 2 sous-catégories : (a) les chatbots qui proposent des programmes d'exercices et (b) les chatbots de tutorat. Concernant la première sous-catégorie (a), ces types de chatbots jouent le rôle de simulateur en permettant la mise en pratique concrète de compétences et de capacités spécifiques. Dans l'apprentissage des langues par exemple, cet outil permet de simuler des conversations dans des contextes organisés par niveau, favorisant ainsi l'acquisition du vocabulaire. Ces simulations sont également souvent utilisées dans le domaine de la santé afin d'imiter le traitement des patients. Par ailleurs, ces chatbots peuvent également participer à l'évaluation des élèves en agissant en tant qu'évaluateur. Ainsi, l'élève bénéficie d'un accompagnement personnalisé grâce au retour d'expérience et à l'adaptation du processus d'apprentissage en fonction de ses besoins.

Quant à la deuxième sous-catégorie (b), ces types de chatbots jouent le rôle d'assistants, tuteurs ou mentors en accompagnant et en apportant des réponses à l'étudiant tout au long de son processus d'apprentissage. Ils peuvent surveiller la compréhension de l'élève, faire des suggestions en cas de besoin, mais également apporter un soutien émotionnel en contribuant à contrôler le comportement de l'élève au cours de son cursus en en lui apportant un renforcement motivationnel. Dans le même registre, le chatbot peut également amener l'élève à réfléchir sur ses stratégies métacognitives pour l'aider à réguler ses processus d'apprentissage. Le chatbot peut également jouer le rôle de camarade de classe en offrant un soutien pour faciliter les

apprentissages et le partage des connaissances. Ainsi, grâce à l'ensemble des processus évoqués, la perte de motivation ou l'abandon peuvent être évités grâce à cette nouvelle forme d'accompagnement.

Dans le cadre de ce mémoire, nous allons nous intéresser plus particulièrement à l'évaluation de ce dernier type de chatbot (chatbots assistants). En effet, dans le cadre des formations à distance, l'accompagnement et l'encadrement des apprenants apparaît comme essentiel afin de soutenir leur engagement et ainsi réduire le taux d'abandon (Brassard & Teutsch, 2014). Nous avons donc pour premier objectif de trouver une méthode d'évaluation permettant de rendre compte dans quelles mesures ces chatbots assistants répondent aux besoins des apprenants.

## 2.3 Comment évaluer un chatbot ?

L'émergence des chatbots éducatifs étant relativement récente, la littérature en lien avec l'évaluation des chatbots éducatifs n'en est encore qu'à ses prémices (Kukulska-Hulme et al., 2021). Néanmoins, il est possible de trouver des exemples concrets et des pistes d'utilisation dans le domaine de l'éducation (Hobert, 2019).

### *2.3.1 Les méthodes d'évaluation des chatbots*

Hobert (2019) identifie 4 procédures d'évaluation : (1) Wizard-Of-Oz, (2) les études en laboratoire, (3) les études sur le terrain et (4) la validation technique. Outre ces 4 approches, Morrissey & Kirakowski (2013) proposent d'utiliser la technique de l'incident critique afin d'analyser les dialogues entre le chatbot et l'utilisateur.

#### *2.3.1.1 Wizard-Of-Oz*

La première approche utilise des implémentations logicielles fonctionnelles au stade de prototype. Une approche courante pour ce type d'évaluation des chatbots pédagogiques consiste à mener des expériences de type « Wizard-of-Oz ». Cette évaluation de type expérimentale donne accès aux participants à une version semi-fonctionnelle du chatbot à évaluer. Toutes les fonctionnalités n'étant pas encore implémentées, c'est un opérateur à distance qui prend en charge le contrôle et répond aux demandes des utilisateurs. L'objectif de l'évaluation est de donner aux utilisateurs l'impression d'interagir avec un chatbot. Ainsi, en utilisant cette approche, on peut évaluer l'acceptation de la technologie ou encore la facilité d'utilisation du système.

#### *2.3.1.2 Études en laboratoire et études sur le terrain*

La seconde approche est relative aux études en laboratoire et la troisième traite des études sur le terrain. Pour ces deux types d'études, le chatbot doit en général être fonctionnel et la majeure partie des fonctionnalités doit être effectuée. Dans la plupart des cas, les implémentations sont souvent plus avancées dans le cadre des études sur le terrain afin de ne pas nécessiter l'intervention d'un opérateur à distance. En définitive, la conduite d'études en laboratoire est adaptée pour tester des prototypes dans un stade avancé du processus de développement. En revanche, les études de

terrain nécessitent un état d'avancement plus abouti et proche de la mise en production.

#### *2.3.1.3 Validation technique*

Il existe une quatrième approche d'évaluation qui n'implique pas d'interaction entre le logiciel et des participants humains : les validations techniques qui permettent de vérifier l'exactitude des algorithmes. Les chercheurs entrent des *inputs* et vérifient le langage naturel du système en examinant si l'*output* est correct et conforme aux attentes. Ce type d'évaluation algorithmique permet d'évaluer l'adéquation du système du langage naturel même au stade de prototype. Selon Hobert (2019), cette dernière approche est particulièrement importante dans l'évaluation des agents conversationnels car c'est un facteur de réussite déterminant.

#### *2.3.1.4 Technique de l'incident critique*

Dans cette cinquième méthode d'évaluation, l'utilisateur est amené à tester l'outil puis, il lui est demandé d'identifier les moments particuliers de l'expérience qui ont posé problème. Pour ce faire, l'utilisateur dispose de la transcription de l'interaction où il a pour tâche de souligner, dans un premier temps, les aspects de la conversation qui lui ont semblé particulièrement problématiques et d'en expliquer les raisons puis, dans un second temps, les aspects qui, au contraire, lui ont semblé convaincants. Ainsi, grâce à cette méthodologie, il est possible d'identifier les forces et les faiblesses du chatbot. En revanche, cette méthode ne permet pas de donner une indication précise sur comment les utilisateurs ont perçu le niveau de gravité de l'erreur et l'impact sur la conversation.

Dans le cadre de ce travail, nous allons nous intéresser à l'approche « Wizard-of-Oz » ainsi qu'aux études sur le terrain afin de pouvoir évaluer les chatbots dès le stade de prototype.

### *2.3.2 Les outils d'évaluation des chatbots*

Hobert (2019) identifie 5 instruments de mesure différents : (1) les questionnaires post-test utilisés pour évaluer la technologie, (2) les questionnaires pré et post-test utilisés pour évaluer l'apprentissage, (3) les entretiens qualitatifs, (4) les logs d'utilisation et (5) les transcriptions.

D'abord, notons (1) les questionnaires post-test utilisés pour évaluer l'acceptation, l'utilisation et l'utilisabilité de la technologie, et (2) les questionnaires pré et post-test utilisés pour évaluer le succès de l'apprentissage en comparant les connaissances avant et après l'utilisation du système. Outre les données quantitatives, les questionnaires permettent également de recueillir des données qualitatives de tous types, sous la forme de prise d'informations déclarées par exemple (Temperman et al., 2021), et ont l'avantage de pouvoir être administrés à un nombre important de participants (Lallemand & Gronier, 2015) tout en offrant une facilité d'analyse (Temperman et al., 2021). Toutefois, les questionnaires sont moins adaptés pour la

collecte de certaines données qualitatives comme l'expression des émotions. Dans tous les cas, la formulation des items quantitatifs ou qualitatifs est importante afin d'éviter qu'ils soient mal interprétés (Lallemand & Gronier, 2015).

Ensuite, relevons (3) les entretiens qualitatifs avec les utilisateurs qui permettent de connaître leur expérience, leur ressenti, leurs besoins ainsi que leur comportement avec le chatbot. Lallemand & Gronier (2015) précisent qu'afin d'obtenir les informations souhaitées, un guide d'entretien peut être créé et préparé en amont. Ainsi, « contrairement aux questionnaires, les entretiens laissent une grande liberté d'expression aux participants et une grande flexibilité aux concepteurs » (p.118). En revanche, cette méthode est plus chronophage en raison, d'une part, du temps à consacrer à l'entrevue, et d'autre part, du traitement manuel des réponses (codage) pour vérifier l'exactitude des données récoltées.

Enfin, il y a (4) les logs d'utilisation et (5) les transcriptions qui sont des outils de mesure techniques permettant l'obtention de données statistiques. Les fichiers de logs permettent d'obtenir des données quantitatives sur l'utilisation du chatbot. Les transcriptions quant à elles, permettent d'analyser les dialogues pour évaluer l'interaction entre les utilisateurs et le chatbot. Cette analyse des discussions permet d'obtenir des corrections de logique et de comportement à apporter pour que l'interface ait un langage le plus naturel possible et puisse répondre de manière correcte aux différentes demandes des apprenants. Notons toutefois que l'analyse de la conversation demande un investissement de temps important car les données doivent être traitées manuellement (en codant les réponses du chatbot).

En plus de ces outils d'évaluation spécifiques aux chatbots, Lallemand & Gronier (2015) citent deux autres méthodes d'évaluation : (6) les focus group et (7) l'observation. Le focus group est une réunion d'un petit groupe d'individus qui participent à une discussion sur un sujet ou une question donnée (Rinkus et al., 2020, p.327) et au cours duquel ils discutent ensemble de leurs expériences et ressentis, ce qui permet de collecter des données qualitatives concernant notamment leurs opinions ou leurs besoins. L'avantage de ces discussions est qu'elles permettent de collecter un nombre important de données tout en réduisant le temps de passation. En effet, « les échanges au sein du groupe et les émulations entre les membres qui en découlent, sont souvent source d'une grande profusion de données verbales qui nécessiteraient plusieurs entretiens individuels » (Lallemand & Gronier, 2015, p.140). Les focus group permettent donc de récolter un maximum de données dans un laps de temps court. Toutefois, ils ne réduisent pas la quantité de données à analyser et ne permettent pas « d'obtenir des informations plus intimes et donc plus riches » (Lallemand & Gronier, 2015, p.141).

Enfin, les observations peuvent se faire en laboratoire ou sur le terrain. L'observation sur le terrain vise à observer l'utilisateur dans son environnement naturel afin de récolter des données qualitatives relatives à ses comportements, attitudes,

verbalisations, interactions sociales ou technologiques, besoins, envies, etc. L'avantage que présente cette technique de collecte des données par rapport aux entretiens et focus group est qu'elle permet « d'accéder aux comportements non verbaux dont les participants n'ont pas toujours conscience » (p.160), permettant ainsi de relever l'écart entre ce que déclare faire un individu et ce qu'il fait réellement. Toutefois, avec cette méthode, il est possible d'être confronté à d'autres biais tels que la présence de l'observateur qui peut venir modifier le comportement de l'utilisateur observé ou encore l'observateur qui peut avoir une mauvaise interprétation de la situation qu'il observe.

Hobert (2019) constate que les méthodes les plus fréquemment utilisées sont les enquêtes quantitatives avant et/ou après une étude sur le terrain ou en laboratoire.

### *2.3.3 Choix de l'outil pour évaluer les chatbots assistants utilisés en formation à distance*

Dans le cadre de ce travail, nous allons élaborer un questionnaire, car au regard des différents éléments relevés dans la littérature, le questionnaire est la forme la plus utilisée pour l'évaluation de logiciels pédagogiques (Hû, Trigano & Crozat, 2001) ainsi que pour évaluer les chatbots (Hobert, 2019 ; Pérez, Daradoumis & Puig, 2020).

Cette méthode nous semble également appropriée dans le cadre de notre recherche car, le questionnaire pourra, d'une part, être diffusé à une plus large échelle (Lallemand & Gronier, 2015) et d'autre part, permettre une analyse des résultats plus simple (Temperman et al., 2021) et moins chronophage que pour d'autres méthodes comme les entretiens par exemple. En outre, Redet et al. (2016), contrairement à Lallemand & Gronier (2015), préconisent l'utilisation du questionnaire afin d'évaluer le ressenti des utilisateurs. En effet, « les outils d'auto-évaluation subjective apparaissent comme indispensables pour interroger l'émotion en UX, car il s'agit d'une gamme d'outils qui permet d'interroger les aspects émotionnels de l'expérience », en plus d'être un « moyen privilégié pour l'obtention de données standardisées » (Redet et al., p.3).

Pour construire un questionnaire, Coron (2020) met en évidence 3 étapes : (1) définir les objectifs d'évaluation, (2) créer les questions et (3) tester le questionnaire (p.36). Nous allons, dans un premier temps, rechercher dans la littérature les critères et items d'évaluation à prendre en considération, selon la méthodologie plus détaillée de O'Brien & Toms (2010) afin de nous permettre d'évaluer les chatbots utilisés comme assistants de cours en formation à distance de la manière la plus complète possible.

## 2.4 Définition des critères et des items d'évaluation

Dans le cadre de l'évaluation des chatbots assistants de cours, nous souhaitons évaluer la qualité de l'expérience utilisateur avec le chatbot afin de s'assurer que celui-ci réponde aux besoins des apprenants. La question n'est donc pas seulement de se

concentrer sur l'efficacité, l'efficience ou la satisfaction, mais de prendre en considération la capacité de l'outil dans l'engagement et l'expérience des utilisateurs. En effet, les utilisateurs décident d'utiliser certains outils en fonction notamment de leur pertinence et de leur utilité (Bétrancourt, 2007) c'est pourquoi, il faut en tenir compte lors de leur conception.

### *2.4.1 Expérience utilisateur*

La notion d'expérience utilisateur renvoie à différents aspects de l'interaction entre l'utilisateur et un produit que Barcenilla & Bastien (2009) résument en 3 dimensions : (1) les qualités instrumentales qui correspondent à l'utilité perçue et aux autres composantes de l'utilisabilité et (2) les qualités non instrumentales « (esthétique, valeurs véhiculées, facteurs motivationnels, etc.) ». Ces deux dimensions réfèrent à la qualité du produit, et influencent la troisième dimension : (3) les réactions émotionnelles telles que des « sentiments subjectifs, expressions motrices et comportementales, réactions physiologiques, évaluation cognitive » (p.324).

Ces dimensions sont détaillées dans le modèle d'évaluation de Hû, Trigano & Crozat (1999, 2001) qui proposent une structure globale d'évaluation d'un logiciel de formation composé de 6 critères d'évaluation : (1) la qualité technique, (2) l'ergonomie, (3) les documents multimédias, (4) la scénarisation, (5) les outils pédagogiques et (6) les impressions générales. L'EMPI (Évaluation des logiciels Multimédia Pédagogiques Interactifs) associe à chacun de ces 6 critères généraux, des méta critères qui ont pour objectif de détailler et d'analyser plus finement le critère évalué. Un dernier niveau de précision est également prévu par les auteurs qui divisent encore ces méta critères en sous-critères d'évaluation. Cette décomposition permet de formuler des questions pour évaluer ces différentes dimensions (voir Annexe 1). Toutefois, en fonction du type de logiciel évalué, l'ensemble des thèmes énoncés ne s'appliquent pas forcément. Les auteurs proposent d'évaluer les impressions générales et le ressenti des utilisateurs à l'aide de paires de mots. Cette forme d'item permet de « donner une note dite instinctive [...] sur une échelle qualitative sans position neutre pour contraindre l'évaluateur à prendre une décision » (p.263-264).

Un questionnaire d'évaluation (AttrakDiff2), proposé par Hassenzahl et al. (2003), permet de mesurer l'expérience utilisateur d'un dispositif de manière globale. Ce questionnaire utilise également des items sous forme de paires de mots est administré aux utilisateurs après leur utilisation de l'environnement à évaluer. Le questionnaire a été développé en allemand par Hassenzahl et al. (2003) puis traduit en français par Lallemand et al. (2015). Il existe une version longue comprenant 28 items ainsi qu'une version courte qui se compose de 10 items. Pour chacun, une échelle de discrimination sémantique en 7 points est proposée (voir Annexe 2). L'évaluation est composée de 4 sous-échelles :

- Qualité pragmatique : concerne la partie dite « instrumentale » de l'outil, la clarté du système ainsi que sa structure. On s'intéresse à la réalisation de tâches ou d'objectifs, ce qui reflète l'utilité et l'utilisabilité de l'outil ;

- Qualité hédonique-identification : à l'inverse de la qualité pragmatique, concerne l'utilisateur. On s'intéresse à comprendre comment l'outil encourage l'utilisateur ;
- Qualité hédonique-stimulation : concerne le développement des connaissances et des compétences des utilisateurs de l'outil ;
- Attractivité : correspond à l'appréciation globale de l'outil selon la qualité perçue.

Laugwitz et al. (2008) proposent également un questionnaire d'évaluation pour la mesure de l'expérience utilisateur : le « User Experience Questionnaire (UEQ) ». Celui-ci comprend 26 items répartis en 6 sous-échelles :

- Attractivité : correspond aux impressions générales envers le système. Les utilisateurs aiment-ils ou non le système ?
- Efficacité : Est-il possible d'utiliser le système de manière rapide et efficace ? L'interface utilisateur semble-t-elle organisée ?
- Clarté : Est-il facile de comprendre comment fonctionne le système ou de se familiariser simplement avec le système ?
- Fiabilité : L'utilisateur a-t-il le sentiment de contrôler l'interaction ? L'interaction avec le système est-elle sûre et prévisible ?
- Stimulation : L'usage du système est-il intéressant et stimulant ? L'utilisateur est-il motivé à en poursuivre l'usage ?
- Nouveauté : Le design du système est-il innovant et créatif ? Le système retient-il l'attention de l'utilisateur ?

Schrepp & Thomaschewski (2019) proposent une extension de ce questionnaire (UEQ+). Les items sont répartis en 16 sous-critères ((1) l'attractivité, (2) l'efficacité, (3) la transparence, (4) la maniabilité, (5) la stimulation, (6) l'originalité, (7) la confiance, (8) le toucher, (9) l'acoustique, (10) la flexibilité, (11) l'utilité, (12) la valeur, (13) l'esthétique visuelle, (14) l'utilisation intuitive, (15) la crédibilité et (16) la qualité de contenu).

Les items utilisés dans ces « User Experience Questionnaire » sont également présentés sous forme de paires de mots et l'utilisateur doit se positionner sur une échelle en 7 points. Ce questionnaire a été traduit dans plus de 30 langues dont une version en français (voir Annexe 3).

Dans le cadre de l'évaluation de chatbots utilisés dans le domaine éducatif qui font l'objet de ce mémoire, nous allons rechercher dans la littérature les critères d'évaluation spécifiques qui leur sont dédiés. Notre second objectif est de s'assurer de prendre en compte suffisamment de critères afin de pouvoir évaluer le chatbot de la manière la plus complète possible.

### *2.4.2 Critères d'évaluation des chatbots éducatifs selon Hobert (2019)*

À notre connaissance, seul Hobert (2019) s'est intéressé à l'évaluation des chatbots dans le domaine de l'éducation. L'auteur identifie 7 critères dans la littérature : (1) acceptation et adoption, (2) réussite de l'apprentissage, (3) augmentation de la motivation, (4) autres effets bénéfiques, (5) utilisabilité, (6) justesse technique et (7) autres facteurs psychologiques. Le tableau 4, ci-dessous, récapitule la classification présente dans l'étude de Hobert (2019) et permet de mettre en évidence les liens entre les différents critères ou objectifs d'évaluation que l'auteur juge utiles pour une évaluation globale des chatbots pédagogiques, ainsi que les procédures et les instruments de mesure cités précédemment.

Tableau 4 - Classification de la revue de littérature qui inclut les dimensions procédures, instruments de mesure et objectifs d'évaluation de Hobert (2019)

| | | **Objectifs d'évaluation** | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | **Acceptation et adoption** | **Réussite de l' apprentissage** | **Augmentation de la motivation** | **Autres effets bénéfiques** | **Utilisabilité** | **Justesse technique** | **Autres facteurs psychologiques** |
| **Procédures** | Expérience Wizard-of-Oz avec une simulation du chatbot pédagogique non-fonctionnel | x | | | | x | | x |
| | Étude en laboratoire avec le chatbot pédagogique fonctionnel en conditions contrôlées | x | x | x | x | x | | x |
| | Étude sur le terrain avec le chatbot pédagogique fonctionnel sur site | x | x | x | x | x | | x |
| | Validation technique d'algorithmes ou des composants du logiciel sans dispositif expérimental | | | | | | x | |
| **Instruments de mesure** | Enquête quantitative après une étude en laboratoire ou sur le terrain avec les utilisateurs du chatbot pédagogique (post-test) | x | x | x | x | x | | x |
| | Enquête quantitative avant et après une étude en laboratoire ou sur le terrain avec les utilisateurs du chatbot éducatif (pré/post-test) | x | x | x | x | x | | x |
| | Étude qualitative sous forme d'entretiens avec les utilisateurs du chatbot | | x | | | x | | x |
| | Analyse du dialogue au moyen des transcriptions de l'interaction des apprenants avec le chatbot pédagogique | x | | | | | x | |
| | Analyse des fichiers de logs techniques générés durant l'utilisation du chatbot pédagogique | | | | | | x | |
| **Légende** : x indique une intersection dominante identifiée dans la revue de littérature. | | | | | | | | |

#### *2.4.2.1 Acceptation, adoption et acceptabilité*

Hobert (2019) constate que parmi les objectifs d'évaluation les plus présents dans la littérature, on retrouve l'acceptation et l'adoption de la technologie. En effet, « environ 30% des études analysées se sont concentrées sur cet objectif d'évaluation principal » (Hobert, 2019, p.264).

L'acceptabilité étudie la manière dont « l'utilisateur s'approprie une technologie, c'est-à-dire qu'il l'adopte, l'adapte et l'intègre à son activité » (Barcenilla & Bastien, 2009, p.318) en analysant le contexte d'utilisation du système et les comportements des utilisateurs (p.311).

Afin d'évaluer l'acceptabilité d'un système, Venkatesh et al. (2003) proposent le modèle « Unified Theory of Acceptance and Use of Technology » (UTAUT). L'objectif de ce modèle est de donner un point de vue complet sur le critère de l'acceptabilité grâce à l'intégration de différentes théories présentes dans la littérature en lien avec l'acceptabilité et l'utilisation des technologies. Pour ce faire, ils regroupent 8 modèles préexistants : la théorie de l'action raisonnée (TAR) (Ajzen & Fishbein, 1975) ; le modèle d'acceptation de la technologie (TAM) (Davis et al., 1989) ; la théorie du comportement planifiée (TCP) (Ajzen, 1991) ; la théorie du comportement interpersonnel (Triandis, 1980) ; le modèle de l'utilisation du PC (MPCU) (Thompson et al., 1991) ; la théorie de la diffusion des innovations (DI) (Rogers, 1983) ; le modèle de la motivation (Davis et al., 1992) et la théorie cognitive sociale (TCS) (Bandura, 1989 ; Compeau & Higgins, 1995). Ces modèles se basent sur les variables ayant été les plus significatives pour décrire l'acceptation de la technologie et leur influence sur le comportement d'utilisation. Le questionnaire se compose de 31 items (voir Annexe 4).

Le regroupement des modèles aboutit aux sous-critères suivants évalués dans l'UTAUT de Venkatesh et al. (2003). La Figure 1 illustre les liens entre ces sous-critères avec les données de genre, d'âge, d'expérience et d'utilisation volontaire du système :

- Performance attendue
- Effort attendu
- Attitude vers les technologies
- Influence sociale
- Conditions de facilitation
- Efficacité personnelle
- Anxiété
- Comportement d'intention d'utiliser le système

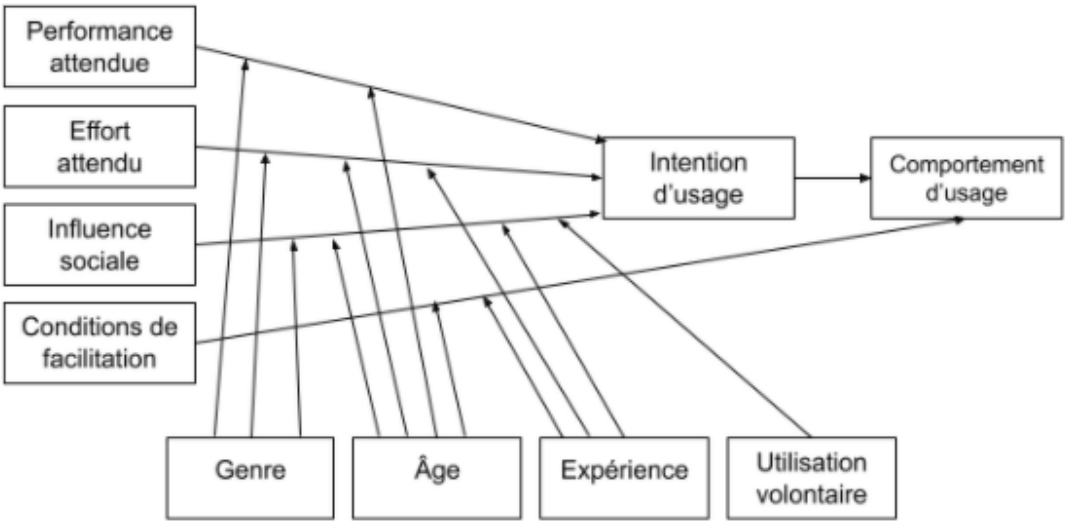

Fig. 1 - Modèle de recherche (Venkatesh et al., 2003)

D'autre part, Caron & Heutte (2017) proposent une échelle d'acceptation instrumentale appliquée aux Technologies de l'Information et de la Communication pour l'Enseignement (TICE). Ce questionnaire est construit sur deux échelles allant de 1 (« Pas du tout d'accord ») à 7 (« Tout à fait d'accord »). La première contient une liste de questions à destination des formateurs et la deuxième une liste de questions à destination des apprenants (voir Annexe 5). Le modèle rend compte des facteurs qui influencent l'utilité et l'usage d'un système, comme illustré dans la figure 2.

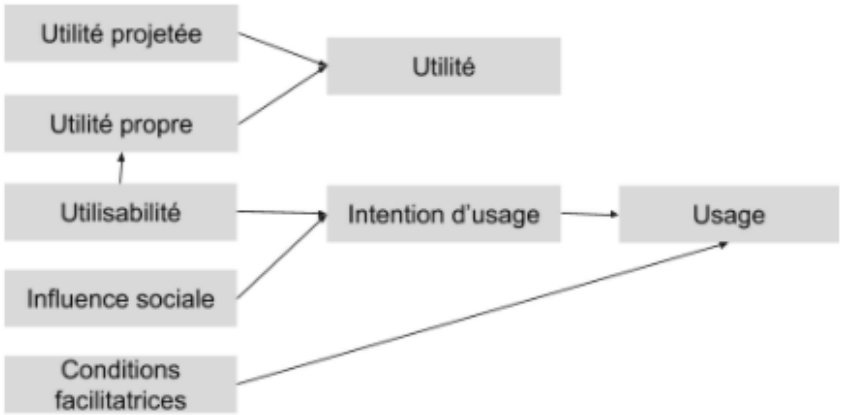

Fig. 2 - Modèle de l'Acceptation Instrumentale appliquée aux TICE (Caron & Heutte, 2017)

Notons que l'échelle destinée aux formateurs ne s'applique pas à notre public cible car le questionnaire sera utilisé par les utilisateurs finaux, à savoir les étudiants. Nous conservons alors uniquement l'échelle destinée aux apprenants. Les sous-critères évalués par cette seconde sous-échelle du modèle de Caron & Heutte (2017) sont :

- Utilité de façon générale
- Utilité projetée sur le formateur pour enseigner
- Utilité apprenant
- Utilisabilité
- Injonction
- Utilité pour moi-même

#### *2.4.2.2 Réussite de l'apprentissage*

L'utilisation d'un chatbot pour la réussite de l'apprentissage est étroitement liée à l'effet de la motivation (Hobert, 2019, p.264). L'auteur constate que seules 33% des études se sont concentrées sur l'évaluation de la réussite.

Toutefois, selon l'auteur, ce critère ne devrait pas être évalué en phase de développement car c'est une tâche complexe qui nécessite des instruments de mesures élaborés. En effet, pour pouvoir évaluer la réussite de l'apprentissage, il est nécessaire que le chatbot soit pleinement opérationnel, y compris le traitement du langage naturel (p. 267).

Dans le cadre de ce travail, nous allons donc exclure ce critère car nous souhaitons pouvoir évaluer un chatbot dès le stade de prototype afin de proposer des pistes d'amélioration à son concepteur.

#### *2.4.2.3 Motivation*

Bien que la motivation figure « au premier rang parmi les facteurs qui influencent la réussite » (Rebah & Dabove, 2017), Hobert (2019) constate que les recherches ne mesurent pas le lien entre l'augmentation de la motivation et l'effet sur la réussite de l'apprentissage, c'est pourquoi il choisit de mettre en évidence ce critère de manière distincte (p.264). Rebah & Dabove (2017) suggèrent également d'analyser les aspects motivationnels dans les apprentissages car ceux-ci peuvent être corrélés à des aspects psychologiques en lien avec les tâches à réaliser ou encore à des aspects relationnels et sociaux. La motivation des apprenants joue effectivement un rôle clé dans les apprentissages car elle génère différentes réactions chez l'apprenant comme la curiosité, la persévérance ou encore la performance (Vallerand et al., 1989, p.1). De plus, « les apprenants sont plus impliqués dans la construction de leurs savoirs en y participant activement grâce aux outils numériques mis à leur disposition » car cela « permet aux étudiants d'optimiser la tâche d'apprendre, et ce de façon autonome, organisée et structurée ». Pour Brandtzaeg et al. (2017) étudier la motivation à utiliser un agent conversationnel revient à « aider les développeurs à faciliter de meilleures interactions avec les chatbots » (p.1) et permet également « de surmonter les difficultés liées à l'adoption de la technologie » (p.2).

Vallerand et al. (1989) relèvent l'existence de trois formes principales de motivation. La première concerne la motivation intrinsèque qui correspond au fait de réaliser une tâche parce que celle-ci nous procure du plaisir. La deuxième forme est la motivation de type extrinsèque, où l'individu réalise la tâche pour une raison externe, par exemple, l'obligation de réaliser l'activité. Enfin, la troisième forme concerne l'amotivation qui correspond à l'absence de motivation à réaliser la tâche car l'utilisateur ne ressent pas de lien entre la tâche à réaliser et les tenants et aboutissants ou bénéfices de celle-ci.

Afin d'identifier et d'évaluer les principaux facteurs de motivation d'utilisation des chatbots, Brandtzaeg et al. (2017) proposent un questionnaire composé de 17 items

évaluant l'utilisation des chatbots et les expériences vécues, la durée d'utilisation, la motivation ainsi que des données démographiques (p.6). Les questions en lien avec la motivation sont posées sous forme de questions ouvertes afin de cerner les principales motivations des utilisateurs. Ce type de questions permet à l'utilisateur de rédiger sa réponse de manière libre, ce qui encourage les participants à fournir une description plus détaillée (p.5) :

- « Qu'est-ce qui vous motive à utiliser un chatbot ? »
- « Quelle est votre principale raison d'utiliser un chatbot ? »

Vallerand et al. (1989) proposent également une échelle de motivation en éducation (EME) permettant de mesurer les différents types de motivation. « C'est l'un des outils de mesure les plus utilisés en éducation » (Heutte, Caron, Fenouillet & Vallerand, 2016, p. 101) L'outil se compose de 7 sous-échelles :

- Amotivation : correspond à l'absence de motivation et reflète le manque de compréhension lors de la réalisation de la tâche ;
- Régulation externe : l'individu est motivé par un facteur externe comme par exemple une récompense après avoir accompli la tâche ;
- Régulation introjectée : l'individu est motivé par un facteur interne comme par exemple se prouver à lui-même qu'il en est capable ;
- Régulation identifiée : l'individu ressent que la réalisation de la tâche fait partie d'une suite logique d'événements et il est motivé par la poursuite de cet objectif, par exemple dans une formation en plusieurs niveaux où la réussite du niveau précédent conditionne le passage au niveau supérieur ;
- Intrinsèque connaissance : grâce à la réalisation de la tâche, l'individu ressent du plaisir à continuer la poursuite de ses objectifs ;
- Intrinsèque accomplissement : grâce à la réalisation de la tâche, l'individu ressent du plaisir à se prouver à lui-même qu'il en est capable ;
- Intrinsèque sensation : grâce à la réalisation de la tâche, l'individu ressent du plaisir de manière générale.

Sur la base de cette échelle de motivation en éducation (EME), Fenouillet, Heutte & Vallerand (2015) proposent l'échelle de motivation en formation des adultes (EMFA) spécialement conçue pour la formation des adultes et utilisable dans des dispositifs de formation en ligne. Elle est composée de 24 items, chacun associé à une échelle en 7 points allant de « Ne correspond pas du tout » à « Correspond très fortement » (voir Annexe 6).

#### *2.4.2.4 Autres effets bénéfiques sur les processus d'apprentissage : comportement et humanité*

Hobert (2019) relève que d'autres objectifs d'évaluation ont également été évalués dans la littérature mais ne peuvent pas être regroupés sous un terme spécifique, c'est pourquoi il choisit de créer le terme générique « autres effets bénéfiques ». Cette catégorie vise à évaluer des critères qui ne sont pas explicitement attribuables au succès de l'apprentissage ou à la motivation. Ces effets correspondent aux impacts

positifs que le chabot pédagogique peut apporter pour aider les apprenants à comprendre un sujet donné en analysant la manière dont la conversation naturelle du chatbot influence positivement la motivation (p.264). En effet, le caractère naturel et humanisant de l'interaction est l'une des caractéristiques essentielles des agents conversationnels, car il influence leur acceptation auprès des utilisateurs (Morrissey & Kirakowski, 2013, p. 87-88).

L'aspect humain (humanité) d'un chatbot est mesuré par la capacité du système de dialogue à comprendre le sens et l'intention des demandes ainsi qu'à répondre de manière appropriée aux indices sociaux (humeur) de l'utilisateur (Radziwill & Benton, 2017) et à faire croire aux utilisateurs qu'ils interagissent avec une personne réelle et non pas virtuelle (Morrissey & Kirakowski, 2013, p.88) en utilisant un langage le plus naturel possible. AbuShawar et al. (2016) proposent d'évaluer ce critère en comparant un corpus de conversation au niveau lexical, syntaxique et sémantique d'une part, entre l'ordinateur et l'utilisateur et d'autre part, avec un corpus existant de dialogues réels entre humains. Pour ce faire, les auteurs préconisent d'utiliser la technique des incidents critiques citée plus haut et proposent deux méthodes pour leur évaluation : la première, (1) la méthode « glass box » qui évalue techniquement une conversation en fonction de la grammaire, de la syntaxe, de la structure des phrases ainsi que la pertinence des réponses, et la deuxième, (2) la méthode « black box » qui mesure la satisfaction des utilisateurs (Morrissey & Kirakowski, 2013, p.89). Dans leur étude, plusieurs thèmes ou pistes d'amélioration reviennent de manière récurrente :

- Maintien du thème de la discussion
- Réponse à une question spécifique
- Répondre à des indices sociaux (*e.g.* être capable de répondre à des questions personnelles telles que « Comment allez-vous ? » ou « Parlez-moi de vous »)
- Utiliser un registre linguistique approprié
- Salutations et personnalité
- Donner des indices de conversation (*e.g.* afin d'obtenir une clarification dans les situations où le chatbot ne dispose pas d'assez d'informations pour pouvoir répondre de manière appropriée)
- Réponses inappropriées et contrôle des conséquences.

Kirakowski, O'Donnell & Yiu (2009) proposent alors un questionnaire basé sur l'évaluation de ces 7 thèmes, composé de 23 items pour lesquels l'utilisateur doit se positionner sur une échelle de fréquence en 5 points : « toujours », « souvent », « parfois », « rarement » et « jamais ». Une question ouverte, posée à la fin de l'évaluation, demande au participant « En quoi ce chatbot est le plus performant ? ». Cette question permet d'évaluer le caractère naturel du chatbot qui est défini par ces 4 caractéristiques :

- Rigueur : mesure les capacités grammaticales et syntaxiques. Le chatbot doit utiliser une grammaire et une orthographe correctes de façon constante et adopter un registre linguistique approprié avec l'utilisateur ;

- Conscience : mesure la manière dont le chatbot semble répondre au cours de la conversation et la pertinence de ses réponses. Le chatbot doit être capable de suivre le cours de la conversation, de maintenir le thème de la discussion, de repérer les indices appropriés, de poser des questions pertinentes et y répondre ;
- Manières : évalue les habitudes de conversation et la capacité de se comporter de manière polie. Le chatbot doit faire preuve de bonnes manières, montrer qu'il a de bonnes habitudes de conversation, qu'il peut contrôler ses propos si une conversation semble s'égarer, et qu'il a une personnalité appropriée afin de développer une relation de confiance avec l'utilisateur ;
- Originalité : mesure la capacité du chatbot d'être original et de prendre des initiatives au cours de la conversation. Le chatbot doit faire preuve d'originalité. Il doit proposer des informations intéressantes en lien avec le thème de la conversation et il doit pouvoir prendre des initiatives en suggérant par exemple de passer à d'autres thèmes de discussion en lien avec la conversation.

#### *2.4.2.5 Utilisabilité et utilité*

L'utilité et l'utilisabilité sont deux critères étroitement liés qui s'influencent l'un l'autre car chacun d'eux permet de déterminer la satisfaction de l'utilisateur et la fréquence d'utilisation de l'outil (Lund, 2001 ; Tricot, 2003). De plus, AbuShawar et al. (2016) considèrent que ces critères jouent également un rôle dans l'engagement et la confiance des utilisateurs ainsi que dans l'acceptabilité du système. Hobert (2019) constate néanmoins qu'il existe peu d'études qui analysent l'effet de l'acceptation de la technologie et l'utilisation effective du système. En effet, seules 25% des recherches évaluent l'utilisabilité ou la facilité d'utilisation des agents conversationnels pédagogiques (p.264) qui a tendance à être négligée au profit des mesures de performance (Lund, 2001, p.1).

L'utilité est définie selon Nogry & al. (2004) comme « l'adéquation entre les fonctions fournies par le système et celles nécessaires à l'utilisateur pour atteindre les objectifs » (p.267). Ainsi, évaluer l'utilité consiste à observer si les fonctionnalités du système permettent à l'utilisateur d'accomplir sa tâche.

Dans le contexte éducatif, évaluer l'utilité est nécessaire afin de déterminer dans quelle mesure l'objectif d'apprentissage a été atteint (Tricot, 2003, p.395), observer si les fonctionnalités du système permettent à l'utilisateur d'accomplir sa tâche (Tricot et al., 2003, p.395) et « se demander si le nouvel outil constitue un gain (en performance, en rapidité d'apprentissage, en satisfaction...) par rapport à ce qui existait avant son introduction » (Bétrancourt, 2007, p.81).

L'utilisabilité renvoie à l'adéquation de l'outil avec les besoins de l'utilisateur et correspond à la relation coût-effort que doit fournir l'utilisateur pour réaliser son but à l'aide du système ou autrement dit, à sa "maniabilité" (Tricot & al., 2003) (*e.g.* le temps d'apprentissage, l'efficience du système notamment concernant le nombre d'erreurs

rencontrées). En d'autres termes, la satisfaction de l'utilisateur et donc la facilité d'utilisation.

Mesurer l'utilisabilité permet d'analyser l'expérience et le comportement des utilisateurs ainsi que la « convivialité du produit » (Lund, 2001, p.1). En effet, si un utilisateur utilise un produit, c'est qu'il existe un besoin et qu'il est utile pour réaliser une tâche (Barcenilla & Bastien, 2009) car les fonctions fournies par le système sont en adéquation avec les fonctions nécessaires pour atteindre les objectifs (Nogry et al., 2004, p.39). Dans le cadre de la conception d'un outil, le fait d'améliorer sa facilité d'utilisation permet d'améliorer son utilité et vice versa (Lund, 2001, p.3).

Pour évaluer le critère d'utilisabilité, Brooke (1996) propose de le mesurer à l'aide de l'échelle d'utilisabilité d'un système : « System Utilisability Scale » (SUS). Ce questionnaire se veut rapide à compléter et facile à comprendre car il ne se compose que de 10 items(voir Annexe 7), ce qui permet une passation rapide. Chaque item est associé à une échelle en 5 niveaux allant de « Pas du tout d'accord » à « Tout à fait d'accord » et permet de discriminer un mauvais système d'un bon système. Cette conclusion est possible grâce au calcul du score qui se fait à partir de la moyenne des différents items. Ce score est compris entre 0 et 100. Le score moyen est de 68/100. À partir de 52 points sur 100, l'auteur considère que le système est acceptable et à partir du seuil de 86 points sur 100, il l'estime excellent.

Davis (1989), quant à lui, mesure l'utilisabilité et l'utilité en proposant un questionnaire comportant 12 items permettant d'objectiver ces deux variables, sur la base de son modèle d'acceptation des technologies (Technology Acceptance Model) (TAM), illustré en figure 3. Chaque item est évalué à l'aide d'une échelle de Likert en 7 points. Ces 12 items sont décomposés en 2 parties. Les 6 premiers items concernent l'utilité perçue et s'intéressent à la rapidité, la performance, la productivité, la facilité, l'efficacité et le transfert. Les 6 derniers items concernent la facilité d'utilisation perçue et portent sur l'apprentissage, la facilité d'utilisation, la clarté, la contrôlabilité, l'habileté et la flexibilité (voir Annexe 8).

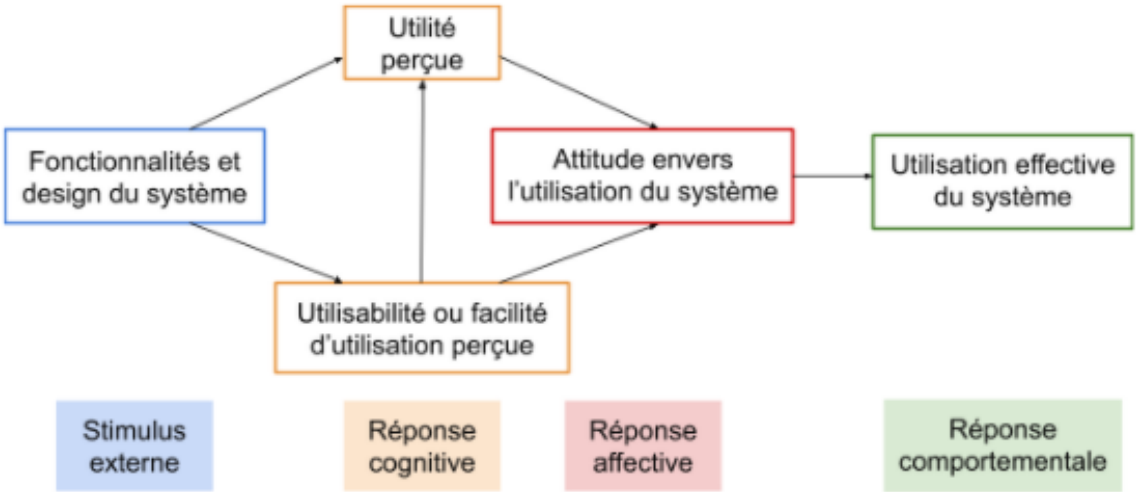


Fig. 3 - Technology Acceptance Model (Davis, 1989)

Plus récemment, Lund (2001) a développé le questionnaire USE (Usefulness, Satisfaction, and Ease of use) comportant 30 items. Pour chaque item, l'utilisateur doit se positionner sur une échelle de Likert en 7 points. Les sous-critères évalués par le questionnaire de Lund (2001) sont (1) l'utilité, (2) la facilité d'utilisation, (3) la facilité d'apprentissage et (4) la satisfaction (voir Annexe 9).

#### *2.4.2.6 Justesse technique : efficacité et performances*

Le critère de justesse technique vise à évaluer le chatbot pédagogique sur le plan technique en vérifiant ses fonctionnalités afin de valider si celui-ci répond de manière correcte à ce qui lui est demandé (Hobert, 2019).

Ces aspects sont liés à l'utilisabilité du système (Brangier & Barcenilla, 2003, p.43). En effet, l'efficacité du produit correspond à sa capacité à atteindre ou à réaliser le but pour lequel il a été conçu (Nogry et al., 2004). Dans la mesure où le système répond correctement et efficacement aux requêtes, il sera utilisé par les utilisateurs (Brangier & Barcenilla, 2003, p.76).

Ainsi, l'efficacité du chatbot peut être mesurée grâce à ses performances à répondre aux différents objectifs, comme par exemple sa capacité à contrôler ses réponses et à répondre de manière correcte à des données inconnues (Radziwill & Benton, 2017). Pour Brangier & Barcenilla (2003), « les deux mesures associées à l'efficacité sont la réussite de la tâche (partielle ou complète) et la qualité de la performance. Classiquement, deux mesures associées à l'efficacité sont distinguées : la réussite de la tâche, à savoir la capacité à atteindre minimalement, partiellement ou totalement les objectifs fixés ; et la qualité de la performance » (p.44).

#### *2.4.2.7 Autres facteurs psychologiques : émotions, sentiment de présence et engagement*

Hobert (2019) constate que d'autres facteurs psychologiques tels que le plaisir, les interactions sociales ou l'engagement sont également étudiés dans le cadre de la conception des chatbots pédagogiques.

Le design émotionnel est de plus en plus pris en compte dans les phases de conception en ergonomie des interfaces (Redet et al., 2016). L'intérêt de s'intéresser aux émotions réside dans le fait que certains comportements peuvent être utilisés en tant que critères d'acceptabilité des produits. Ainsi, pour prédire l'utilisation des outils, les émotions sont donc un facteur à prendre en compte parce qu'elles peuvent « améliorer l'apprentissage et la réalisation de différentes tâches » (Lockner et al., 2014). Lallemand (2015) constate également que les émotions ont un lien étroit avec l'expérience de l'utilisateur et l'utilisabilité car elles créent des comportements d'utilisation différents selon que le système génère des émotions positives ou négatives. Il apparaît donc nécessaire d'évaluer ces émotions ressenties lors de l'utilisation de l'outil. En effet, selon Dupré et al. (2018), si le produit suscite une

émotion, alors cela aura une influence sur la volonté de l'utilisateur à le réutiliser. Le concepteur a donc tout intérêt à examiner les caractéristiques de l'outil et les réactions qu'il génère auprès des utilisateurs afin de modifier les aspects qui engendrent des effets négatifs (Barcenilla & Bastien, 2009).

Les premiers modèles qui visent à analyser les impressions et le ressenti des utilisateurs s'intéressent de manière globale à leur satisfaction et « sont basées sur l'hypothèse selon laquelle les comportements des utilisateurs sont le résultat d'une équation entre l'utilité perçue, la facilité d'utilisation et la satisfaction procurée par l'utilisation ». Davis (1989) (TAM) et Venkatesh et al. (2003) (UTAUT) intègrent cette dimension dans leurs modèles d'évaluation mais cela reste implicite (Dupré et al., 2018, p.314).

Selon Ekman (1982) cité par Lallemand (2015), il existe 6 émotions de base comprenant 4 émotions négatives : (1) la peur, (2) la colère, (3) le dégoût et (4) la tristesse ; ainsi que 2 émotions positives : (5) la joie et (6) l'intérêt/la surprise. Il existe d'autres émotions qui, pour Plutchik (1980) cité par Lallemand (2015), sont un mélange de plusieurs émotions de base. Il illustre sa théorie dans « la roue des émotions » en montrant les différentes combinaisons et niveaux d'intensité des émotions ressenties.

Lallemand (2015) distingue 3 composantes permettant d'évaluer les émotions : « une composante physiologique, principalement liée à l'activation des systèmes endocriniens et nerveux autonomes ; une composante subjective, qui correspond à l'évaluation que le sujet fait de son propre état émotionnel ; et une composante expressive et comportementale, qui renvoie notamment aux expressions faciales, postures ou tonalité de la voix » (p.1). Pour évaluer chaque composante, il est nécessaire d'utiliser des outils et des méthodes distinctes qui sont synthétisés dans le tableau 5, ci-dessous.

Tableau 5 - Tableau récapitulatif des composantes des émotions et de leur évaluation selon Lallemand (2015)

| Composantes | Physiologique | Subjective | Expressive et comportementale |
|---|---|---|---|
| **Mesures** | ● Réponse électrodermale<br>● Fréquence cardiaque<br>● Fréquence respiratoire<br>● Électroencéphalographie | ● Mesure des dimensions émotionnelles (le plaisir, l'activation et la dominance)<br>● Mesure des émotions discrètes à l'aide de labels émotionnels, adjectifs ou phrases à l'aide d'échelles afin de situer le degré d'intensité ressenti | ● Facial Action Coding System<br>● Électromyographie |
| **Avantages** | ● Mesurer les émotions de manière non invasive et « objective » | ● Facilité d'analyse<br>● Accessibilité en fonction du public à évaluer | ● Permet l'identification des expressions faciales et des émotions de base |

| Composantes | Physiologique | Subjective | Expressive et comportementale |
|---|---|---|---|
| **Inconvénients** | • Interprétation complexe des données<br>• Induction de réactions émotionnelles parasites lors de l'évaluation | | • Logiciels peu fiables<br>• Difficultés d'analyse des émotions plus complexes |

Dans le cadre du questionnaire d'évaluation des chatbots assistants, nous allons évaluer la composante émotionnelle subjective car, ce type de mesure peut s'intégrer dans un questionnaire et, comme le mentionne Lallemand (2015), celle composante permet de recueillir plus facilement des données et simplifie également leur analyse.

Afin d'évaluer les émotions ressenties, Bradley & Lang (1994) proposent une échelle auto-rapportée : Self Assessment Manikin (SAM) (voir Annexe 10). Il est demandé à l'utilisateur de se positionner selon la valence de l'émotion ressentie (positive/négative) qui correspond au degré de satisfaction ; le niveau d'excitation (excité/calme) qui « fait référence au degré d'éveil du sujet » (Gil, 2009, p.16) et le niveau de dominance (contrôle, émotion dominante) qui indique la capacité du sujet à pouvoir influencer ou contrôler la situation.

Kay & Loverock (2008), quant à eux, proposent une échelle afin de mesurer les émotions associées à l'apprentissage de tâches informatiques. La Computer Emotion Scale se compose de 12 items classés selon 4 émotions : (1) la joie, (2) la tristesse, (3) l'anxiété et (4) la colère (voir Annexe 11).

Outre les aspects émotionnels, rappelons qu'un chatbot assistant de cours a pour objectif de jouer le rôle tuteur en offrant aux apprenants une certaine forme de présence et de soutien (Richardson & Swan, 2003). Dans la littérature, cette notion de présence est particulièrement présente dans le domaine de la formation à distance (Anderson et al., 2001 ; Fiock, 2020 ; Jézégou, 2010 ; Lowenthal & Lowenthal (2010)).

Dans le domaine de la formation, la présence « renvoie au rôle joué par le formateur dans les interactions sociales qu'il entretient à distance avec les apprenants pour soutenir la présence cognitive et la présence socio-affective » (Jézégou, 2010, p.264). Pour Dow et al. (2007), la présence correspond à un sentiment subjectif d'être connecté dans une expérience avec le numérique (p.1476). Ces auteurs relèvent que la présence a un effet positif sur la satisfaction des étudiants et des formateurs, de même que sur l'apprentissage. En effet, les étudiants dont la perception de la présence est élevée, s'engagent davantage dans leurs apprentissages.

Afin de mesurer le sentiment de présence dans le domaine des environnements virtuels, Witmer & Singer (1998) ont développé l'un des premiers questionnaires qui a été traduit et validé par le Laboratoire de Cyberpsychologie de l'Université du Québec

(Robillard et al., 2002) (voir Annexe 12). Ce questionnaire est composé de 5 sous-échelles :

- Réalisme : correspond à la capacité de l'interface d'être la plus naturelle possible ;
- Possibilité d'agir : correspond à la compétence de l'utilisateur à interagir avec l'environnement et à la possibilité de contrôler ses activités dans l'environnement virtuel ;
- Qualité de l'interface : correspond à la perception de son efficacité ;
- Possibilité d'examiner : correspond à la capacité de l'utilisateur d'être concentré dans l'environnement et sa capacité à réaliser des tâches ;
- Auto-évaluation de la performance : correspond au sentiment perçu par l'utilisateur concernant son expérience avec l'environnement ;
- Auditif : concerne les aspects sonores de l'interface ;
- Haptique : correspond au sens du toucher et à la possibilité d'examiner les objets visuellement dans l'environnement ou à les manipuler.

D'autre part, Dow et al. (2007) observent que la présence et l'engagement sont liés car plus un utilisateur aura un sentiment de présence ou d'immersion dans l'activité, plus il se sentira engagé (p.1476). Les auteurs décrivent l'engagement comme le degré d'implication ou d'intérêt d'une personne dans une activité. Pour O'Brien, Cairns & Hall (2018), ce critère d'engagement dans les technologies doit être pris en compte dans l'analyse et l'évaluation de la qualité de l'expérience utilisateur. Ils proposent alors l'échelle UES (User Engagement Scale) qui permet d'évaluer les attributs de l'engagement en contexte numérique (Fontaine et al., 2019, p.82). Fontaine et al. (2019) proposent une traduction française de cette échelle de mesure d'engagement. La version courte (short form) se compose de 12 items et comprend 4 sous-échelles (voir Annexe 13) :

- Attention soutenue : correspond au sentiment d'être absorbé dans la tâche et de perdre la notion du temps ;
- Convivialité perçue : correspond au ressenti suite à l'utilisation de l'outil et au degré d'effort nécessaire pour réaliser la tâche ;
- Attrait esthétique : correspond à l'attractivité et à l'effet visuel de l'interface ;
- Gratification : correspond au succès global de l'interaction avec l'outil, l'implication ressentie et la volonté des utilisateurs de l'utiliser à nouveau et/ou de le recommander à d'autres utilisateurs.

En conclusion, nous avons identifié 10 critères pertinents dans la littérature (expérience utilisateur ; acceptabilité ; motivation ; comportement-humanité ; utilisabilité ; utilité ; efficacité-performances ; émotions ; présence ; engagement). Au regard des modèles d'évaluation présentés par les différents auteurs, les items présents dans ces échelles vont nous permettre de concevoir un questionnaire permettant d'évaluer un chatbot créé dans le but d'assister les enseignants dans le cadre d'une formation à distance.

# 3 Objectif de la recherche

Dans le cadre de ce mémoire, nous avons pour objectif de concevoir un outil d'évaluation mesurant l'expérience des utilisateurs avec un chatbot utilisé comme assistant de cours dans une formation hybride ou à distance. En effet, les utilisateurs d'outils numériques ont des attentes, non seulement concernant leurs fonctionnalités, mais également leur attractivité. Par conséquent, les concepteurs d'outils numériques devraient identifier les caractéristiques qui facilitent les expériences positives des utilisateurs pour les appliquer à leur propre conception (O'Brien & Toms, 2010), d'où l'intérêt principal de les évaluer du point de vue des utilisateurs plutôt qu'uniquement d'un point de vue des performances (Peras, 2018).

# 4 Contexte de la recherche

Le présent mémoire est réalisé de manière conjointe avec Simon Jöhr qui a pour objectif de concevoir un chatbot assistant pour le cours ADID (Apprentissage Digital et formation à Distance) dispensé sous forme hybride dans le cadre du Master in Learning and Teaching Technology (MALTT) de la Faculté de Psychologie et des Sciences de l'Éducation de l'Université de Genève.

Ce chatbot sera une ressource pédagogique interactive destinée à accompagner les élèves et il a pour objectif de soutenir le processus d'engagement des étudiants dans leurs apprentissages, en leur apportant des réponses automatiques pour le cours via une application de messagerie instantanée disponible 7j/7 et 24h/24.

Ce tuteur virtuel récolte des données utiles pour les enseignants tels que leurs émotions et leurs ressentis concernant le cours. Les fonctionnalités de ce chatbot sont détaillées dans le mémoire de Simon Jöhr (en cours de publication). Notons que dans le cadre de son travail de mémoire, les fonctionnalités ont été limitées et scénarisées afin de pouvoir proposer un prototype du chatbot en cours de développement. Le robot comprend, entre autres, les fonctionnalités suivantes :

- Apparait automatiquement lorsque la page web du cours est consultée ;
- Demande et enregistre le prénom de l'utilisateur afin de pouvoir le réutiliser ensuite pour l'ensemble des conversations ;
- Propose des contenus multimédias divers (*e.g.* Images, GIF, etc.)
- Renvoie vers des pages du cours (*e.g.* le forum de discussion)
- Répond de manière automatique en fonction de règles prédéfinies, selon les mots déclencheurs utilisés par l'utilisateur ;
- Demande aux apprenants de situer leur état émotionnel sur une échelle afin de récolter leurs émotions et leur niveau de motivation.

Dans le cadre de notre recherche, nous allons utiliser le prototype de ce chatbot afin de nous permettre de tester notre questionnaire en contexte réel.

# 5 Méthodologie

Pour nous permettre d'évaluer l'expérience d'utilisation avec les chatbots assistants de cours, nous avons opté pour la création d'un questionnaire d'évaluation à l'attention des utilisateurs. Dans cette première phase de conception du questionnaire, nous nous sommes basés sur la littérature afin de déterminer les critères d'évaluation à prendre en compte, selon la méthodologie proposée par O'Brien & Toms (2010).

Dans la seconde phase, nous allons procéder à une double validation des critères et des items issus d'échelles de mesures déjà existantes, en les soumettant (1) aux experts et (2) aux utilisateurs, comme suggéré par Bouletreau et al. (1999). Le chapitre « Méthodologie » renseigne sur ce processus d'élaboration de notre questionnaire ainsi que sur la validation cognitive des critères et des items relevés dans la littérature.

Enfin, la dernière phase, consiste à tester notre questionnaire afin de s'assurer de sa validité qualitative (Bouletreau et al., 1999).

Le chapitre « Résultats » présente les items retenus à l'issue des pré-tests effectués et détaille les résultats de l'évaluation d'ADIDBot. Les données obtenues grâce au questionnaire créé dans le cadre de ce mémoire permettent d'analyser le ressenti des utilisateurs et fournissent au concepteur des pistes d'amélioration pour son chatbot afin qu'il répondre de la meilleure manière possible aux besoins des utilisateurs.

## 5.1 Une méthodologie inspirée par O'Brien & Toms (2010)

Pour rappel, l'objectif de cette recherche est d'élaborer un questionnaire auto-rapporté d'évaluation des chatbots utilisés en tant qu'assistants de cours dans des formations à distance. Pour ce faire, nous avons suivi les deux phases proposées par O'Brien & Toms (2010, p.53) et procédé à plusieurs validations de l'outil, comme le révèle la figure 4.

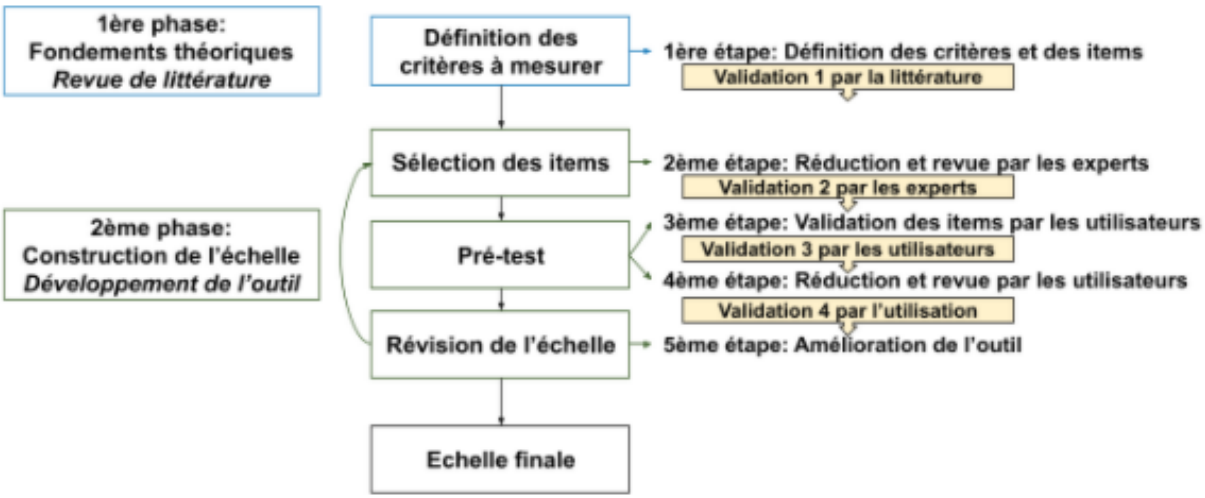


Fig. 4 - La méthodologie utilisée et adaptée selon O'Brien & Toms (2010)

Dans la première phase, nous avons défini les fondements théoriques, c'est-à-dire explicité les différentes dimensions à évaluer par le questionnaire en réalisant une revue de littérature. Dans la seconde phase et sur la base des critères retenus dans la littérature, nous avons développé une échelle qui permet d'évaluer le plus de dimensions possibles relatives à la qualité du chatbot. Pour y parvenir, nous avons suivi les trois étapes proposées par les auteurs. D'abord, les items pertinents ont été sélectionnés sur la base de l'analyse documentaire, en utilisant des échelles déjà existantes pour évaluer les dimensions souhaitées. Les critères et les items y afférents ont été soumis à une validation qualitative composée d'une vérification par les experts et un pré-test par les utilisateurs (Bouletreau et al., 1999). Enfin, les items retenus ont été utilisés pour jauger la qualité de l'outil dans le contexte étudié.

L'étape du pré-test proposée par O'Brien & Toms (2010) inclut deux validations, tel que suggéré par Bouletreau et al. (1999). Premièrement, afin de s'assurer de la pertinence des critères et des items évalués, une première révision est effectuée par les experts du domaine (Elangovan & Sundaravel, 2021). Deuxièmement, afin de valider leur compréhension, les items sont soumis aux utilisateurs pour évaluer la validité cognitive de chaque énoncé du questionnaire du point de vue de leur clarté (Karabenick et al., 2007). Ces deux étapes, permettent de pouvoir pallier au biais d'utilisation d'un questionnaire identifié par Lallemand & Gronier (2015) (éviter que la formulation des items soit mal interprétée).

Ainsi, le questionnaire passera au travers de plusieurs étapes de validation : (1) par la littérature, en utilisant des items issus d'échelles préexistantes et déjà validées, (2) par les experts du domaine pour s'assurer de la pertinence des critères et des items, et enfin (3) par les utilisateurs afin de valider la compréhension des items. Ces validations ont été réalisées de manière itérative, comme le révèle la figure 4, dans le but d'améliorer le questionnaire au fur et à mesure de sa conception. Une quatrième et dernière validation du questionnaire est prévue (4) à la fin du questionnaire d'évaluation et a pour objectif d'obtenir de nouvelles pistes d'amélioration de l'outil.

Dans les sections suivantes, nous allons détailler les différentes étapes réalisées afin de concevoir notre questionnaire ainsi que les items retenus pour aboutir à la version finale.

### *5.1.1 1ère étape : Définition des critères d'évaluation du Chatbot et des items y afférents*

Le questionnaire se compose de critères (dimensions d'évaluation) et d'items qui permettent de les mesurer.

Dans la première phase, nous avons exploré la littérature afin d'identifier les critères d'évaluation pertinents déjà observés par les chercheurs. Dans la seconde phase, ces différents critères ont été combinés et les items y relatifs ont été sélectionnés afin de

disposer d'un questionnaire qui soit le plus complet possible pour évaluer les chatbots assistants de cours.

Ainsi, nous avons identifié 10 critères d'évaluation pertinents (voir Tableau 6) qui peuvent s'appliquer, dans le domaine éducatif, aux outils numériques ainsi qu'aux chatbots : (1) Expérience utilisateur, (2) Acceptabilité, (3) Utilisabilité, (4) Utilité, (5) Motivation, (6) Comportement, (7) Émotions, (8) Efficacité et performances, (9) Engagement, et (10) Présence. Pour chaque critère, les items proposés par différents auteurs ont été sélectionnés comme étant pertinents pour l'évaluation des chatbots assistants. Ces items sont au nombre de 203 et se présentent sous trois formes. Premièrement, des items sous forme de questions ouvertes qui permettent aux utilisateurs de donner leur opinion. L'analyse de ce type d'items s'apparente à une analyse qualitative. Deuxièmement, des phrases affirmatives proposées aux utilisateurs. Ceux-ci doivent se positionner sur une échelle afin d'indiquer leur degré d'adhésion à l'énoncé. Enfin, des items sous forme de paires de mots. Les utilisateurs doivent se positionner sur un pôle afin d'indiquer leur ressenti. L'analyse de ces deux dernières formes d'items s'apparente à une analyse quantitative.

Tableau 6 - Critères, modèles et outils d'évaluation existants dans la littérature

| Critères | | Sous-critères | Modèles et outils d'évaluation |
|---|---|---|---|
| **1** | **Expérience utilisateur** | • Attractivité<br>• Confiance<br>• Crédibilité<br>• Impressions générales<br>• Originalité<br>• Qualité hédonique-identification<br>• Qualité hédonique-stimulation<br>• Qualité pragmatique<br>• Stimulation<br>• Transparence<br>• Valeur | • Extension of the User Experience Questionnaire (UEQ+) de Schrepp & Thomaschewski (2019) (voir Annexe 3)<br>• AttrakDiff 2 de Hassenzahl et al. (2003) (voir Annexe 2)<br>• EMPI (Évaluation des logiciels Multimédia Pédagogiques Interactifs) de Hu, Trigano & Crozat (2001) (voir Annexe 1) |
| **2** | **Acceptabilité** | • Anxiété<br>• Conditions de facilitation<br>• Influence sociale<br>• Injonction | • Unified Theory of Acceptance and Use of Technology (UTAUT) de Venkatesh et al. (2003) (voir Annexe 4)<br>• Modèle de l'Acceptation Instrumentale appliquée aux TICE de Caron & Heutte (2017) (voir Annexe 5) |
| **3** | **Utilisabilité** | • Efforts attendus<br>• Facilité d'apprentissage<br>• Facilité d'utilisation perçue<br>• Intention d'utiliser le système<br>• Maniabilité<br>• Performances attendues<br>• Satisfaction | • Unified Theory of Acceptance and Use of Technology (UTAUT) de Venkatesh et al. (2003) (voir Annexe 4)<br>• TAM (Technology Acceptance Model) de Davis (1989) (voir Annexe 8)<br>• System Utilisability Scale (SUS) de Brooke (1996) (voir Annexe 7)<br>• Modèle de l'Acceptation Instrumentale appliquée aux TICE de Caron & Heutte (2017) (voir Annexe 5) |

| Critères | | Sous-critères | Modèles et outils d'évaluation |
|---|---|---|---|
| | | • Utilisation intuitive | • USE (Usefulness, Satisfaction, and Ease of use) de Lund (2001) (voir Annexe 9)<br>• Extension of the User Experience Questionnaire (UEQ+) de Schrepp & Thomaschewski (2019) (voir Annexe 3) |
| **4** | **Utilité** | • Crédibilité<br>• Utilité apprenant<br>• Utilité générale<br>• Utilité perçue<br>• Utilité pour moi-même | • Extension of the User Experience Questionnaire (UEQ+) de Schrepp & Thomaschewski (2019) (voir Annexe 3)<br>• TAM (Technology Acceptance Model) de Davis (1989) (voir Annexe 8)<br>• Modèle de l'Acceptation Instrumentale appliquée aux TICE de Caron & Heutte (2017) (voir Annexe 5)<br>• USE (Usefulness, Satisfaction, and Ease of use) de Lund (2001) (voir Annexe 9) |
| **5** | **Motivation** | | • Échelle de motivation en éducation (EME) de Vallerand et al. (1989)<br>• Échelle de motivation en formation des adultes (EMFA) de Fenouillet, Heutte & Vallerand (2015) (voir Annexe 6)<br>• Modèle de Brandtzaeg et al. (2017) |
| **6** | **Comportement** | | • Modèle d'évaluation du caractère naturel de Morrissey & Kirakowski (2013) |
| **7** | **Émotions** | • Anxiété<br>• Colère<br>• Domination<br>• Joie<br>• Plaisir<br>• Stimulation<br>• Tristesse | • Self Assessment Manikin (SAM) de Bradley & Lang (1994) (voir Annexe 10)<br>• Computer Emotion Scale de Kay & Loverock (2008) (voir Annexe 11) |
| **8** | **Efficacité et performances** | | • Extension of the User Experience Questionnaire (UEQ+) de Schrepp & Thomaschewski (2019) (voir Annexe 3) |
| **9** | **Engagement** | • Attention soutenue<br>• Convivialité perçue<br>• Gratification | • UES (User Engagement Scale) de O'Brien, Cairns & Hall (2018) (voir Annexe 13) |
| **10** | **Présence** | • Réalisme | • Questionnaire sur l'état de présence de Witmer & Singer (1998) (voir Annexe 12) |

#### *5.1.1.1 Critère 1 : Expérience utilisateur*

Pour le critère général de l'expérience utilisateur qui mesure le ressenti de l'utilisateur lors de l'utilisation de l'outil, les items suivants, issus des modèles d'évaluation UEQ+ de Schrepp & Thomaschewski (2019), présent en annexe 3, AttrakDiff 2 de Hassenzahl et al. (2003), en annexe 2 et EMPI de Hû, Trigano & Crozat (2001), en annexe 1, ont été retenus :

Tableau 7 - Items retenus pour l'évaluation de l'expérience utilisateur

| Items | Modèle d'évaluation | Sous-critère |
|---|---|---|
| Rassurant / Déroutant | Hû, Trigano & Crozat (2001) | Impressions générales |
| Luxuriant / Dépouillé | Hû, Trigano & Crozat (2001) | Impressions générales |
| Ludique / Académique | Hû, Trigano & Crozat (2001) | Impressions générales |
| Actif / Passif | Hû, Trigano & Crozat (2001) | Impressions générales |
| Simple / Complexe | Hû, Trigano & Crozat (2001) | Impressions générales |
| Innovant / Traditionnel | Hû, Trigano & Crozat (2001) | Impressions générales |
| Technique / Humain | Hassenzahl et al. (2003) | Qualité pragmatique |
| Compliqué / Simple | Hassenzahl et al. (2003) | Qualité pragmatique |
| Pas pratique / Pratique | Hassenzahl et al. (2003) | Qualité pragmatique |
| Fastidieux / Efficace | Hassenzahl et al. (2003) | Qualité pragmatique |
| Imprévisible / Prévisible | Hassenzahl et al. (2003) | Qualité pragmatique |
| Confus / Clair | Hassenzahl et al. (2003) | Qualité pragmatique |
| Incontrôlable / Maîtrisable | Hassenzahl et al. (2003) | Qualité pragmatique |
| M'isole / Me sociabilise | Hassenzahl et al. (2003) | Qualité hédonique-identification |
| Amateur / Professionnel | Hassenzahl et al. (2003) | Qualité hédonique-identification |
| De mauvais goût / De bon goût | Hassenzahl et al. (2003) | Qualité hédonique-identification |
| Bas de gamme / Haut de gamme | Hassenzahl et al. (2003) | Qualité hédonique-identification |
| M'exclut / M'intègre | Hassenzahl et al. (2003) | Qualité hédonique-identification |
| Me sépare des autres / Me rapproche des autres | Hassenzahl et al. (2003) | Qualité hédonique-identification |
| Non présentable / Présentable | Hassenzahl et al. (2003) | Qualité hédonique-identification |
| Conventionnel / Original | Hassenzahl et al. (2003) | Qualité hédonique-stimulation |
| Sans imagination / Créatif | Hassenzahl et al. (2003) | Qualité hédonique-stimulation |
| Prudent / Audacieux | Hassenzahl et al. (2003) | Qualité hédonique-stimulation |
| Conservateur / Novateur | Hassenzahl et al. (2003) | Qualité hédonique- |

| Items | Modèle d'évaluation | Sous-critère |
|---|---|---|
| | | stimulation |
| Ennuyeux / Captivant | Hassenzahl et al. (2003) | Qualité hédonique-stimulation |
| Peu exigeant / Challenging | Hassenzahl et al. (2003) | Qualité hédonique-stimulation |
| Commun / Nouveau | Hassenzahl et al. (2003) | Qualité hédonique-stimulation |
| Déplaisant / Plaisant | Hassenzahl et al. (2003) | Attractivité |
| Laid / Beau | Hassenzahl et al. (2003) | Attractivité |
| Désagréable / Agréable | Hassenzahl et al. (2003) | Attractivité |
| Rebutant / Attirant | Hassenzahl et al. (2003) | Attractivité |
| Mauvais / Bon | Hassenzahl et al. (2003) | Attractivité |
| Repoussant / Attrayant | Hassenzahl et al. (2003) | Attractivité |
| Décourageant / Motivant | Hassenzahl et al. (2003) | Attractivité |
| Fâcheuse / Réjouissante | Schrepp & Thomaschewski (2019) | Attractivité |
| Mauvaise / Bonne | Schrepp & Thomaschewski (2019) | Attractivité |
| Désagréable / Agréable | Schrepp & Thomaschewski (2019) | Attractivité |
| Pas sympathique / Sympathique | Schrepp & Thomaschewski (2019) | Attractivité |
| Incompréhensible / Compréhensible | Schrepp & Thomaschewski (2019) | Transparence |
| Difficile à apprendre / Facile à apprendre | Schrepp & Thomaschewski (2019) | Transparence |
| Compliquée / Facile | Schrepp & Thomaschewski (2019) | Transparence |
| Confuse / Claire | Schrepp & Thomaschewski (2019) | Transparence |
| Pas intéressant / Intéressant | Schrepp & Thomaschewski (2019) | Stimulation |
| Ennuyeux / Passionnant | Schrepp & Thomaschewski (2019) | Stimulation |
| De mauvaise qualité / Précieux | Schrepp & Thomaschewski (2019) | Valeur |
| Soporifique / Stimulant | Schrepp & Thomaschewski (2019) | Stimulation |
| Dépourvue d'imagination / Créative | Schrepp & Thomaschewski (2019) | Originalité |
| Conventionnelle / Originale | Schrepp & Thomaschewski (2019) | Originalité |
| Conventionnelle / Inédite | Schrepp & Thomaschewski (2019) | Originalité |
| Conservatrice / Innovatrice | Schrepp & Thomaschewski (2019) | Originalité |

| Items | Modèle d'évaluation | Sous-critère |
|---|---|---|
| Peu sûr / Sûr | Schrepp & Thomaschewski (2019) | Confiance |
| Pas sérieux / Sérieux | Schrepp & Thomaschewski (2019) | Confiance |
| Non fiable / Fiable | Schrepp & Thomaschewski (2019) | Confiance |
| Non transparent / Transparent | Schrepp & Thomaschewski (2019) | Confiance |
| Non présentable / Présentable | Schrepp & Thomaschewski (2019) | Valeur |
| Pas de bon goût / De bon goût | Schrepp & Thomaschewski (2019) | Valeur |
| Non élégante / Élégantes | Schrepp & Thomaschewski (2019) | Valeur |
| Invraisemblables / Vraisemblables | Schrepp & Thomaschewski (2019) | Crédibilité |
| Pas sérieuses / Sérieuses | Schrepp & Thomaschewski (2019) | Crédibilité |
| Imprécises / Précises | Schrepp & Thomaschewski (2019) | Crédibilité |

#### *5.1.1.2 Critère 2 : Acceptabilité*

Pour le critère d'acceptabilité qui permet de mesurer si l'outil sera adopté ou non par les utilisateurs, nous avons retenu les items d'évaluation suivants, issus des modèles de l'UTAUT de Venkatesh et al. (2003), présent en annexe 4 et de l'Acceptation Instrumentale appliquée aux TICE de Caron & Heutte (2017), en annexe 5 :

Tableau 8 - Items retenus pour l'évaluation de l'acceptabilité

| Items | Modèle d'évaluation | Sous-critère |
|---|---|---|
| L'utilisation du chatbot est une bonne idée. | Venkatesh et al. (2003) | Influence sociale |
| Le chatbot rend l'apprentissage plus intéressant. | Venkatesh et al. (2003) | Influence sociale |
| Travailler avec le chatbot est amusant. | Venkatesh et al. (2003) | Influence sociale |
| J'aime utiliser le chatbot. | Venkatesh et al. (2003) | Influence sociale |
| Des gens qui influencent mon comportement pensent que je devrais utiliser le système. | Venkatesh et al. (2003) | Influence sociale |
| Les gens qui sont importants pour moi pensent que je devrais utiliser le chatbot. | Venkatesh et al. (2003) | Influence sociale |
| Une personne spécifique (ou un groupe) est disponible pour apporter de l'aide en cas de difficultés avec le chatbot. | Venkatesh et al. (2003) | Conditions de facilitation |
| En général, l'institution encourage l'utilisation du chatbot. | Venkatesh et al. (2003) | Conditions de facilitation |
| J'ai les ressources nécessaires pour utiliser le chatbot. | Venkatesh et al. (2003) | Conditions |

| Items | Modèle d'évaluation | Sous-critère |
|---|---|---|
| | | de facilitation |
| J'ai les connaissances nécessaires pour utiliser le chatbot. | Venkatesh et al. (2003) | Conditions de facilitation |
| Le chatbot n'est pas compatible avec d'autres systèmes que j'utilise. | Venkatesh et al. (2003) | Conditions de facilitation |
| Je pourrais réaliser une tâche en utilisant le chatbot... S'il n'y a personne autour de moi pour me dire quoi faire. | Venkatesh et al. (2003) | Anxiété |
| Si je peux appeler quelqu'un à l'aide si je suis bloqué. | Venkatesh et al. (2003) | Anxiété |
| Si j'ai beaucoup de temps à disposition pour réaliser la tâche à l'aide du chatbot. | Venkatesh et al. (2003) | Anxiété |
| Si je n'ai que l'assistance prédéfinie à disposition. | Venkatesh et al. (2003) | Anxiété |
| J'appréhende l'utilisation du chatbot. | Venkatesh et al. (2003) | Anxiété |
| J'ai peur de penser que je peux perdre un grand nombre d'informations en utilisant le système ou en cliquant sur un mauvais bouton. | Venkatesh et al. (2003) | Anxiété |
| J'hésite à utiliser le système par peur de faire des erreurs que je ne peux pas corriger. | Venkatesh et al. (2003) | Anxiété |
| L'utilisation du chatbot m'intimide un peu. | Venkatesh et al. (2003) | Anxiété |
| J'utilise le chatbot parce que l'on me le demande. | Caron & Heutte (2017) | Injonction |
| J'utilise le chatbot parce que les autres apprenants l'utilisent. | Caron & Heutte (2017) | Injonction |
| J'utilise le chatbot parce que j'espère de son usage une certaine reconnaissance au sein de mon établissement. | Caron & Heutte (2017) | Injonction |

#### *5.1.1.3 Critère 3 : Utilisabilité*

Pour le critère d'utilisabilité qui mesure la capacité de l'outil à atteindre ses objectifs, nous avons retenu les items d'évaluation suivants, issus des modèles de Venkatesh et al. (2003), présent en annexe 4, Davis (1989), en annexe 8, Brooke (1996), en annexe 7, Caron & Heutte (2017), en annexe 5, Lund (2001), en annexe 9 et Schrepp & Thomaschewski (2019), en annexe 3 :

Tableau 9 - Items retenus pour l'évaluation de l'utilisabilité

| Items | Modèle d'évaluation | Sous-critère |
|---|---|---|
| Je trouve l'utilisation du chatbot utile dans mon apprentissage. | Venkatesh et al. (2003) | Performances attendues |
| Utiliser le chatbot me permet d'accomplir les tâches plus rapidement. | Venkatesh et al. (2003) | Performances attendues |
| Utiliser le chatbot augmente ma productivité. | Venkatesh et al. (2003) | Performances attendues |

| Items | Modèle d'évaluation | Sous-critère |
|---|---|---|
| Mon interaction avec le chatbot est claire et compréhensible. | Venkatesh et al. (2003) | Efforts attendus |
| C'est facile pour moi de devenir expérimenté en utilisant le chatbot. | Venkatesh et al. (2003) | Efforts attendus |
| Je trouve que le chatbot est facile à utiliser. | Venkatesh et al. (2003) | Efforts attendus |
| Apprendre à utiliser le chatbot est facile pour moi. | Venkatesh et al. (2003) | Efforts attendus |
| J'ai l'intention d'utiliser le chatbot ces prochains mois. | Venkatesh et al. (2003) | Intention d'utiliser le système |
| Je pense que je vais utiliser le chatbot ces prochains mois. | Venkatesh et al. (2003) | Intention d'utiliser le système |
| Je planifie d'utiliser le chatbot ces prochains mois. | Venkatesh et al. (2003) | Intention d'utiliser le système |
| Apprendre à utiliser le chatbot sera facile pour moi. | Davis (1989) | Facilité d'utilisation perçue |
| Je trouve qu'il est facile de demander au chatbot de faire ce que je veux. | Davis (1989) | Facilité d'utilisation perçue |
| Mon interaction avec le chatbot est claire et compréhensible. | Davis (1989) | Facilité d'utilisation perçue |
| Je trouve le chatbot flexible dans nos interactions. | Davis (1989) | Facilité d'utilisation perçue |
| Il me sera facile de devenir habile dans l'utilisation du chatbot. | Davis (1989) | Facilité d'utilisation perçue |
| Je trouve que le chatbot est facile à utiliser. | Davis (1989) | Facilité d'utilisation perçue |
| Je voudrais utiliser ce système fréquemment. | Brooke (1996) | Utilisabilité |
| Ce système est inutilement complexe. | Brooke (1996) | Utilisabilité |
| Ce système est facile à utiliser. | Brooke (1996) | Utilisabilité |
| J'aurais besoin du soutien d'un technicien pour être capable d'utiliser ce système. | Brooke (1996) | Utilisabilité |
| Les différentes fonctionnalités de ce système sont bien intégrées. | Brooke (1996) | Utilisabilité |
| Il y a trop d'incohérences dans ce système. | Brooke (1996) | Utilisabilité |

| Items | Modèle d'évaluation | Sous-critère |
|---|---|---|
| La plupart des gens apprendront à utiliser ce système très rapidement. | Brooke (1996) | Utilisabilité |
| Ce système est très lourd à utiliser. | Brooke (1996) | Utilisabilité |
| Je me suis senti·e très en confiance en utilisant ce système. | Brooke (1996) | Utilisabilité |
| J'ai eu besoin d'apprendre beaucoup de choses avant de pouvoir utiliser ce système. | Brooke (1996) | Utilisabilité |
| C'est facile d'apprendre à utiliser le chatbot. | Caron & Heutte (2017) | Utilisabilité |
| C'est facile d'utiliser le chatbot pour faire ce que je veux faire. | Caron & Heutte (2017) | Utilisabilité |
| Je trouve le chatbot facile à utiliser. | Caron & Heutte (2017) | Utilisabilité |
| Le chatbot est facile à utiliser. | Lund (2001) | Utilisabilité |
| Le chatbot est simple à utiliser. | Lund (2001) | Utilisabilité |
| Le chatbot est convivial. | Lund (2001) | Utilisabilité |
| Le chatbot nécessite le moins d'étapes possibles pour accomplir ce que je veux faire avec lui. | Lund (2001) | Utilisabilité |
| Le chatbot est flexible. | Lund (2001) | Utilisabilité |
| Utiliser le chatbot ne nécessite pas d'effort. | Lund (2001) | Utilisabilité |
| Je peux utiliser le chatbot sans instruction écrite. | Lund (2001) | Utilisabilité |
| Je ne rencontre pas d'incohérence quand j'utilise le chatbot. | Lund (2001) | Utilisabilité |
| Le chatbot convient à la fois aux utilisateurs réguliers et aux utilisateurs occasionnels. | Lund (2001) | Utilisabilité |
| Je peux corriger mes erreurs rapidement et facilement. | Lund (2001) | Utilisabilité |
| Je peux utiliser le chatbot avec succès à chaque fois. | Lund (2001) | Utilisabilité |
| J'ai appris à utiliser le chatbot facilement. | Lund (2001) | Facilité d'apprentissage |
| Je me rappelle facilement comment utiliser le chatbot. | Lund (2001) | Facilité d'apprentissage |
| Il est facile d'apprendre à utiliser le chatbot. | Lund (2001) | Facilité d'apprentissage |
| Je suis rapidement devenu habile avec le chatbot. | Lund (2001) | Facilité d'apprentissage |
| Je suis satisfait du chatbot. | Lund (2001) | Satisfaction |
| Je recommanderais le chatbot à un ami. | Lund (2001) | Satisfaction |

| Items | Modèle d'évaluation | Sous-critère |
|---|---|---|
| Le chatbot est amusant à utiliser. | Lund (2001) | Satisfaction |
| Le chatbot fonctionne comme je veux qu'il fonctionne. | Lund (2001) | Satisfaction |
| Le chatbot est merveilleux. | Lund (2001) | Satisfaction |
| J'ai l'impression que j'ai besoin d'avoir un chatbot. | Lund (2001) | Satisfaction |
| Le chatbot est agréable à utiliser. | Lund (2001) | Satisfaction |
| Laborieuse / Aisée | Schrepp & Thomaschewski (2019) | Utilisation intuitive |
| Non logique / Logique | Schrepp & Thomaschewski (2019) | Utilisation intuitive |
| Pas évidente / Évidente | Schrepp & Thomaschewski (2019) | Utilisation intuitive |
| Incohérente / Cohérente | Schrepp & Thomaschewski (2019) | Utilisation intuitive |
| Imprévisible / Prévisible | Schrepp & Thomaschewski (2019) | Maniabilité |
| Gênante / Utile | Schrepp & Thomaschewski (2019) | Maniabilité |
| Peu sûre / Sûr | Schrepp & Thomaschewski (2019) | Maniabilité |
| Pas conforme aux attentes / Conforme aux attentes | Schrepp & Thomaschewski (2019) | Maniabilité |

#### *5.1.1.4 Critère 4 : Utilité*

Pour le critère d'utilité qui mesure si les fonctionnalités du système permettent à l'utilisateur d'accomplir sa tâche, nous avons retenu les items suivants, tirés des modèles de Schrepp & Thomaschewski (2019), présent en annexe 3, Davis (1989), en annexe 8, Caron & Heutte (2017), en annexe 5 et Lund (2001), en annexe 9 :

Tableau 10 - Items retenus pour l'évaluation de l'utilité

| Items | Modèle d'évaluation | Sous-critère |
|---|---|---|
| Non bénéfique / Bénéfique | Schrepp & Thomaschewski (2019) | Utilité et Crédibilité |
| Pas utile / Utile | Schrepp & Thomaschewski (2019) | Utilité |
| Non avantageuse / Avantageuse | Schrepp & Thomaschewski (2019) | Utilité |
| Non rentable / Rentable | Schrepp & Thomaschewski (2019) | Utilité |

| Items | Modèle d'évaluation | Sous-critère |
|---|---|---|
| Utiliser le chatbot me permettra d'accomplir les tâches plus rapidement. | Davis (1989) | Utilité perçue |
| L'utilisation du chatbot augmentera mes performances. | Davis (1989) | Utilité perçue |
| L'utilisation du chatbot augmentera ma productivité. | Davis (1989) | Utilité perçue |
| L'utilisation du chatbot améliorera mon efficacité. | Davis (1989) | Utilité perçue |
| Utiliser le chatbot faciliterait mon travail. | Davis (1989) | Utilité perçue |
| Je trouve que le chatbot est un outil utile dans mon travail. | Davis (1989) | Utilité perçue |
| L'utilisation du chatbot dans mes études me permet d'accomplir des tâches plus rapidement. | Caron & Heutte (2017) | Utilité générale |
| L'utilisation du chatbot me rend plus efficace. | Caron & Heutte (2017) | Utilité générale |
| Je trouve le chatbot utile pour mon projet d'étude. | Caron & Heutte (2017) | Utilité générale |
| Le chatbot me permet d'apprendre plus rapidement. | Caron & Heutte (2017) | Utilité apprenant |
| Avec le chatbot j'ai plus de facilité pour accéder à la formation reçue. | Caron & Heutte (2017) | Utilité apprenant |
| Le chatbot améliore ma capacité à appliquer la formation reçue. | Caron & Heutte (2017) | Utilité apprenant |
| Avec le chatbot c'est plus facile d'apprendre. | Caron & Heutte (2017) | Utilité apprenant |
| Le chatbot est utile pour mon apprentissage. | Caron & Heutte (2017) | Utilité apprenant |
| Avec le chatbot, je comprends mieux mon projet d'étude. | Caron & Heutte (2017) | Utilité pour moi-même |
| Le chatbot m'a permis de mieux mettre en œuvre mon projet d'étude. | Caron & Heutte (2017) | Utilité pour moi-même |
| Le chatbot m'aide à être plus efficace. | Lund (2001) | Utilité |
| Le chatbot m'aide à être plus productif | Lund (2001) | Utilité |
| Le chatbot est utile. | Lund (2001) | Utilité |
| Le chatbot me permet d'avoir plus de contrôle sur les activités de ma vie quotidienne. | Lund (2001) | Utilité |
| Le chatbot me permet de faire plus facilement ce que je veux faire. | Lund (2001) | Utilité |
| Le chatbot me fait gagner du temps quand je l'utilise. | Lund (2001) | Utilité |
| Le chatbot répond à mes besoins. | Lund (2001) | Utilité |

| Items | Modèle d'évaluation | Sous-critère |
|---|---|---|
| Le chatbot fait tout ce que je veux qu'il fasse. | Lund (2001) | Utilité |

#### *5.1.1.5 Critère 5 : Motivation*

Pour évaluer la motivation, les items quantitatifs de l'échelle de motivation en éducation proposés par Fenouillet, Heutte & Vallerand (2015) n'ont pas été retenus car ils sont trop centrés sur la formation et concernent finalement peu l'outil. Ces items permettront en revanche d'analyser les réponses qualitatives des utilisateurs aux questions ouvertes au regard des dimensions de l'échelle de motivation (EME) (voir Annexe 6).

Les items retenus pour le questionnaire ont été tirés de l'étude de Brandtzaeg et al. (2017) :

Tableau 11 - Items retenus pour évaluer la motivation

| Items |
|---|
| Qu'est-ce qui vous motive à utiliser un chatbot ? |
| Quelle est votre principale raison d'utiliser un chatbot ? |

#### *5.1.1.6 Critère 6 : Comportement*

Pour le critère de comportement, en lien avec le langage naturel, la personnalité et l'humanité du chatbot, nous avons retenu l'item suivant, tiré de l'étude de Kirakowski, O'Donnell & Yiu (2009) :

Tableau 12 - Item retenu pour l'évaluation du comportement et de l'humanité du chatbot

| Items |
|---|
| En quoi ce chatbot est le plus performant ? |

#### *5.1.1.7 Critère 7 : Émotions*

Pour évaluer les émotions, nous avons identifié les items suivants, issus des modèles de Bradley & Lang (1994) (voir Annexe 10) et Kay & Loverock (2008) (voir Annexe 11) :

Tableau 13 - Items retenus pour évaluer les émotions

| Items | Modèle d'évaluation | Sous-critère |
|---|---|---|
| Malheureux / Heureux | Bradley & Lang (1994) | Plaisir |
| Insatisfait / Satisfait | Bradley & Lang (1994) | Plaisir |
| Lenteur / Rapidité | Bradley & Lang (1994) | Stimulation |
| Guidé / Autonome | Bradley & Lang (1994) | Domination |

| Items | Modèle d'évaluation | Sous-critère |
|---|---|---|
| Satisfait | Kay & Loverock (2008) | Joie |
| Découragé | Kay & Loverock (2008) | Tristesse |
| Anxieux | Kay & Loverock (2008) | Anxiété |
| Irritable | Kay & Loverock (2008) | Colère |
| Excité | Kay & Loverock (2008) | Joie |
| Dépité | Kay & Loverock (2008) | Tristesse |
| Peu sûr de moi | Kay & Loverock (2008) | Anxiété |
| Frustré | Kay & Loverock (2008) | Colère |
| Curieux | Kay & Loverock (2008) | Joie |
| Impuissant | Kay & Loverock (2008) | Anxiété |
| Nerveux | Kay & Loverock (2008) | Anxiété |
| En colère | Kay & Loverock (2008) | Colère |

#### *5.1.1.8 Critère 8 : Efficacité et performances*

L'efficacité fait partie des critères évalués par l'UEQ+ de Schrepp & Thomaschewski (2019) (voir Annexe 3). Pour les critères d'efficacité et de performances, nous avons retenu les 4 items suivants :

Tableau 14 - Items retenus pour évaluer l'efficacité et les performances

| Items |
|---|
| Lent / Rapide |
| Inefficace / Efficace |
| Pas pratique / Pratique |
| Surchargé / Ordonné |

#### *5.1.1.9 Critère 9 : Engagement*

Pour évaluer l'engagement, nous avons retenu les items tirés de O'Brien, Cairns & Hall (2018) (voir Annexe 13) :

Tableau 15 - Items retenus pour l'évaluation de l'engagement

| Items | Sous-critère |
|---|---|
| Cette expérience m'a tellement absorbé que j'en ai perdu la notion du temps. | Attention soutenue |
| Le temps a filé lorsque j'ai utilisé le chatbot. | Attention soutenue |
| J'ai été absorbé par cette expérience. | Attention soutenue |
| J'ai éprouvé de la frustration lors de l'utilisation du chatbot. | Convivialité perçue |

| Items | Sous-critère |
|---|---|
| J'ai trouvé que le chatbot était difficile à utiliser. | Convivialité perçue |
| L'utilisation du chatbot était pénible. | Convivialité perçue |
| L'utilisation du chatbot en valait la peine. | Gratification |
| Mon expérience s'est avérée gratifiante | Gratification |
| J'ai été intéressé par cette expérience. | Gratification |

#### *5.1.1.10 Critère 10 : Présence*

Pour évaluer la présence, les items ci-dessous ont été retenus. Dans le modèle d'évaluation de la présence selon Witmer & Singer (1998) (voir Annexe 12), seuls les aspects génériques sur l'environnement présents dans ce questionnaire peuvent s'appliquer dans le cadre de l'évaluation des interactions avec un agent conversationnel. De ce fait, nous avons retenu uniquement ces 2 items :

Tableau 16 - Items retenus pour l'évaluation de la présence

| Items | Sous-critère |
|---|---|
| Dans quelle mesure vos interactions avec l'environnement vous semblaient-elles naturelles ? | Réalisme |
| Dans quelle mesure les expériences que vous avez vécues dans l'environnement virtuel ressemblaient-elles à celles de l'environnement réel ? | Réalisme |

### *5.1.2 2ème étape : Réduction et revue par les experts*

Afin d'assurer la validité du contenu du questionnaire, la liste des items sélectionnés et classés par critère a été soumise à l'appréciation d'un expert de chatbots éducatifs. Ce dernier a alors évalué, pour chaque item, la clarté (l'item est-il compréhensible ?) et la pertinence (l'item est-il représentatif du critère choisi et adapté pour évaluer un chatbot assistant de cours ?). En outre, les items présentés au sein d'un même critère ont été comparés entre eux, mais aussi, avec ceux d'autres critères d'évaluation proposés. À l'issue de cette analyse, l'expert a ainsi émis des suggestions afin de reformuler ou supprimer certains items d'une part, et de supprimer ou d'adapter certains critères d'autre part. Les items retenus ont ensuite été soumis à la 3ème étape.

### *5.1.3 3ème étape : Validation des items par les utilisateurs*

Après la revue et la réduction des items par l'expert, nous avons procédé à la validation de leur compréhension et de leur pertinence par des utilisateurs du chatbot. Pour ce faire, 8 étudiants du Master MALTT (Université de Genève) ont volontairement accepté de participer à notre recherche en prenant part à un entretien individuel. Le formulaire de consentement figure à l'annexe 18.

L'entretien s'est déroulé à distance, via Zoom, en 2 parties. Le guide d'entretien est présent en annexe 19. Dans la première, les utilisateurs ont été invités à prendre en main le prototype d'un chatbot nommé ADIDBot. Celui-ci, en phase de développement, sera utilisé dans le cadre du cours Apprentissage Digital et formation à Distance (ADID) dispensé de manière hybride, et sera une ressource pédagogique interactive destinée à soutenir le processus d'engagement des étudiants dans leurs apprentissages, en discutant de manière personnalisée avec eux afin de leur apporter des réponses automatiques pour le cours via une application de messagerie instantanée disponible 7j/7 et 24h/24. Le dialogue textuel de 2-3 minutes entre chaque étudiant et ADIDBot avait un double objectif. D'une part, savoir si le chatbot, en cours de développement, était fonctionnel. D'autre part, s'imprégner de ce qu'est un chatbot et savoir comment il fonctionne.

Dans la deuxième partie de l'entretien, les utilisateurs ont été invités à réaliser une activité de tri de cartes au cours de laquelle ils devaient classer les items retenus par l'expert (à l'étape 2) selon les trois catégories suivantes :

1. Pertinents pour l'évaluation du chatbot
2. Non pertinents pour l'évaluation du chatbot
3. Items non compréhensibles

Ces étudiants ont réalisé le tri de carte de manière autonome en exprimant à haute voix leurs pensées et les raisons pour lesquelles l'item était classé dans telle ou telle catégorie.

Selon Lallemand et al. (2016), le tri de cartes est une méthode qui permet aux participants de catégoriser l'information, mais également de définir si l'item a été compris et s'il a été assimilé ou non. Cette activité a l'avantage d'être ludique, participative, et permet une meilleure implication des utilisateurs. À travers ce tri, nous nous sommes intéressés aux représentations des participants à la recherche afin de pouvoir adapter le questionnaire en fonction de leur compréhension des items. Nous avons par ailleurs analysé les différences et les similitudes dans les catégorisations des différents répondants afin de vérifier leur homogénéité. Cependant, cette méthode peut être chronophage et il n'est pas toujours possible d'obtenir un résultat homogène et identique pour tous les participants.

Enfin, nous avons discuté des résultats de chaque tri de cartes en demandant aux participants s'ils voyaient d'autres items pertinents à ajouter pour évaluer le chatbot. Ces entrevues, d'une durée totale de 45 à 60 minutes, ont permis, d'une part, de valider la compréhension des items du point de vue des utilisateurs afin de s'assurer qu'il n'y ait pas d'ambiguïté pour que tout futur utilisateur du questionnaire comprenne et évalue la même chose ; et d'autre part, de recueillir des suggestions d'amélioration des items pour le questionnaire.

### 5.1.4 4ème étape : Réduction et revue par les utilisateurs puis construction du questionnaire

Afin de s'assurer de la compréhension des items par les utilisateurs, ceux-ci ont été évalués, tout comme pour les experts, du point de vue de leur clarté (l'item est-il compréhensible ?) et de leur pertinence (l'item est-il représentatif du critère choisi et adapté pour évaluer un chatbot assistant de cours ?). Suite au tri de cartes, les items ont été classés, par les utilisateurs, dans l'une des trois catégories précitées (Pertinents pour l'évaluation du chatbot, Non pertinents pour l'évaluation du chatbot et Items non compréhensibles). À l'issue de cette analyse, les critères jugés non pertinents et ceux non compréhensibles ont été supprimés, et les items validés ont été structurés de manière logique, afin de constituer la version finale du questionnaire d'évaluation.

Enfin, le questionnaire créé a été testé par les utilisateurs dans le cadre de l'évaluation d'ADIDBot. Cette dernière validation permettra de rendre compte de l'efficacité du questionnaire et de le réviser selon la 5ème étape prévue dans méthodologie adaptée de O'Brien & Toms (2010). Pour ce faire, les utilisateurs ont été invités à répondre à 4 questions, habituellement posées dans les questionnaires de satisfaction (la première en se positionnant sur une échelle de Likert en 5 points allant de « 1 - Ne permet pas du tout d'évaluer le chatbot » à « 5 - Évalue tout fait le chatbot », et les 3 suivantes sous forme de question ouverte) :

1. Selon vous, est-ce que le questionnaire auquel vous venez de répondre permet d'évaluer le chatbot de façon optimale ?
2. En quoi le questionnaire d'évaluation permet d'évaluer le chatbot ?
3. En quoi le questionnaire d'évaluation ne permet pas d'évaluer le chatbot ?
4. Quelles pourraient être les pistes d'amélioration du questionnaire d'évaluation ?

De plus, l'analyse des réponses des utilisateurs au questionnaire permettra à Simon Jöhr, le concepteur d'ADIDBot, d'obtenir d'éventuelles pistes d'amélioration pour son chatbot réalisé dans le cadre de son mémoire de master.

### 5.1.5 Test du questionnaire en contexte réel pour l'évaluation d'ADIDBot

#### 5.1.5.1 Collecte des données

Les apprenants ont été amenés à utiliser le chatbot en cours de développement durant la dernière période de l'année académique qui est constituée d'une semaine de cours en présence et de 4 semaines à distance, au terme desquelles un travail de groupe sommatif doit être rendu. Les utilisateurs ont été invités à utiliser ADIDBot chaque semaine. Ce chatbot, commandité par les enseignants du cours, permet de répondre aux demandes des étudiants et a pour tâche de s'intéresser aux ressentis des utilisateurs dans leurs travaux de groupe durant la semaine de cours en présence ainsi que durant toute la période à distance.

Au terme de cette période d'utilisation, le questionnaire réalisé dans le cadre de ce mémoire a été soumis aux étudiants afin de pouvoir le tester de manière concrète. Le questionnaire a été réalisé dans Qualtrics puis envoyé à l'ensemble des étudiants du cours (15 étudiants au total). Néanmoins, seuls 6 utilisateurs y ont répondu (voir Annexe 17). Ainsi, le questionnaire d'évaluation du chatbot a été administré à 6 utilisateurs. La population se compose de 4 femmes, 1 homme et 1 sans indication de genre. 4 participants sont âgés de 25 à 34 ans et 2 de 35 à 44 ans. Le temps de passation est compris entre 4 et 12 minutes. 5 participants ont répondu aux questions à la fin du questionnaire en lien avec son évaluation (étape 4) et 1 répondant n'a pas souhaité y répondre.

#### *5.1.5.2 Analyse des données*

Afin de pouvoir analyser de manière uniforme les questions, il est nécessaire de définir une codification commune pour l'ensemble des items du questionnaire. Nous allons donc nous baser sur l'échelle de Likert en 7 points allant de « 1 - Pas du tout d'accord » à « 7 - Tout à fait d'accord » car ce type d'échelle est souvent utilisée comme nous avons pu le constater dans les modèles issus de la littérature, par exemple dans l'AttrakDiff2 de Hassenzahl et al. (2003). Cette homogénéité dans la méthode permet d'établir une base de comparaison identique pour chaque critère malgré la différence de forme des items.

Pour permettre la comparaison entre les items constitués de phrases et ceux constitués de paires de mots, l'échelle de l'ensemble des items sera composée de 7 possibilités de positionnement de l'utilisateur. Pour les items sous forme de paires de mots, l'utilisateur doit se positionner sur un pôle. L'échelle se compose comme suit : Item A – 3 – 2 – 1 – 0 – 1 – 2 – 3 Item B (*e.g.* Pas pratique 3 – 2 – 1 – 0 – 1 – 2 – 3 Pratique). Plus l'utilisateur se place à une extrémité, plus la valeur augmente. Les valeurs ont donc été recodées afin d'attribuer également un score allant de 1 à 7. Le pôle évaluant l'outil de manière négative aura la valeur 1 et le pôle qui l'évalue de manière positive, la valeur 7. La codification se présente comme suit : Item A (négatif) – 1 – 2 – 3 – 4 – 5 – 6 – 7 Item B (positif) (*e.g.* Pas pratique 1 – 2 – 3 – 4 – 5 – 6 – 7 Pratique), la valeur neutre ayant pour valeur : 4.

Les questions ouvertes présentes pour le critère de la motivation génèrent des données de type qualitatives. Pour les analyser, il conviendra de situer les réponses de l'utilisateur en les comparant avec les énoncés présents dans l'échelle de motivation en formation des adultes (EMFA) de Fenouillet, Heutte & Vallerand (2015) afin de déterminer de quel type de motivation il s'agit (voir Annexe 6). Pour chaque type de motivation, une valeur a été attribuée afin de pouvoir les comparer aux données quantitatives (voir Tableau 17).

Tableau 17 - Échelle pour la codification des réponses aux items du critère de motivation

| Échelle de motivation | Valeur attribuée |
|---|---|
| Amotivation | 1 |
| Régulation externe | 2 |
| Régulation introjectée | 3 |
| Régulation identifiée | 4 |
| Intrinsèque connaissance | 5 |
| Intrinsèque accomplissement | 6 |
| Intrinsèque sensation | 7 |

Pour la question ouverte évaluant le critère du comportement, les réponses seront situées sur l'échelle de Kirakowski, O'Donnell & Yiu (2009) (Rigueur, Conscience, Manières, Originalité). Pour chaque type de comportement, une valeur a été attribuée afin de pouvoir les comparer aux données quantitatives. Les valeurs de l'échelle, au nombre de 4, étant toutes positives, leur contraire a été codifié afin de rendre compte du mauvais comportement (voir Tableau 18) et permettre ainsi la comparaison avec les autres items du questionnaire.

Tableau 18 - Échelle pour la codification des réponses aux items du critère de comportement

| Critères | Valeur attribuée |
|---|---|
| Manque d'originalité | 1 |
| Mauvaises manières | 2 |
| Absence de conscience | 3 |
| Rigueur | 4 |
| Conscience | 5 |
| Manières | 6 |
| Originalité | 7 |

Pour calculer le résultat du panel, il convient d'additionner le score obtenu de tous les items pour chaque critère d'évaluation afin d'obtenir son score global (Lallemand et al., 2015).

$$\textbf{Score global} = \Sigma\ \textbf{Valeur du critère évalué, de l'ensemble des participants}$$

Ensuite, il convient de calculer la moyenne en divisant le score global par le nombre d'items qui évalue ce critère d'évaluation.

$$\textbf{Moyenne} = \frac{\textbf{Score global}}{\textbf{Nombre d'items}}$$

Enfin, cette moyenne permet de comparer les résultats des différents participants et d'interpréter l'adhésion globale du groupe étudié. Pour cela, il est nécessaire de calculer le seuil minimal pour chaque critère. Le seuil minimal correspond à la valeur de l'item multipliée par le nombre de participants.

Cette manière de procéder permet de pouvoir comparer les différents résultats entre eux.

Le tableau ci-dessous illustre la manière d'interpréter les résultats du questionnaire. Le seuil est calculé sur la base d'une participation de 6 personnes. Dans le cas où le nombre de participants diffère, il conviendra alors de calculer le seuil sur la base de la formule suivante :

**Seuil = Valeur x Nombre de participants**

Tableau 19 - Échelle pour l'interprétation des résultats du questionnaire

| Valeur | Libellés pour les items sous forme de phrase | Libellés pour les items sous forme de paire de mots | Seuil minimal (pour 6 participants) |
|---|---|---|---|
| 1 | Pas du tout d'accord | +++ | 6 |
| 2 | En désaccord | ++ | 12 |
| 3 | Plutôt en désaccord | + | 18 |
| 4 | Ni en accord ni en désaccord | 0 | 24 |
| 5 | Plutôt d'accord | + | 30 |
| 6 | D'accord | ++ | 36 |
| 7 | Tout à fait d'accord | +++ | 42 |

# 6 Résultats

De la revue de littérature, nous avons retenu 203 items qui permettent d'évaluer les 10 critères d'évaluation (acceptabilité, efficacité-performances, émotions, engagement, expérience utilisateur, comportement, motivation, présence, utilité et utilisabilité) issus d'échelles préexistantes.

Nous avons procédé aux deux étapes pour évaluer qualitativement le questionnaire réalisé : (1) un jugement d'experts et (2) un pré-test auprès des utilisateurs (Bouletreau et al., 1999). La première étape a été effectuée par un expert de chatbot. Les items et les critères y relatifs lui ont été transmis par email. Ce dernier nous a fait un retour par écrit avec ses commentaires et a conservé 66 items, en excluant les items du critère d'engagement car ceux-ci étaient similaires à ceux déjà présents dans l'utilité et l'utilisabilité, réduisant ainsi le nombre de critères évalués à 9.

Une seconde étape a été réalisée avec les utilisateurs finaux du chatbot ADID durant des entretiens à l'aide de la méthodologie du tri de cartes. Les participants ont gardé 36 items.

À la suite de l'analyse des entretiens menés pour cette deuxième validation, certains items ont finalement été écartés à cause de la non-compréhension du lien entre l'item et le critère qu'il doit évaluer. D'autres items en lien avec la dimension éducative n'ont en revanche pas été supprimés car ce thème est peu abordé par les autres items conservés. Ainsi, la version finale du questionnaire se compose de 35 items et 9 critères d'évaluation.

Ce questionnaire a ensuite été utilisé pour évaluer ADIDBot. Les utilisateurs ont relevé une bonne expérience utilisateur de par la présence du chatbot, son utilisabilité et son efficacité. En revanche, son utilité, son acceptabilité, la motivation à l'utiliser et l'émotion qu'il génère sont les dimensions ayant un score le plus faible.

## 6.1 Des feedbacks des experts

L'expert a identifié plusieurs éléments dans la liste d'items proposés, notamment 1) la présence d'items similaires énoncés par les différents auteurs (par exemple, « Je trouve l'utilisation du chatbot utile dans mon apprentissage » et « Le chatbot est utile pour mon apprentissage » ; « Je trouve que le chatbot est facile à utiliser » et « Je trouve le chatbot facile à utiliser ») ; 2) Certains items étaient trop vagues (*e.g.* « L'utilisation du chatbot est une bonne idée ») ; 3) Certains items n'étaient pas suffisamment compréhensibles (*e.g.* « Pas évident – Évident »). Sur la base du constat de l'expert, la liste des items retenus a été modifiée :

1) parmi les items similaires, seul un a été conservé (*e.g.* « Simple – Complexe » a été supprimé et « Ce chatbot est facile à utiliser » a été conservé ; « Ce chatbot est inutilement complexe », « J'aurai besoin du soutien d'un technicien pour être capable d'utiliser ce chatbot », « La plupart des gens apprendront à

utiliser ce chatbot très rapidement » et « Ce chatbot est lourd à utiliser » ont été supprimés car le critère d'utilisabilité est évalué par l'item : « Ce chatbot est facile à utiliser » qui lui, a été gardé) ;

2) les items trop vagues ont été supprimés (*e.g.* « Luxuriant – Dépouillé ») ;
3) les items non compréhensibles ont été retirés (*e.g.* « Pas évident – Évident »).

Finalement, 66 items répartis en 9 critères ont été retenus par l'expert (voir Tableau 20). Les autres items de la revue de littérature ont été supprimés (voir Annexe 14). Les sous-critères sont également précisés dans le tableau 20 car ils permettent d'ajouter un niveau de détail supplémentaire pour l'analyse des données.

Tableau 20 - Liste des items conservés suite au feedback des experts

| Items | Critères | Sous-critères |
|---|---|---|
| Rassurant – Déroutant | Expérience utilisateur | Impressions générales |
| Ludique – Académique | Expérience utilisateur | Impressions générales |
| Actif – Passif | Expérience utilisateur | Impressions générales |
| Innovant – Traditionnel | Expérience utilisateur | Impressions générales |
| Dépourvu d'imagination – Créatif | Expérience utilisateur | Qualité hédonique-stimulation |
| Désagréable – Agréable | Expérience utilisateur | Attractivité |
| Fâcheux – Réjouissant | Expérience utilisateur | Attractivité |
| Mauvaise – Bonne | Expérience utilisateur | Attractivité |
| Pas sympathique – Sympathique | Expérience utilisateur | Attractivité |
| Incompréhensible – Compréhensible | Expérience utilisateur | Transparence |
| Pas intéressant – Intéressant | Expérience utilisateur | Stimulation |
| Ennuyeux – Passionnant | Expérience utilisateur | Stimulation |
| De mauvaise qualité – De bonne qualité | Expérience utilisateur | Valeur |
| Soporifique – Stimulant | Expérience utilisateur | Stimulation |
| Dépourvu d'imagination – Créatif | Expérience utilisateur | Originalité |
| Conventionnel – Original | Expérience utilisateur | Originalité |
| Conservateur – Innovateur | Expérience utilisateur | Originalité |
| Pas sérieux – Sérieux | Expérience utilisateur | Confiance |
| Non fiable – Fiable | Expérience utilisateur | Confiance |
| Imprécis – Précis | Expérience utilisateur | Crédibilité |
| Le chatbot rend l'apprentissage plus intéressant. | Acceptabilité | Influence sociale |
| Travailler avec le chatbot est amusant. | Acceptabilité | Influence sociale |
| J'aime utiliser le chatbot. | Acceptabilité | Influence sociale |
| J'appréhende l'utilisation du chatbot. | Acceptabilité | Anxiété |
| J'utilise le chatbot parce que l'on me le demande. | Acceptabilité | Injonction |
| J'utilise le chatbot parce que les autres | Acceptabilité | Injonction |

| Items | Critères | Sous-critères |
|---|---|---|
| apprenants l'utilisent. | | |
| Je trouve l'utilisation du chatbot utile pour mon apprentissage. | Utilisabilité | Performances attendues |
| Utiliser le chatbot me permet d'accomplir les tâches plus rapidement. | Utilisabilité | Performances attendues |
| Mon interaction avec le chatbot est claire et compréhensible. | Utilisabilité | Efforts attendus |
| J'ai l'intention d'utiliser le chatbot ces prochains mois. | Utilisabilité | Intention d'utiliser le système |
| Apprendre à utiliser le chatbot sera facile pour moi. | Utilisabilité | Facilité d'utilisation perçue |
| Je voudrais utiliser ce système fréquemment. | Utilisabilité | Utilisabilité |
| Ce chatbot est facile à utiliser. | Utilisabilité | Utilisabilité |
| Les différentes fonctionnalités de ce système sont bien intégrées. | Utilisabilité | Utilisabilité |
| Il y a trop d'incohérences dans ce chatbot. | Utilisabilité | Utilisabilité |
| Je me suis senti·e très en confiance en utilisant ce chatbot. | Utilisabilité | Utilisabilité |
| Je trouve le chatbot facile à utiliser. | Utilisabilité | Utilisabilité |
| Le chatbot permet de réaliser des tâches en peu d'étapes. | Utilisabilité | Utilisabilité |
| Je suis satisfait du chatbot. | Utilisabilité | Satisfaction |
| Je recommanderais le chatbot à un ami. | Utilisabilité | Satisfaction |
| Le chatbot est agréable à utiliser. | Utilisabilité | Satisfaction |
| Laborieuse – Aisée | Utilisabilité | Utilisation intuitive |
| Non logique – Logique | Utilisabilité | Utilisation intuitive |
| Imprévisible – Prévisible | Utilisabilité | Maniabilité |
| Non bénéfique – Bénéfique | Utilité | Utilité et Crédibilité |
| L'utilisation du chatbot me rend plus efficace. | Utilité | Utilité générale |
| Avec le chatbot c'est plus facile d'apprendre. | Utilité | Utilité apprenant |
| Le chatbot est utile. | Utilité | Utilité |
| Le chatbot me permet d'accomplir les tâches plus facilement. | Utilité | Utilité |
| Le chatbot me fait gagner du temps. | Utilité | Utilité |
| Le chatbot correspond à ce que j'ai besoin. | Utilité | Utilité |
| Le chatbot fait tout ce que j'attends de lui. | Utilité | Utilité |
| Qu'est-ce qui vous motive à utiliser un chatbot ? | Motivation | Motivation |
| Quelle est votre principale raison d'utiliser un chatbot ? | Motivation | Motivation |

| Items | Critères | Sous-critères |
|---|---|---|
| En quoi ce chatbot est le plus performant ? | Comportement | Humanité |
| Déçu.e – Stimulé.e | Émotions | Plaisir |
| Insatisfait – Satisfait | Émotions | Plaisir |
| Guidé – Autonome | Émotions | Dominance |
| Impatient.e – Calme | Émotions | Joie |
| Stressé.e – Serein.e | Émotions | Joie |
| Lent – Rapide | Efficacité et performances | Efficacité |
| Inefficace – Efficace | Efficacité et performances | Efficacité |
| Pas pratique – Pratique | Efficacité et performances | Efficacité |
| Surchargé – Ordonné | Efficacité et performances | Efficacité |
| Naturel – Robotisé | Présence | Réalisme |
| Environnement réel – Environnement fictif | Présence | Réalisme |

## 6.2 Des feedbacks des entretiens couplés au tri de cartes

Les items qui ont été retenus (voir Tableau 20) ont été cognitivement pré-testés par des utilisateurs du chatbot via des entretiens couplés à un tri de cartes dont la retranscription figure en annexe 20. Les feedbacks des huit participants aux entretiens de validation du questionnaire ont permis de révéler que certains items n'étaient pas pertinents dans l'évaluation du chatbot ou encore que certains items n'étaient pas compréhensibles pour les utilisateurs. Par exemple, l'item « J'utilise le chatbot parce que les autres apprenants l'utilisent » a été classé dans la catégorie « Non pertinent » par 7 utilisateurs sur 8. Autre exemple, l'item « J'utilise le chatbot parce qu'on me le demande » a été jugé incompréhensible par 2 utilisateurs et non pertinent pour les 6 autres utilisateurs. Sur la base du classement des items des utilisateurs, les items ayant été désignés comme étant utiles pour au moins 6 utilisateurs ont été conservés. En revanche, les autres items ont été supprimés (Voir Annexes 15 et 16).

Suite au pré-test effectué par les utilisateurs, 36 items (18 items sous forme de phrases et 18 items sous forme de paires de mots) ont été retenus (voir Tableau 21).

Tableau 21 - Items retenus par les utilisateurs

| Items | Critères | Sous-critères |
|---|---|---|
| Travailler avec le chatbot est amusant. | Acceptabilité | Influence sociale |
| J'aime utiliser le chatbot. | Acceptabilité | Influence sociale |
| Lent – Rapide | Efficacité et performances | Efficacité |
| Inefficace – Efficace | Efficacité et performances | Efficacité |

| Items | Critères | Sous-critères |
|---|---|---|
| Pas pratique – Pratique | Efficacité et performances | Efficacité |
| Surchargé – Ordonné | Efficacité et performances | Efficacité |
| Insatisfait – Satisfait | Émotions | Plaisir |
| Rassurant – Déroutant | Expérience utilisateur | Impressions générales |
| Ludique – Académique | Expérience utilisateur | Impressions générales |
| Désagréable – Agréable | Expérience utilisateur | Attractivité |
| Pas sympathique – Sympathique | Expérience utilisateur | Attractivité |
| Incompréhensible – Compréhensible | Expérience utilisateur | Transparence |
| Pas intéressant – Intéressant | Expérience utilisateur | Stimulation |
| De mauvaise qualité – De bonne qualité | Expérience utilisateur | Valeur |
| Non fiable – Fiable | Expérience utilisateur | Confiance |
| Imprécis – Précis | Expérience utilisateur | Crédibilité |
| Naturel – Robotisé | Présence | Réalisme |
| Je trouve l'utilisation du chatbot utile pour mon apprentissage. | Utilisabilité | Performances attendues |
| Utiliser le chatbot me permet d'accomplir les tâches plus rapidement. | Utilisabilité | Performances attendues |
| Mon interaction avec le chatbot est claire et compréhensible. | Utilisabilité | Efforts attendus |
| J'ai l'intention d'utiliser le chatbot ces prochains mois. | Utilisabilité | Intention d'utiliser le système |
| Ce chatbot est facile à utiliser. | Utilisabilité | Utilisabilité |
| Les différentes fonctionnalités de ce système sont bien intégrées. | Utilisabilité | Utilisabilité |
| Je trouve le chatbot facile à utiliser. | Utilisabilité | Utilisabilité |
| Je suis satisfait du chatbot. | Utilisabilité | Satisfaction |
| Le chatbot est agréable à utiliser. | Utilisabilité | Satisfaction |
| Laborieuse – Aisée | Utilisabilité | Utilisation intuitive |
| Non logique – Logique | Utilisabilité | Utilisation intuitive |
| Le chatbot permet de réaliser des tâches en peu d'étapes. | Utilisabilité | Utilisabilité |
| Non bénéfique – Bénéfique | Utilité | Utilité et Crédibilité |
| L'utilisation du chatbot me rend plus efficace. | Utilité | Utilité générale |
| Avec le chatbot c'est plus facile d'apprendre. | Utilité | Utilité apprenant |
| Le chatbot est utile. | Utilité | Utilité |
| Le chatbot me permet d'accomplir les tâches plus facilement. | Utilité | Utilité |
| Le chatbot me fait gagner du temps. | Utilité | Utilité |
| Le chatbot correspond à ce que j'ai besoin. | Utilité | Utilité |

## 6.3 De la version finale du questionnaire

Après analyse des entretiens, certains items ont finalement été écartés malgré l'utilité perçue par au moins 6 utilisateurs. Ces items sont :

- « Surchargé – Ordonné »
- « Laborieux – Aisé »
- « Les différentes fonctionnalités de ce chatbot sont bien intégrées »
- « Rassurant – Déroutant »

En effet, pour les utilisateurs, « Surchargé et Ordonné » se réfère à l'aspect esthétique de l'interface. Ce n'est pas la dimension que nous souhaitions évaluer (N.B : les participants ont été anonymisés selon les lettres de l'alphabet grec).

> *Zeta : « Surchargé - ordonné ; on est plus sur l'interface, l'aspect de l'interface ».*
>
> *Beta : « Surchargé - ordonné ; oui. Bah par exemple, à la fin, quand il m'a donné plein de messages, c'était un peu surchargé ».*

L'item « Laborieux – Aisé » a été retiré car il ne semble pas avoir été bien compris par deux des utilisateurs.

> *Beta : « Laborieux - aisé ; du coup ça, ça serait la facilité ou pas ? Laborieux – aisé … aisé […] Je pense que j'ai pas compris les termes, du coup… »*

L'item « Les différentes fonctionnalités de ce chatbot sont bien intégrées » a été écarté car en réalité, dans le chatbot évalué, il n'y avait pas de fonctionnalités intégrées à proprement parler. De plus, pour deux utilisateurs, cet item a été jugé comme étant non compréhensible.

Enfin, l'item « Rassurant – Déroutant » n'a pas été sélectionné car il semble que les utilisateurs n'utilisent pas cet item pour évaluer le critère souhaité.

> *Beta : « Rassurant - déroutant ; Ouais, ça, ça va avec le caractère du robot ».*

De plus, il est jugé non prioritaire par rapport aux autres items :

> *Zeta : « Rassurant - déroutant ; ça ça touche plus […] les émotions donc j'crois pas que c'est prioritaire ».*

En outre, l'item « Le chatbot rend l'apprentissage plus intéressant » a été conservé malgré le fait que ce critère ait été jugé utile pour uniquement 5 participants, car la dimension éducative n'est que trop peu abordée dans les autres items.

Au final, le questionnaire comporte au total 35 items (voir Tableau 22) répartis comme suit :

- 3 questions ouvertes ;
- 17 items sous forme de phrases ;
- 15 items sous forme de paires de mots.

Tableau 22 - Questionnaire final

| Consigne | Item | Critère évalué | Échelle |
|---|---|---|---|
| Question : | Qu'est-ce qui vous motive à utiliser un chatbot ? | Motivation | Analyse sur la base de l'échelle de motivation de Fenouillet, Heutte & Vallerand (2015) |
| Question : | Quelle est votre principale raison d'utiliser un chatbot ? | Motivation | |
| Sur une échelle de 1 (Pas du tout d'accord) à 5 (Tout à fait d'accord), dans quelle mesure êtes-vous d'accord avec les affirmations suivantes ? | Le chatbot est utile. | Utilité | Pas du tout d'accord (1)<br>En désaccord (2)<br>Plutôt en désaccord (3)<br>Ni d'accord, ni en désaccord (4)<br>Plutôt d'accord (5)<br>D'accord (6)<br>Tout à fait d'accord (7) |
| | Ce chatbot est facile à utiliser. | Utilisabilité | |
| | J'aime utiliser le chatbot. | Acceptabilité | |
| Pour chaque paire de mot, placez une croix dans la case qui se rapproche le plus de votre ressenti par rapport au chatbot. | Pas pratique - Pratique | Efficacité | Échelle : A 3-2-1-0-1-2-3 B<br>Codification :<br>... (1)<br>.. (2)<br>. (3)<br>0 (4)<br>. (5)<br>.. (6)<br>... (7) |
| | De mauvaise qualité - De bonne qualité | Efficacité | |
| | Lent - Rapide | Efficacité | |
| | Inefficace - Efficace | Efficacité | |
| Concernant l'utilisation de mes informations et données personnelles, le chatbot est... | Non fiable - Fiable | Expérience utilisateur (Confiance) | |
| Dans quelle mesure vos interactions avec le chatbot vous semblaient-elles naturelles ? | Naturelles - Robotisées | Présence | |
| Les informations et données que le chatbot met à ma disposition sont... | Non bénéfiques - Bénéfiques | Expérience utilisateur (Crédibilité) | |
| | Imprécises - Précises | Expérience utilisateur (Crédibilité) | |
| | Non logique - Logique | Utilisabilité | |
| | Incompréhensible - Compréhensible | Expérience utilisateur (Transparence) | |
| | Pas intéressant - Intéressant | Expérience utilisateur (Stimulation) | |
| Question : | En quoi ce chatbot est-il le plus performant ? | Comportement | Analyse sur la base de l'échelle de comportement de Kirakowski, O'Donnell & Yiu (2009) |

| Consigne | Item | Critère évalué | Échelle |
|---|---|---|---|
| Sur une échelle de 1 (Pas du tout d'accord) à 5 (Tout à fait d'accord), dans quelle mesure êtes-vous d'accord avec les affirmations suivantes ? | Utiliser le chatbot me permet d'accomplir les tâches plus rapidement. | Utilisabilité | Pas du tout d'accord (1)<br>En désaccord (2)<br>Plutôt en désaccord (3)<br>Ni d'accord, ni en désaccord (4)<br>Plutôt d'accord (5)<br>D'accord (6)<br>Tout à fait d'accord (7) |
| | Le chatbot me permet d'accomplir les tâches plus facilement. | Utilité | |
| | Le chatbot permet de réaliser des tâches en peu d'étapes. | Utilisabilité | |
| | L'utilisation du chatbot me rend plus efficace. | Utilité | |
| | Je trouve le chatbot utile pour mon apprentissage. | Utilisabilité | |
| | Avec le chatbot, c'est plus facile d'apprendre. | Utilité | |
| Pour chaque paire de mot, placez une croix dans la case qui se rapproche le plus de votre ressenti par rapport au chatbot.<br>Échelle : A 3-2-1-0-1-2-3 B | Ludique - Académique | Impressions générales (Expérience utilisateur) | Échelle : A 3-2-1-0-1-2-3 B<br>Codification :<br>... (1)<br>.. (2)<br>. (3)<br>0 (4)<br>. (5)<br>.. (6)<br>... (7) |
| | Satisfait.e - Insatisfait.e | Émotions | |
| | Désagréable - Agréable | Expérience utilisateur (Attractivité) | |
| | Pas sympathique - Sympathique | Expérience utilisateur (Attractivité) | |
| Sur une échelle de 1 (Pas du tout d'accord) à 5 (Tout à fait d'accord), dans quelle mesure êtes-vous d'accord avec les affirmations suivantes ? | Travailler avec le chatbot est amusant. | Acceptabilité | Pas du tout d'accord (1)<br>En désaccord (2)<br>Plutôt en désaccord (3)<br>Ni d'accord, ni en désaccord (4)<br>Plutôt d'accord (5)<br>D'accord (6)<br>Tout à fait d'accord (7) |
| Sur une échelle de 1 (Pas du tout d'accord) à 5 (Tout à fait d'accord), dans quelle mesure êtes-vous d'accord avec les affirmations suivantes ?<br>De manière générale, … | Le chatbot est plaisant à utiliser. | Utilisabilité | |
| | Je trouve le chatbot facile à utiliser. | Utilisabilité | |
| | Je suis satisfait.e du chatbot. | Utilisabilité | |
| | Le chatbot correspond à ce que j'ai besoin. | Utilité | |
| | Mon interaction avec le chatbot est claire et compréhensible. | Utilisabilité | |
| | Le chatbot rend l'apprentissage plus intéressant. | Acceptabilité | |
| | J'ai l'intention d'utiliser le chatbot ces prochains mois. | Utilisabilité | |

Afin de procéder à la dernière validation du questionnaire réalisé, celui-ci a été utilisé afin d'évaluer ADIDBot, un chatbot en cours de conception pour le cours ADID dispensé sous forme hybride dans le cadre du Master MALTT à l'Université de Genève. Les réponses des utilisateurs auront une double utilité : valider et améliorer le questionnaire réalisé, et fournir au concepteur d'ADIDBot des pistes d'amélioration pour son outil.

## 6.4 De l'évaluation du questionnaire

Grâce aux quatre questions posées aux utilisateurs en fin de questionnaire, nous avons pu obtenir des données concernant son efficacité. La première question avait pour objectif de mesurer si le questionnaire réalisé permettait d'évaluer les différentes dimensions du chatbot (« Selon vous, est-ce que le questionnaire auquel vous venez de répondre permet d'évaluer le chatbot de façon optimale ? »). De leurs réponses (voir Tableau 23), il en ressort que l'avis des utilisateurs est mitigé. En effet, pour 3 participants, le questionnaire permet d'évaluer le chatbot de façon optimale (2 évaluent bien et 1 évalue tout à fait), tandis que pour les 3 autres participants, le questionnaire ne le permet pas (2 pas tout à fait et 1 pas du tout).

Tableau 23 – Retour des utilisateurs sur le questionnaire

| **Participant n°** | *Selon vous, est-ce que le questionnaire auquel vous venez de répondre permet d'évaluer le chatbot de façon optimale?* | *En quoi le questionnaire d'évaluation permet d'évaluer le chatbot?* | *En quoi le questionnaire d'évaluation NE permet PAS d'évaluer le chatbot?* | *Quelles pourraient être les pistes d'amélioration du questionnaire d'évaluation?* |
|---|---|---|---|---|
| **1** | Évalue bien le chatbot | Pas de réponse | Pas de réponse | Pas de réponse |
| **2** | Ne permet pas tout à fait d'évaluer le chatbot | Sur les points d'ergonomie et de logique des réponses. | Sur l'utilité du chatbot. Nous l'avons utilisé pour le tester pas par besoin. Et ses réponses ne sont pas très utiles en l'état actuel pour nos études. | Éviter les questions sur l'utilité du chatbot. |
| **3** | Évalue tout à fait le chatbot | Le questionnaire est complet, il reprend tous les items pertinents pour l'évaluation du chatbot. | je trouve qu'il évalue bien. Je ne vois pas ce qu'il manque. | aucune. il me semble complet |
| **4** | Ne permet pas tout à fait d'évaluer le chatbot | Sur les aspects ergonomiques | En partie sur le contenu, c'est-à-dire qu'une partie des questions ne sont pas adaptées à son contenu actuel (plutôt centré sur l'évaluation | Les 2 premières questions se ressemblent (ambiguïté). |

| Participant n° | *Selon vous, est-ce que le questionnaire auquel vous venez de répondre permet d'évaluer le chatbot de façon optimale?* | *En quoi le questionnaire d'évaluation permet d'évaluer le chatbot?* | *En quoi le questionnaire d'évaluation NE permet PAS d'évaluer le chatbot?* | *Quelles pourraient être les pistes d'amélioration du questionnaire d'évaluation?* |
|---|---|---|---|---|
| | | | de la satisfaction et des inquiétudes vis-à-vis du cours ADID) mais pourront prendre + de sens dans le futur avec l'ajout de contenus relatifs à de l'appui pédagogique | |
| **5** | Évalue bien le chatbot | Les différents aspects visés par le questionnaire permettent d'évaluer le chatbot | Les questions portent sur le chatbot de manière générale, on peut moins s'exprimer sur les interactions, par exemple (ex. l'utilité des différentes questions posées par le chatbot) | Pas grand-chose, il est très clair et très bien organisé |
| **6** | Ne permet pas du tout d'évaluer le chatbot | Le questionnaire évalue si cela nous plaît de parler à un chatbot (notre réticence à l'utiliser ou non, si on a eu une bonne expérience). Il permet aussi d'évaluer les points que les utilisateurs semblent insatisfaits. | Je n'ai pas pu bien évaluer le chatbot, car je n'ai pas eu réellement l'impression de l'avoir utilisé. Je n'avais pas forcément de questions à lui poser, ou bien, il ne répondait pas vraiment à mes questions (il me proposait surtout d'aller sur le Forum). Il y avait une différence entre l'étape 1 et 2, mais l'étape 2 et 3 me semblaient similaires. | Des questions ouvertes qui proposent aux utilisateurs de proposer des points d'amélioration du chatbot. |

La deuxième question s'intéresse aux dimensions évaluées par le questionnaire (« En quoi le questionnaire d'évaluation permet d'évaluer le chatbot ? »). Les participants 2 et 4 indiquent que le questionnaire évalue les aspects ergonomiques. Le participant 2 ajoute la dimension logique des réponses. Pour le participant 6, plusieurs éléments sont évalués par le questionnaire : « […] si cela nous plaît de parler à un chatbot (notre réticence à l'utiliser ou non, si on a eu une bonne expérience). Il permet aussi d'évaluer les points [à propos desquels] les utilisateurs semblent insatisfaits ».

Par ailleurs, 2 des 5 participants ayant répondu à la troisième question (« En quoi le questionnaire d'évaluation NE permet PAS d'évaluer le chatbot ? ») mentionnent que le questionnaire ne permet pas d'évaluer l'utilité du chatbot (Participant 2 : « Sur l'utilité

du chatbot [...] ses réponses ne sont pas très utiles en l'état actuel pour nos études ». Participant 5 : « [...] on peut moins s'exprimer sur les interactions par exemple l'utilité des différentes questions posées par le chatbot »).

Enfin, la dernière question demande aux participants de proposer d'éventuelles pistes d'améliorations du questionnaire. Le participant 2 suggère de ne pas s'intéresser à l'utilité du chatbot. Le participant 4 relève que « Les 2 premières questions se ressemblent (ambiguïté) ». Pour rappel, les deux premiers items du questionnaires (« Qu'est-ce qui vous motive à utiliser un chatbot ? » et « Quelle est votre principale raison d'utiliser un chatbot ? ») ont pour objectif de mesurer les facteurs de motivation. Enfin, le participant 6 propose d'ajouter : « des questions ouvertes qui proposent aux utilisateurs de proposer des points d'amélioration du chatbot ».

## 6.5 Évaluation d'ADIDBot

6 participants ont participé à l'évaluation d'ADIDBot en répondant au questionnaire réalisé (voir Tableau 22). Leurs réponses ont été analysées en calculant, dans un premier temps, le score global, qui correspond aux valeurs attribuées par les différents participants aux items d'évaluation d'un critère donné. Dans un second temps, ce score global a été divisé par le nombre d'items évalués pour ce critère afin d'obtenir une moyenne (voir Tableau 24).

Tableau 24 - Scores obtenus par critère

| | Acceptabilité | Comportement | Efficacité | Émotions | Expérience utilisateur | Motivation | Présence | Utilisabilité | Utilité | Impressions Générales |
|---|---|---|---|---|---|---|---|---|---|---|
| **Score global** | 85 | 36 | 127 | 21 | 210 | 28 | 36 | 302 | 136 | 9 |
| **Nombre d'items** | 3 | 1 | 4 | 1 | 6 | 2 | 1 | 10 | 6 | 1 |
| **Moyenne** | 28,33 | 36 | 31,75 | 21 | 35 | 14 | 36 | 30,2 | 22,67 | 9 |

De manière générale, les critères qui sont les moins bien évalués sont compris entre la valeur minimum de 6 et sont inférieurs à la moyenne qui se situe à 24. Pour les critères les mieux évalués, ceux-ci s'approchent de la valeur maximale de 42, comme le montre la figure 5, pour laquelle la moyenne de chaque dimension est regardée selon le classement.

Fig. 5 – Valeur minimale, moyenne et valeur maximale

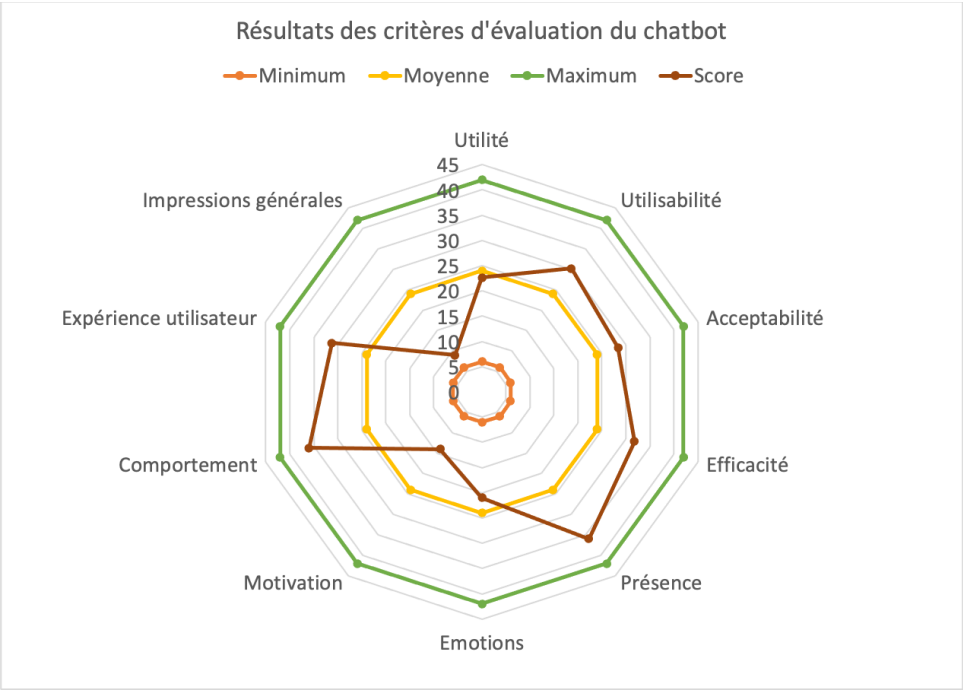


Fig. 6 - Résultats des critères d'évaluation du chatbot

Si l'on regarde plus spécifiquement les résultats des différents critères (voir Figure 6), les critères d'utilité et de l'émotion ont été évalués plutôt négativement par les utilisateurs. En effet, l'utilité du chatbot totalise un score global de 136 pour 6 items, soit une moyenne de 22,67. Les émotions générées par le chatbot totalisent un score global de 21. On peut en déduire que l'émotion générée par le chatbot est plutôt négative. Il en va de même pour le critère de motivation qui totalise le score de 14. En effet, l'analyse des réponses aux questions ouvertes en lien avec le critère de motivation (voir Tableau 25) indique que les utilisateurs se situent plutôt sur le versant de la régulation, ce qui montre une motivation intrinsèque plutôt faible. D'ailleurs, l'un des participants a même mentionné qu'il se sentait obligé de l'utiliser. Pour la moitié d'entre eux, l'utilisation du chatbot répond à une sollicitation externe afin de nous permettre de collecter des données pour notre recherche.

Tableau 25 - Réponses des participants pour l'évaluation du critère de motivation

| *Question* 1 **Participant n°** | *Qu'est-ce qui vous motive à utiliser le chatbot?* | **Type de motivation** Échelle de motivation en éducation (EME) de Vallerand & al. (1989) |
|---|---|---|
| **1** | Obligation | **Amotivation** |
| **2** | Vous aider à récolter des données :). Sinon, de manière générale pour les chatbots, c'est pour avoir une réponse. | **Régulation externe** |
| **3** | Il est sympa | **Régulation externe** |
| **4** | Curiosité, appui pédagogique | **Régulation introjectée** |
| **5** | Le fait de recevoir une réponse d'une entité | **Régulation externe** |
| **6** | Le travail de mémoire d'Aurélie et Simon | **Régulation externe** |

| *Question* 2 **Participant n°** | *Quelle est votre principale raison d'utiliser le chabot?* | **Type de motivation** Échelle de motivation en éducation (EME) de Vallerand & al. (1989) |
|---|---|---|
| **1** | Pour tester | **Régulation externe** |
| **2** | Vous aider à récolter des données :) sinon de manière générale parce que je ne peux pas obtenir de réponse autrement | **Régulation externe** |
| **3** | Participer à son évaluation | **Régulation identifiée** |
| **4** | Lorsque j'ai besoin d'un renseignement | **Régulation externe** |
| **5** | L'étude dans laquelle il est utilisé | **Régulation identifiée** |
| **6** | Aider Aurélie et Simon | **Régulation externe** |

D'autre part, les utilisateurs ne sont ni en accord ni en désaccord par rapport au niveau d'acceptabilité du chatbot qui se situe dans la moyenne. Le critère d'acceptabilité du chatbot totalise un score global de 85 pour 3 items ; soit, une moyenne de 28,33.

En revanche, les critères d'utilisabilité, d'efficacité et d'expérience utilisateur ont été plutôt bien évalués par les utilisateurs. En effet, l'utilisabilité du chatbot totalise un score global de 302 pour 10 items. Soit, une moyenne de 30,2. L'efficacité du chatbot totalise un score global de 127 pour 4 items. Soit, une moyenne de 31,75. On peut en conclure que les utilisateurs sont plutôt d'accord avec son efficacité.

La présence et le comportement sont les dimensions qui totalisent le plus haut score avec un score global de 36. Ce qui indique que le ressenti des utilisateurs est positif et que le chatbot répond correctement aux requêtes des utilisateurs (voir Tableau 26). On peut en déduire que le chatbot a su capter les utilisateurs grâce au côté ludique identifié à l'aide du sous-critère d'évaluation concernant les impressions générales du chatbot, qui totalise un score de 9.

Tableau 26 - Réponses des participants pour l'évaluation du critère de comportement

| *Question* / **Participant n°** | *En quoi le chatbot est-il le plus performant ?* | **Type de comportement** Kirakowski, O'Donnell & Yiu (2009) |
|---|---|---|
| **1** | Humanité des réponses | **Manières** |
| **2** | Il est rapide et logique dans ses réponses. | **Conscience** |
| **3** | Sa proximité dans la relation est vraiment sympa. | **Manières** |
| **4** | Ludique | **Manières** |
| **5** | Offrir une expérience conviviale | **Manières** |
| **6** | L'humour | **Originalité** |

En résumé, le chatbot est jugé comme étant facile d'utilisation, il a une bonne présence, on peut lui faire confiance. Il est transparent et attractif tout en étant ludique. En revanche, l'utilité n'est pas particulièrement bonne et il n'est que moyennement accepté et crédible. Il ne stimule que moyennement les utilisateurs.

# 7 Discussion

L'objectif de ce travail est de concevoir un outil permettant d'évaluer un chatbot utilisé comme assistant de cours dans le cadre d'une formation à distance. Pour ce faire, un questionnaire a été réalisé en intégrant les critères d'évaluation proposés par Hobert (2019) et les items y relatifs issus d'échelles déjà existantes dans la littérature.

Afin de s'assurer de la compréhension des items retenus pour l'évaluation des chatbots assistants, deux pré-tests cognitifs ont été réalisés : le premier par un expert du domaine et le second par les utilisateurs finaux.

Enfin, la version finale du questionnaire a été utilisée afin d'évaluer ADIDBot, un chatbot en cours de conception pour un cours donné de manière hybride à l'Université de Genève.

## 7.1 La construction de l'outil au regard de la littérature

Au regard de la littérature, nous avons déterminé que l'outil le plus approprié est un questionnaire. En effet, l'analyse des données est plus aisée et moins chronophage que pour d'autres méthodes (Temperman et al., 2021) et nous a permis de mettre en lumière les critères les mieux et les moins bien évalués par les utilisateurs. Néanmoins, le questionnaire n'est que moyennement approprié pour recueillir de manière efficiente les ressentis des utilisateurs. Contrairement à ce que Redet et al. (2016) indiquent dans leur recherche (p.3), la collecte de certaines données qualitatives comme l'expression des émotions est plus difficile à formuler par écrit. Pour pouvoir obtenir un retour d'information plus détaillé, il serait intéressant de coupler la passation du questionnaire avec des entretiens avec les utilisateurs (Hobert, 2019, p.268).

Pour construire notre questionnaire, nous nous sommes basés sur les critères d'évaluation des chatbots pédagogiques proposés par Hobert (2019). Son cadre théorique contient des critères pertinents sans toutefois présenter un outil qui permette concrètement de les évaluer. Nous avons donc recherché dans la littérature des items d'évaluation issus d'échelles déjà existantes et idéalement déjà traduites en français qui permettent d'évaluer ces différents critères. Néanmoins, la disponibilité des items était souvent limitée. En effet, peu d'auteurs appliquent le principe de réplicabilité : beaucoup ne fournissent pas l'ensemble du matériel qu'ils ont eux-mêmes utilisé dans leurs recherches. Les items permettant d'évaluer les différentes dimensions d'évaluation ne sont que trop rarement dévoilés par les auteurs, ce qui a rendu la revue de littérature d'autant plus complexe.

Les critères « Autres effets bénéfiques » et « Autres facteurs psychologiques » de Hobert (2019) pouvant également englober d'autres dimensions pertinentes à évaluer, il a été nécessaire de prendre en compte des critères d'évaluation supplémentaires dans la littérature en lien avec l'évaluation d'outils numériques, permettant ainsi d'appréhender les chatbots utilisés dans un contexte éducatif dans leur globalité. Ainsi,

nous avons identifié qu'il était important d'évaluer l'expérience utilisateur de manière générale et plus spécifiquement, les corrélations entre l'acceptation de la technologie et l'utilisation effective du système (Hobert, 2019, p.264). En complément à ces aspects liés à l'outil, il est apparu nécessaire de s'intéresser au ressenti des utilisateurs, à travers les aspects émotionnels et motivationnels. En effet, dans les formations à distance, nous constatons l'importance de l'accompagnement des apprenants à l'aide d'outils numériques comme les chatbots afin de motiver et engager les apprenants dans leurs apprentissages (Nissen, 2009). Ainsi, les critères de présence et d'engagement doivent également être pris en compte lors de l'évaluation de ces outils afin de mieux comprendre les facteurs qui influencent leur utilisation, et permettre de ce fait aux concepteurs de créer des outils en adéquation avec les besoins effectifs des utilisateurs. En outre, dans le cadre spécifique des chatbots, l'aspect humain, le langage naturel, la personnalité et le comportement du chatbot sont autant d'aspects qui peuvent influencer l'utilisation du chatbot (Morrissey & Kirakowski, 2013).

De par le large éventail de critères à évaluer, la liste des items d'évaluation était conséquente. La méthodologie utilisée afin de la réduire a consisté en premier lieu à solliciter un expert du domaine afin de rendre compte de la pertinence des critères et des items à évaluer dans le cadre d'un chatbot assistant de cours, selon les recommandations de Elangovan & Sundaravel (2021). En second lieu, les items ont été soumis aux utilisateurs afin de valider leur compréhension selon les recommandations de Karabenick et al. (2007).

## 7.2 La revue des critères et des items par l'expert VS par les utilisateurs

Une première validation a été réalisée, à distance, par un expert du domaine. Pour ce faire, un fichier Excel contenant l'ensemble des items, classés par critères, a été envoyé à l'expert, qui l'a retourné avec ses annotations. La démarche a nécessité un traitement minutieux car les commentaires figuraient sur chaque item. Afin de pouvoir proposer une deuxième version des items, il a été nécessaire de traiter chaque item individuellement en interprétant les remarques figurant sur le document. Cette manière de procéder diffère de celle utilisée avec les utilisateurs.

Pour la validation par les utilisateurs, celle-ci s'est déroulée sous forme d'entretiens individuels à distance (via Zoom) en manipulant les items selon la méthodologie du tri de cartes. L'accent a principalement été mis sur la validation de la compréhension des items. Cette manière de procéder est ludique (Lallemand et al., 2016) et a été appréciée par les utilisateurs qui ont été très actifs dans le classement des différents items. De plus, le traitement des données a été facilité, contrairement à la revue des items par l'expert, puisque les items ont été classés directement par les utilisateurs selon s'ils étaient (1) Pertinents pour l'évaluation du chatbot, (2) Non pertinents pour l'évaluation du chatbot ou (3) Non compréhensibles. Néanmoins, au vu du grand

nombre d'items, cette démarche a été extrêmement chronophage. La durée moyenne de passation des entretiens se situe autour des 45 minutes.

En conclusion, une discussion autour des items semble être nécessaire afin de s'assurer des raisons de suppression ou de reformulation. Pour l'expert, une discussion aurait été bénéfique et aurait éventuellement permis de discuter de la compréhension des items retirés et conservés. La passation sous forme de tri de cartes a été très intéressante et a permis à chaque participant de s'exprimer sur chacun des items présentés. En revanche, le nombre important d'items rend cette passation très coûteuse en temps, ce qui peut avoir un effet sur la concentration des participants, les amenant donc à moins s'impliquer dans l'analyse des derniers items à classer. En effet, selon Mackworth (1948) cité par Hery (2016) « la vigilance décroît fortement dès 30 minutes et les erreurs de détection augmentent de 10 à 15% ».

## 7.3 L'évaluation du chatbot à l'aide du questionnaire et les dimensions à améliorer

A la fin du questionnaire, quatre questions ont été posées aux utilisateurs afin d'évaluer si celui-ci permet d'évaluer globalement le chatbot et d'obtenir d'éventuelles pistes d'amélioration. D'après leurs réponses, les participants sont mitigés quant à l'efficacité du questionnaire.

Nous relevons néanmoins que les questions posées ne semblent pas collecter de manière optimale les informations souhaitées. En effet, la deuxième question qui avait pour objectif de déceler si le questionnaire permettait d'évaluer le chatbot de manière optimale (« En quoi le questionnaire d'évaluation permet d'évaluer le chatbot ? ») ne semble pas avoir été comprise par l'ensemble des participants. Par exemple, les participants 3 et 5 ne donnent pas d'indication sur les dimensions évaluées par le questionnaire (Participant 3 : « Le questionnaire est complet, il reprend tous les items pertinents pour l'évaluation du chatbot ». Participant 5 « Les différents aspects visés par le questionnaire permettent d'évaluer le chatbot »). La troisième question « En quoi le questionnaire d'évaluation NE permet PAS d'évaluer le chatbot ? » avait pour objectif de déceler d'éventuelles dimensions non évaluées ou moins bien évaluées par le questionnaire. Cependant, les participants ont principalement exprimé leur ressenti au sujet d'ADIDBot dans leurs réponses à la troisième question, bien que cela ne soit pas son but premier.

Concernant les pistes d'amélioration proposées, relevons la proposition du participant 6 qui propose d'ajouter « des questions ouvertes qui proposent aux utilisateurs de proposer des points d'amélioration du chatbot ». Ce dernier constat semble en effet être une remarque pertinente car nous avons observé qu'il nous manque des données en lien avec le ressenti des utilisateurs et leurs besoins. Aussi, nous avons constaté que le questionnaire ne permet pas réellement d'obtenir des pistes d'amélioration du questionnaire et semble en effet ne pas être la méthode la plus appropriée. Il serait

alors judicieux de trouver une autre méthode afin de nous permettre d'obtenir ce type de données. Les entretiens semblent, comme vu précédemment, être la forme qui conviendrait le mieux (Lallemand & Gronier, 2015).

Afin d'avoir une idée plus précise de l'efficacité du questionnaire, il nous faudrait d'une part davantage de résultats, et d'autre part, la passation d'entretiens permettrait d'obtenir des indications plus précises au sujet de ce que les utilisateurs pensent concrètement du questionnaire, tel que suggéré par Lallemand & Gronier (2015).

Par ailleurs, nous relevons que le questionnaire n'a pas permis de mettre en évidence le manque d'acceptabilité du chatbot que nous avons relevé grâce au critère de motivation. (Participant 2 : « Nous l'avons utilisé [le chatbot] pour le tester pas par besoin et ses réponses ne sont pas très utiles en l'état actuel pour nos études ». Participant 4 : [le chatbot est] « plutôt centré sur l'évaluation de la satisfaction et des inquiétudes vis-à-vis du cours ADID mais pourront prendre plus de sens dans le futur avec l'ajout de contenus relatifs à de l'appui pédagogique ». Participant 6 : « [...] je n'ai pas eu réellement l'impression de l'avoir [le chatbot] utilisé. Je n'avais pas forcément de questions à lui poser, ou bien, il ne répondait pas vraiment à mes questions (il me proposait surtout d'aller sur le forum) ». Ces éléments nous permettent d'affirmer que le chatbot ne sera pas réutilisé par les utilisateurs. Il conviendrait alors de revoir les items liés au critère d'acceptabilité afin que celui-ci permette mieux de mettre en évidence si les utilisateurs comptent adopter et intégrer l'outil pour leurs apprentissages (Barcenilla & Bastien, 2009).

## 7.4 De l'évaluation d'ADIDBot

Dans le cadre de l'utilisation du questionnaire à l'aide d'ADIDBot, nous avons pu constater que le questionnaire créé dans le cadre de ce mémoire permettait bien d'évaluer un chatbot assistant de cours. En effet, le questionnaire a su mettre en évidence les critères bien évalués par les utilisateurs et ceux qui n'ont pas été bien évalués. Ces dimensions concordent avec ce que les utilisateurs ont explicité durant les entretiens d'utilisation du chatbot ADIDBot (pré-test cognitif).

En revanche, pour des données qualitatives, l'utilisation du questionnaire a ses limites, notamment pour rendre compte du ressenti des utilisateurs. En effet, les réponses fournies par écrit sont courtes, ne sont pas complètes et peuvent être sujettes à différentes interprétations malgré le fait que celles-ci aient été analysées au regard de modèles théoriques préexistants. Afin de pallier à cette problématique, il conviendrait alors de coupler l'utilisation du questionnaire avec la passation systématique d'entretiens, car cette manière de procéder permet d'obtenir des données en lien avec l'expérience, le ressenti, les besoins des utilisateurs. Les interactions directes avec les utilisateurs leur permettent de rendre compte de l'utilisation de l'outil et ainsi de donner des pistes d'amélioration (Lallemand & Gronier, 2015).

De manière générale, nous retenons que le questionnaire évalue le chatbot de manière plutôt satisfaisante grâce, d'une part, aux retours des utilisateurs (les participants 3 et 5 indiquent que le questionnaire « semble complet » et « est très clair et très bien organisé ») et d'autre part, au regard des résultats obtenus qui coïncident avec les propos exprimés par les utilisateurs durant les entretiens.

## 7.5 Pistes d'amélioration pour ADIDBot

Concernant l'évaluation d'ADIDBot, il serait intéressant, pour le concepteur et les enseignants, de travailler sur les critères les moins bien évalués par les utilisateurs (voir Figure 7) afin d'augmenter son acceptabilité et donner envie aux utilisateurs d'utiliser effectivement l'outil.

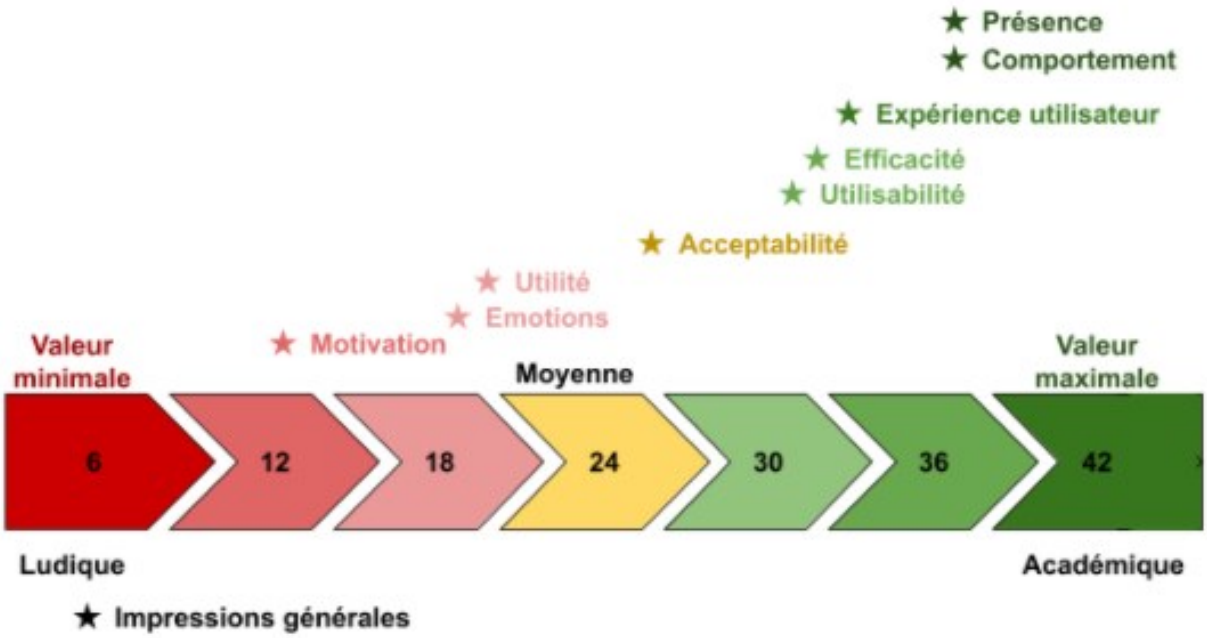


Fig. 7 – Répartition des résultats de l'évaluation d'ADIDBot

L'utilité du chatbot semble être la dimension qui influence en grande partie la motivation des étudiants à utiliser l'outil et par conséquent, son acceptabilité. En effet, les fonctionnalités proposées et leur facilité d'usage jouent un rôle dans l'acceptabilité du produit (Branger & Barcenilla, 2003, p.58). D'ailleurs, au regard des entretiens, il est ressorti que différentes fonctionnalités étaient manquantes, comme par exemple la possibilité de rechercher des informations et des ressources en lien avec le cours.

Les résultats, en lien avec l'aspect motivationnel de l'utilisation d'ADIDBot, montrent un sentiment de devoir utiliser cet outil sans y voir un éventuel bénéfice. Ce constat se retrouve également dans les questions ouvertes à la fin du questionnaire (Participant 2 : « Nous l'avons utilisé [le chatbot] pour le tester pas par besoin et ses réponses ne sont pas très utiles en l'état actuel pour nos études ». Participant 4 : [le chatbot est] « plutôt centré sur l'évaluation de la satisfaction et des inquiétudes vis-à-vis du cours ADID mais pourront prendre plus de sens dans le futur avec l'ajout de contenus relatifs à de l'appui pédagogique ». Participant 6 : « [...] je n'ai pas eu réellement l'impression

de l'avoir [le chatbot] utilisé. Je n'avais pas forcément de questions à lui poser, ou bien, il ne répondait pas vraiment à mes questions (il me proposait surtout d'aller sur le forum) ». Ces éléments nous permettent d'affirmer que le chatbot ne sera pas réutilisé par les utilisateurs. En effet, au regard des formes principales de motivation énoncées par Vallerand et al. (1989), les utilisateurs d'ADIDBot ont majoritairement des sources de motivation extrinsèques (« Le travail de mémoire d'Aurélie et Simon » (participant 6) ou « Vous aider à récolter des données » (participant 2), « L'étude dans lequel [le chatbot] est utilisé » (participant 5)) et même d'amotivation par « l'obligation » mentionné par le participant 1.

Ce résultat va de pair avec l'utilité perçue par les utilisateurs et les émotions que génère le chatbot. Pour Hobert (2019), si le chatbot n'est pas capable de comprendre les intentions des apprenants, ceux-ci peuvent refuser de l'utiliser (p. 267). Selon Tricot (2001, p.3), l'utilité est liée à la pertinence de l'outil, à sa finalité, c'est-à-dire dans quel but le chatbot est utilisé. En outre, les émotions suscitées par le chatbot influencent son acceptabilité. En travaillant sur ces aspects, l'utilisation future de l'outil sera meilleure (Redet et al., 2016).

Par ailleurs, une fois la conception du chatbot finalisée, il serait intéressant de mesurer si l'augmentation de la motivation se traduit également par une augmentation de la réussite de l'apprentissage, tel que suggéré par Hobert (2019) (p.264). Ces données pourraient être obtenues grâce à des entretiens avec les utilisateurs afin de mieux comprendre les effets sur les processus d'apprentissage (Hobert, 2019, p.268).

## 7.6 Limites et points d'amélioration

Concernant les items, nous relevons que les traductions des échelles de Venkatesh et al. (2003), Bradley & Lang (1994) et Kay & Loverock (2008) n'étaient pas disponibles dans la littérature. Ces échelles ont donc été traduites de manière libre comme indiqué dans les annexes 4, 10 et 11. Une piste d'amélioration serait de trouver une traduction qui ait été déjà validée ou de procéder à leur validation (Besson & Haddadj, 2003).

D'autre part, l'une des principales difficultés de ce travail a été de traiter, tirer et réduire la grande quantité d'items présents dans les différentes échelles issues de la littérature, de repérer les items identiques et de s'assurer qu'ils évaluaient bien le même critère d'évaluation. Bien qu'un expert ait contribué à cette tâche, un seul avis a été sollicité. D'autres confrontations avec d'autres experts auraient peut-être pu contribuer à un tri différent. De plus, la forme utilisée pour réaliser le feedback n'est pas optimale (envoi d'un fichier Excel par e-mail). Une discussion ou un premier tri de cartes aurait été plus efficient et homogène au niveau de la méthodologie adoptée.

Le grand nombre d'items à trier a également sollicité davantage de temps pour les utilisateurs. Une réduction plus importante aurait permis d'éventuellement motiver plus de participants aux entretiens.

À l'issue de ces pré-tests, nous notons que les critères évalués ne comportent pas tous le même nombre d'items. Par conséquent, il serait intéressant de proposer un nombre identique d'items par critère afin que s'assurer que cette différence de nombre n'influence pas les résultats du questionnaire.

Concernant le traitement des données qualitatives, leur analyse se base sur la lecture et l'interprétation des données récoltées sous forme écrite. Ces données sont courtes et peuvent être sujettes à interprétation. L'utilisation d'entretiens semble être la forme la plus appropriée afin de récolter ce type de données (Lallemand & Gronier, 2015).

En revanche, malgré le fait que les items utilisés soient déjà validés et issus de la littérature, le questionnaire manque tout de même d'une évaluation quantitative (Bouletreau, 1999) ou d'une validité psychométrique (André et al., 2015). L'intérêt de procéder à ces contrôles réside dans le fait qu'ils permettent de juger de sa fiabilité (André et al., 2015). Dans le cadre de notre recherche, calculer cette validité permettrait éventuellement de mettre en avant cette différence de nombre d'items évalués pour chaque critère et également de rendre compte de la validité du questionnaire. L'intérêt serait de s'assurer que chaque dimension est bien mesurée avec les items proposés.

Au vu de ce qui précède, la présente recherche pourrait être poursuivie sous différents aspects en proposant par exemple une version plus courte du questionnaire. Il serait également intéressant de compléter cette recherche en procédant à une nouvelle analyse des items par un plus grand nombre d'experts et d'utilisateurs afin d'obtenir des données plus significatives.

# 8 Conclusion

Ce mémoire avait pour objectif de créer un outil d'évaluation pour les chatbots utilisés en tant qu'assistants de cours dans une formation hybride et/ou à distance. Les chatbots étant relativement récents dans le domaine éducatif, nous avons constaté qu'il n'existe actuellement que peu d'outils d'évaluation qui leur sont dédiés.

Il a donc été question dans un premier temps de voir ce que propose la littérature afin de déterminer la méthode la plus efficace pour évaluer un chatbot assistant de cours, et d'identifier quels sont les meilleurs critères d'évaluation ainsi que les items qui permettent d'évaluer et de mesurer efficacement la qualité du chabot. Dans un deuxième temps, l'outil a été testé par les utilisateurs afin de l'améliorer. Enfin, dans un troisième temps, cet outil a été utilisé en contexte réel pour évaluer le chatbot développé par Simon Jöhr dans le cadre du cours ADID afin de lui permettre d'obtenir des pistes d'améliorations.

Nous avons relevé dans la littérature que les méthodes utilisées pour évaluer les chatbots sont principalement des questionnaires (Verleger & Pembridge, 2018 ; Santoso et al., 2018 ; Hobert, 2019 ; Smutny & Schereiberova, 2020), des analyses de la satisfaction des utilisateurs (Santoso et al., 2018), des entretiens (Hobert, 2019), des collectes de données (Radziwill & Benton, 2017 ; Savin-Baden et al., 2015) ou des analyses des logs de conversation (Hobert, 2019). Dans le cadre de cette recherche, nous avons déterminé que l'outil le plus approprié est un questionnaire. En effet, cette méthode d'évaluation est la plus utilisée à la fois pour l'évaluation de logiciels pédagogiques (Hû, Trigano & Crozat, 2001) ainsi que pour évaluer les chatbots (Hobert, 2019 et Pérez, Daradoumis & Puig, 2020). Le questionnaire est un outil relativement simple à réaliser. Sa facilité d'analyse et le gain de temps pour l'analyse des données (Temperman et al., 2021) en font des arguments de taille. En revanche, il n'est que moyennement approprié pour recueillir de manière efficiente les ressentis des utilisateurs contrairement à ce que Redet et al. (2016) indiquent dans leur recherche (p.3).

Hobert (2019) met en évidence les critères pertinents pour l'évaluation des chatbots pédagogiques mais ne propose pas d'outil pour les évaluer concrètement. Nous nous sommes donc basés sur cette liste de critères d'évaluation afin de rechercher dans la littérature des échelles déjà existantes permettant d'évaluer ces différents critères.

À l'issue de cette revue de littérature, il convenait de juger de la pertinence et de la compréhension des différents critères et items proposés à l'aide de pré-tests. Pour ce faire, les critères et les items recensés ont été soumis une première fois aux experts dans le domaine éducatif. Après cette première révision, 66 items répartis en 9 critères ont été retenus. Ensuite, ces items ont fait l'objet d'une deuxième validation cognitive par les utilisateurs du chatbot. Suite à cette seconde révision, 36 items ont été conservés par les utilisateurs. Enfin, les items ont été réarrangés et analysés. En

définitive, 35 items répartis en 10 critères (impressions générales, expérience utilisateur, acceptabilité, utilité, utilisabilité, efficacité, émotions, sentiment de présence, comportement, motivation) ont été retenus afin de constituer le questionnaire d'évaluation pour les chatbots utilisés comme assistants de cours.

Sur la base de ces éléments, la présente recherche propose un questionnaire d'évaluation mesurant ces 10 critères à l'aide de 35 items d'évaluation présentés sous forme de phrase, de paires de mots et de questions ouvertes. Pour les deux premiers types d'items, les utilisateurs sont invités à les évaluer en se positionnant sur une échelle de Likert en 7 points allant de « 1 - Pas du tout d'accord » à « 7 - Tout à fait d'accord ». En calculant, le score global des items puis la moyenne par critère, les scores peuvent être comparés entre eux afin de déterminer les critères les mieux évalués ou au contraire les moins bien évalués par les utilisateurs. Pour les items sous forme de questions ouvertes, celles-ci évaluent les critères de motivation et de comportement du chatbot. Afin d'analyser les données qualitatives récoltées, les réponses des utilisateurs ont été analysées : pour la motivation, sur la base de l'échelle de motivation de Fenouillet, Heutte & Vallerand (2015) et pour le comportement, sur la base des caractéristiques du caractère du chatbot de Kirakowski, O'Donnell & Yiu (2009).

Le questionnaire réalisé dans le cadre de ce mémoire a fait l'objet d'une validation qualitative (Bouletreau et al., 1999) mais ne jouit pas d'une validité psychométrique (André et al., 2015).

Enfin, le questionnaire a été utilisé pour évaluer ADIDBot. Il en résulte que les utilisateurs jugent ce chatbot comme étant ludique (en opposition à académique) et sont mitigés quant à son acceptabilité. De plus, les critères d'utilisabilité, d'efficacité et d'expérience utilisateur sont des critères plutôt bien évalués par les utilisateurs. La présence et le comportement sont les critères jouissant du plus haut score, indiquant que le ressenti des utilisateurs est positif. En revanche, son utilité et l'émotion qu'il génère ont été évaluées plutôt négativement par les utilisateurs.

Au regard des entretiens menés lors de l'utilisation et de l'évaluation du chatbot, ces résultats coïncident avec les propos exprimés par les utilisateurs. Nous pouvons donc en déduire que le questionnaire permet effectivement d'évaluer les chatbots assistants de cours. Toutefois, au vu de ce qui précède, il serait intéressant de compléter cette recherche en procédant à la validation du questionnaire réalisé.

# 9 Références


AbuShawar, B., & Atwell, E. (2016). Usefulness, localizability, humanness, and language-benefit: additional evaluation criteria for natural language dialogue systems. *International Journal of Speech Technology*, *19*(2), 373-383.

Ahaji, K., El Hajjami, A., El Mokri, A., Ajana, L., & Chikhaoui, A. (2008). Étude et évaluations d'outils multimédias d'enseignantes et d'enseignants innovants: Une expérience dans le système éducatif marocain. *Revue de l'EPI (Enseignement Public et Informatique)*, 110 [En ligne]. Repéré à https://edutice.archives-ouvertes.fr/edutice-00402582/

Allison, D. A. (2012). Chatbots in the library: is it time? *Library Hi Tech*, *30*(1), 95-107. https://doi.org/10.1108/07378831211213238

Anderson, T., Rourke, L., Garrison, D. R., & Archer, W. (2001). Assessing teaching presence in a computer conference context. *Journal of Asynchronous Learning Networks*, *5*(2) [En ligne]. Repéré à http://www.sloan-c.org/publications/jaln/v5n2/pdf/

André, N., Loye, N. & Laurencelle, L. (2015). La validité psychométrique : un regard global sur le concept centenaire, sa genèse, ses avatars. *Mesure et évaluation en éducation*, *37*(3), 125–148. https://doi.org/10.7202/1036330ar

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction*, *24*(6), 574-594. https://doi.org/10.1080/10447310802205776

Barbier, J. M. (2016). Tendances d'évolution de la formation des adultes. In R. Wittorski (Ed.), *La professionnalisation en formation* (pp. 17-33). Presses universitaires de Rouen et du Havre.

Barcenilla, J., & Bastien, J. (2009). L'acceptabilité des nouvelles technologies : quelles relations avec l'ergonomie, l'utilisabilité et l'expérience utilisateur ? *Le travail humain*, *72*, 311-331. https://doi.org/10.3917/th.724.0311

Besson, D., & Haddadj, S. (2003, juin). *Adaptation internationale des échelles de mesure entre universalisme et culturalisme: application à la mesure de l'environnement de l'entreprise*. Communication présentée à la XIIème Conférence de l'Association Internationale de Management Stratégique (AIMS), Tunisie.

Bétrancourt, M. (2007). L'ergonomie des TICE quelles recherches pour quels usages sur le terrain. In *Transformation des regards sur la recherche en technologie de l'éducation* (pp.77-89). Louvain-la-Neuve : De Boeck Supérieur. https://doi.org/10.3917/dbu.charl.2007.01.0077

Bouletreau, A., Chouaniere, D., Wild, P. & Fontana J. M. (1999). *Concevoir, traduire et valider un questionnaire. A propos d'un exemple, EUROQUEST*. [Rapport de recherche] Notes scientifiques et techniques de l'INRS NS 178, Paris : Institut National de Recherche et de Sécurité (INRS). Repéré à https://hal.archives-ouvertes.fr/hal-01420163

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, *25*(1), 49-59. https://doi.org/10.1016/0005-7916(94)90063-9

Brandtzaeg, P. B., & Følstad, A. (2017). Why people use chatbots. In *International conference on internet science* (pp. 377-392). Springer, Cham. https://doi.org/10.1007/978-3-319-70284-1_30

Brangier, É., & Barcenilla, J. (2003). *Concevoir un produit facile à utiliser*. *Adapter les technologies à l'homme.* Paris : Éditions d'organisation.

Brassard, C., & Teutsch, P. (2014). Proposition de critères de proximité pour l'analyse des dispositifs de formation médiatisée. *Distances et médiations des savoirs*, 2(5). https://doi.org/10.4000/dms.646

Brooke, J. (1996). SUS: a 'quick and dirty' usability scale. In P. W. Jordan, B. Thomas, B. Weerdmeester, & I. L. McClelland (Eds.) *Usability Evaluation In Industry* (pp. 189-194). Grande-Bretagne: T.J. Press (Padstow) Ltd.

Caron, P.-A., & Heutte, J. (2017). Comprendre l'usage que les professeurs des écoles font des TNI et du numérique. (p. 341-352). In N. Guin, B. De Lièvre, M. Trestini & B. Coulibaly (Eds.) In *Actes de la 8e Conférence sur les Environnements Informatiques pour l'Apprentissage Humain - EIAH (Strasbourg, juin 2017)*. Repéré à https://www.researchgate.net/publication/320078827_Comprendre_l%27usage_que_les_professeurs_des_ecoles_font_des_TNI_et_du_numerique

Clark, D. (2018, 10 avril). The Fallacy of 'Robot' Teachers [Billet de blog]. Repéré à https://donaldclarkplanb.blogspot.com/search?q=10+uses+for+Chatbots+in+learning+(with+examples)

Coron, C. (2020). Outil 10. Le questionnaire : les questions. In C. Coron (Ed.), *La Boîte à outils de l'analyse de données en entreprise* (pp. 36-37). Paris : Dunod. Repéré à https://www.cairn.info/la-boite-a-outils-de-l-analyse-de-donnees--9782100808557.htm

Crozat, S., Trigano, P., & Hû, O. (1999). EMPI : Une méthode informatisée pour l'évaluation des didacticiels multimédias. *Revue des Interactions Humaines Médiatisées (RIHM) Journal of Human Mediated Interactions*, *1*(2), 61-87. Repéré à https://edutice.archives-ouvertes.fr/edutice-00000391/document

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319-340. https://doi.org/10.2307/249008

Dow, S., Mehta, M., Harmon, E., MacIntyre, B., & Mateas, M. (2007). Presence and engagement in an interactive drama. In *Proceedings of the SIGCHI conference on Human factors in computing systems (San Jose California, 28 avril - 3 mai 2007).* (pp.1475-1484). https://doi.org/10.1145/1240624.1240847

Dupré D., Dubois M., Tcherkassof A. (2018). Rôle de l'émotion dans l'acceptabilité d'un produit : évaluation des composantes cognitive, motivationnelle et subjective. *Psychologie du Travail et des Organisations*, *24*(4), 313-324. https://doi.org/10.1016/j.pto.2017.07.001

Dussarps, C. (2015). L'abandon en formation à distance. *Distances et médiations des savoirs*, 10 [En ligne]. https://doi.org/10.4000/dms.1039

Elangovan, N. & Sundaravel, E. (2021). Method of preparing a document for survey instrument validation by experts. *MethodsX, 8*(101326). https://doi.org/10.1016/j.mex.2021.101326

Fenouillet F., Heutte J., & Vallerand R.-J. (2015, juin), Validation of the Adult Education Motivation Scale. Communication présentée au *Fourth World Congress on Positive Psychology (IPPA)*, Orlando, Floride. https://doi.org/10.13140/RG.2.1.3511.3040

Fiock, H. (2020). Designing a community of inquiry in online courses. *The International Review of Research in Open and Distributed Learning*, *21*(1), 135-153. https://doi.org/10.19173/irrodl.v20i5.3985

Fontaine, G., Cossette, S., Maheu-Cadotte, M.-A., Mailhot, T., Lavoie, P., Gagnon, M.-P., Dubé, V., & Côté, J. (2019). Traduction, adaptation et évaluation psychométrique préliminaire d'une mesure d'engagement et d'une mesure de charge cognitive en contexte d'apprentissage numérique. *Pédagogie Médicale*, *20*(2), 79-90. https://doi.org/10.1051/pmed/2020009

Garcia Brustenga, G., Fuertes Alpiste, M., & Molas Castells, N. (2018). Briefing Paper: Chatbots in Education. Barcelone : eLearn Center. Universitat Oberta de Catalunya (UOC). https://doi.org/10.7238/elc.chatbots.2018

Gérin-Lajoie, S. (2018). *L'encadrement pédagogique à distance : Comment le réaliser dans l'univers numérique d'aujourd'hui ?* Communication présentée au Séminaire académique pour la mission royale belge, 14 mars 2018, Montréal. Repéré à https://r-libre.teluq.ca/1555/1/Gerin-Lajoie%20%282018%29_Visite%20royale%2018-03-14_Vfinal.pdf

Gil, S. (2009). Comment étudier les émotions en laboratoire. *Revue électronique de psychologie sociale*, *4*, 15-24. Repéré à https://psychologiescientifique.org/wp-content/uploads/2018/02/Gil-2009-Comment-%C3%A9tudier-les-%C3%A9motions-en-laboratoire.pdf

Grislin, M. & Kolski, C. (1996). Évaluation des interfaces homme-machine lors du développement de système interactif. *Technique et Sciences Informatiques (TSI)*, *15*(3), 265-296. Repéré à https://www.researchgate.net/publication/228604487_Evaluation_des_interfaces_homme-machine_lors_du_developpement_des_systemes_interactifs

Gronier, G., & Baudet, A. (2021). Psychometric Evaluation of the F-SUS: Creation and Validation of the French Version of the System Usability Scale. *International Journal of Human–Computer Interaction*, *37*(16), 1571-1582. https://doi.org/10.1080/10447318.2021.1898828

Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In G. Szwillus & J. Ziegler (Eds.), *Mensch & Computer 2003. Berichte des German Chapter of the ACM, vol 57*

(pp. 187-196). Vieweg+Teubner Verlag. https://doi.org/10.1007/978-3-322-80058-9_19

Heutte, J., Caron, P.-A., Fenouillet, F. & Vallerand, R. J. (2016). Étude des liens entre les caractéristiques instrumentales et les différents types de motivations des participants dans un MOOC. *Revue internationale des technologies en pédagogie universitaire / International Journal of Technologies in Higher Education*, *13*(2-3), 94–110. Repéré à https://id.erudit.org/iderudit/1039247ar

Hery, I. (2016, 1 juillet). Concentration et attention en profondeur : durée moyenne et durée idéale. [Billet de blog]. Repéré à https://letempsreconquis.fr/concentration-et-attention-en-profondeur-duree-ideale

Hinderks, A., Schrepp, M., & Thomaschewski, J. (2018). UEQ User Experience Questionnaire. Repéré à https://www.ueq-online.org/

Hobert, S. (2019). How are you, chatbot? evaluating chatbots in educational settings - results of a literature review. In N. Pinkwart & J. Konert (Eds.), DELFI 2019 (pp. 259-270). doi: 10.18420/delfi2019_289

Hû, O., Trigano, P., & Crozat, S. (2001). Une aide à l'évaluation des logiciels multimédias de formation. *Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation, 8*(3), 239-274. https://doi.org/10.3406/stice.2001.1486

Huu Binh, N. (2014). Fiabilité et validité du Modèle d'acceptation de la technologie (TAM) dans le contexte d'apprenants vietnamiens du français comme langue étrangère face aux TIC. *Revue internationale des technologies en pédagogie universitaire, 11*(3), 38–50. https://doi.org/10.7202/1035702ar

Jézégou, A. (2007). La distance en formation : Premier jalon pour une opérationnalisation de la théorie de la distance transactionnelle. *Distances et savoirs, 5*(3), 341-366. https://doi.org/10.3166/ds.5.341-366

Jézégou, A. (2010). Créer de la présence à distance en e-learning : Cadre théorique, définition, et dimensions clés. *Distances et savoirs, 8*(2), 257-274. Repéré à https://www.cairn.info/revue-distances-et-savoirs-2010-2-page-257.htm

Jöhr, S. (En cours de publication). *Élaboration d'un chatbot pour soutenir l'apprentissage du cours ADID de l'Université de Genève (Titre provisoire).* Mémoire de master en sciences et technologies de l'apprentissage, Université de Genève.

Karabenick, S. A., Wolley, M. E., Friedel, J. M., Ammon, B. V., Blazevski, J., Bonney, C. R.(...) (2007). Cognitive Processing of Self-Report Items in Educational Research: Do They Think What We Mean? *Educational Psychologist*, *42*(3), 139-151. https://doi.org/10.1080/00461520701416231

Karsenti, T. (2003). Conditions d'efficacité des formations ouvertes ou à distance (FOAD) en pédagogie universitaire. *Pédagogie médicale, 4*(4), 223-234. Repéré à https://www.pedagogie-medicale.org/articles/pmed/pdf/2003/04/pmed20034p223.pdf

Kay, R. H., & Loverock, S. (2008). Assessing emotions related to learning new software: The computer emotion scale. *Computers in Human Behavior, 24*(4), 1605-1623. https://doi.org/10.1016/j.chb.2007.06.002

Kirakowski, J., O'Donnell, P., Yiu, A. (2009). Establishing the Hallmarks of a Convincing Chatbot-Human Dialogue. In I. Maurtua (Ed.), *Human-Computer Interaction* (pp. 145-154). Rijeka, Croatie: BoD - Books on Demand. Repéré à https://1lib.ch/book/3149494/ba8221

Kukulska-Hulme, A., Bossu, C., Coughlan, T., Ferguson, R., FitzGerald, E., Gaved, M., Herodotou, C., Rienties, B., Sargent, J., Scanlon, E., Tang, J., Wang, Q., Whitelock, D., Zhang, S. (2021). *Innovating Pedagogy 2021: Exploring New Forms of Teaching, Learning and Assessment To Guide Educators and Policy* - Makers. *Open University Innovation Report 9*. Milton Keynes : The Open University.

Laboratoire de Cyberpsychologie de l'UQO (2002). Questionnaire sur l'état de présence. Québec : Université du Québec en Outaouais. Repéré à http://w3.uqo.ca/cyberpsy/wp-content/uploads/2019/04/QEP_vf.pdf

Lallemand, C. (2015, 4 novembre). Évaluer les émotions [Billet de blog]. Repéré à https://uxmind.eu/2015/04/09/evaluer-les-emotions/

Lallemand, C., & Gronier, G. (2015). *Méthodes de design UX : 30 méthodes fondamentales pour concevoir et évaluer les systèmes interactifs*. Paris : Eyrolles.

Lallemand, C., Koenig, V., Gronier, G., & Martin, R. (2015). Création et validation d'une version française du questionnaire AttrakDiff pour l'évaluation de l'expérience utilisateur des systèmes interactifs. *European Review of Applied Psychology, 65*(5), 239-252. https://doi.org/10.1016/j.erap.2015.08.002

Laugwitz B., Held T., Schrepp M. (2008). Construction and Evaluation of a User Experience Questionnaire. In A. Holzinger (Ed.), *Symposium of the Austrian HCI and Usability for Education and Work. USAB 2008 (Graz, Austria, 20.21 novembre 2008). Lecture Notes in Computer Science, vol. 5298,* pp. 63-76. Berlin : Springer. https://doi.org/10.1007/978-3-540-89350-9_6

Lockner, D., Bonnardel, N., Bouchard, C., & Rieuf, V. (2014). Émotion et Design d'Interface. In *Proceedings of the 2014 Ergonomie et Informatique Avancée Conférence - Design, Ergonomie et IHM : quelle articulation pour la co-conception de l'interaction ? (Bidart-Biarritz, 15-17 octobre 2014, pp. 33-40).* Repéré à https://www.researchgate.net/profile/Damien-Lockner/publication/278963577_Emotion_et_Design_d'Interface/links/5588133c08ae1dfa49d22a48/Emotion-et-Design-dInterface.pdf

Lowenthal, D. A., & Lowenthal, P. R. (2010). A mixed methods examination of instructor social presence in accelerated online courses. In L. Kyei-Blankson, J. Blankson, E. Ntuli, & C. Agyeman (Eds.), *Handbook of research on strategic management of interaction, presence and participation in online courses* (pp. 147-159). Hershey : IGI Global. doi :10.4018/978-1-4666-9582-5.ch006

Lund, A. M. (2001). Measuring usability with the USE questionnaire. *Usability interface, 8*(2), 3-6. Repéré à https://www.researchgate.net/profile/Arnold-Lund/publication/230786746_Measuring_Usability_with_the_USE_Questionnaire/links/56e5a90e08ae98445c21561c/Measuring-Usability-with-the-USE-Questionnaire.pdf

Meyer von Wolff, R., Hobert, S., & Schumann, M. (2019). How May I Help You? – State of the Art and Open Research Questions for Chatbots at the Digital Workplace. In *Proceedings of the 52nd Hawaii international conference on system sciences (Maui, January 2019, pp. 95-104)*. Repéré à https://scholarspace.manoa.hawaii.edu/bitstream/10125/59450/1/0010.pdf

Monfort, E., Quillon-Dupré, L., Jouanneaux, N. & Deschaintre-Nazaret, M. (2018). USE-Age Traduction et validation d'un questionnaire d'évaluation de l'acceptation de l'informatique par les personnes âgées. Données préliminaires. Repéré à https://www.researchgate.net/publication/327551613_USE-Age_Traduction_et_validation_d'un_questionnaire_d'evaluation_de_l'acceptation_de_l'informatique_par_les_personnes_agees_Donnees_preliminaires

Morrissey, K., & Kirakowski, J. (2013). 'Realness' in Chatbots: Establishing Quantifiable Criteria. In M. Kurosu (Ed.) *Proceedings of the 15th international conference on Human-Computer Interaction: interaction modalities and techniques (Las Vegas 21-26 juillet 2013)*. Volume 8007, pp. 87-96. Berlin : Springer. https://doi.org/10.1007/978-3-642-39330-3_10

Nieves, B. (2018, 3 mai). IA Conversacional : definición y conceptos básicos. [Billet de blog]. Repéré à https://planetachatbot.com/ia-conversacional-conceptos-basicos-y-la-definicion-107529e213c1

Nissen, E. (2009). Accompagnement dans une formation à distance et dans une formation hybride : Analyse de pratiques. In A. Rivens Mompean & M.-J. Barbot (Eds.), *Dispositifs médiatisés en langues et évolutions professionnelles pour l'accompagnement-tutorat*. (Université Lille 3 : Collection « travaux et recherches » ; pp. 191-210). Villeneuve d'Ascq : CEGES. Repéré à https://hal.archives-ouvertes.fr/hal-00785937/document

Nogry, S., Jean-Daubias, S., Ollagnier-Beldame, M. (2004). Évaluation des EIAH : une nécessaire diversité des méthodes. In *Technologies de l'Information et de la Connaissance dans l'Enseignement Supérieur et l'Industrie (Université de technologie de Compiègne, octobre 2014)* (pp. 39-45). Repéré à http://sup.ups-tlse.fr/documentation/docs/fich_170.pdf#page=39

O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology, 61*(1), 50-69. https://doi.org/10.1002/asi.21229

O'Brien, H. L., Cairns, P., & Hall, M. (2018). A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies, 112*, 28-39. https://doi.org/10.1016/j.ijhcs.2018.01.004

Peras, D. (2018). Chatbot evaluation metrics. In *36th International Scientific Conference on Economic and Social Development (Zagreb, 14-15 décembre 2019, pp. 89-97)*. Repéré

à https://www.researchgate.net/profile/Paula-Gargoles/publication/336923013_COUNTRY_OF_ORIGIN_INFASHION_INDUSTRY_-A_LITERATURE_REVIEW/links/5dbaf7fa299bf1a47b045547/COUNTRY-OF-ORIGIN-INFASHION-INDUSTRY-A-LITERATURE-REVIEW.pdf#page=100

Pérez, J. Q., Daradoumis, T., & Puig, J. M. M. (2020). Rediscovering the use of chatbots in education: A systematic literature review. *Computer Applications in Engineering Education, 28*(6), 1549-1565. https://doi.org/10.1002/cae.22326

Pultrone, B. (2020). *Les mutations du monde du travail et leurs impacts sur les travailleurs peu qualifiés : Paper I : Formation continue et employabilité ; Paper II : Évolution du travail et changement de paradigme.* Mémoire de Master en science du travail, Université catholique de Louvain. Repéré à https://dial.uclouvain.be/memoire/ucl/object/thesis:23721

Radziwill, N. M, & Benton, M. C. (2017). Evaluating quality of chatbots and intelligent conversational agents. *Software Quality Professional, 19*(3), 25-36.

Rebah, H. B, et Dabove, G., M. (2017). Étude de la motivation autodéterminée des étudiants dans le contexte d'une activité pédagogique faisant appel à Facebook comme plateforme d'échange, *Distances et médiations des savoirs*, *17.* https://doi.org/10.4000/dms.1758

Redet, J., Vian, M., Esposito, N., & Tijus, C. (2016). Comparaison d'outils adaptés au terrain pour l'évaluation du ressenti à l'usage des IHM. In *Actes de Ergo'IA (Conférence internationale en ergonomie et informatique avancée) (Bidart-Biarritz, juillet 2016)*. Repéré à http://nicolasesposito.fr/publications_fichiers/redet2016outils.pdf

Richardson, J. C., & Swan, K. (2003). Examining social presence in online courses in relation to students' perceived learning and satisfaction. *Journal of Asynchronous Learning Networks*, 7(1), 68−88. Repéré à https://www.researchgate.net/profile/Karen-Swan-3/publication/228953009_An_Examination_of_Social_Presence_in_Online_Courses_in_Relation_to_Students'_Perceived_Learning_and_Satisfaction/links/548083760cf22525dcb5f28e/An-Examination-of-Social-Presence-in-Online-Courses-in-Relation-to-Students-Perceived-Learning-and-Satisfaction.pdf

Rinkus, M.A., Donovan, S. M., Hall, T.E., & O'Rourke, M. (2020). Using a survey to initiate and sustain productive group dialogue in focus groups. *International Journal of Social Reseach Methodology*, *24*(3), 327-340. https://doi.org/10.1080/13645579.2020.1786240

Robillard, G., Bouchard, S., Renaud, P., & Cournoyer, L.-G. (2002). Validation canadienne-française de deux mesures importantes en réalité virtuelle : L'Immersive Tendencies Questionnaire et le Presence Questionnaire. Poster présenté au *25e congrès annuel de la Société Québécoise pour la Recherche en Psychologie 2002, Trois-Rivières*. Repéré à http://w3.uqo.ca/cyberpsy/wp-content/uploads/2019/04/QEP_vf.pdf

Ronce, S. (2018). *Comprendre et développer un Chatbot : Messagerie instantanée, paiements en ligne, intelligence artificielle*. Paris : Eyrolles.

Sadiq, M., Talbi, M., Benichou, E., & Janati-Idrissi, R. (2009). Méthode d'évaluation ergo-pédagogique des supports multimédias de formation. *EpiNet : Revue électronique de l'EPI*, *118*. Repéré à http://www.epi.asso.fr/revue/articles/a0910d.htm

Sauvé, L. (avec la coll. de C. Papi, S. Gérin-Lajoie, G. Desjardins, Racette, N., D. Paquelin, & T. Laferrière) (2020). *Regard des apprenant.es universitaires sur les modes d'organisation et d'encadrement pédagogique en formation à distance et en ligne*. Québec : Université TÉLUQ. Repéré à https://r-libre.teluq.ca/2281/1/Rapport%20scientifique_FQRSC_vfinale.pdf

Santoso, H. A., Winarsih, N. A. S., Mulyanto, E., Sukmana, S. E., Rustad, S., Rohman, M. S., ... Firdausillah, F. (2018). Dinus Intelligent Assistance (DINA) chatbot for university admission services. In *International Seminar on Application for Technology of Information and Communication (Semarang, 21-22 septembre 2018, pp. 417-423*). https://doi.org/10.1109/ISEMANTIC.2018.8549797

Savin-Baden, M., Tombs, G. & Bhakta, R. (2015). Beyond robotic wastelands of time: Abandoned pedagogical agents and new pedalled pedagogies, *E-Learning and Digital Media*, 12(3-4), 295–314. https://doi.org/10.1177/2042753015571835

Schrepp, M., & Thomaschewski, J. (2019). UEQ+ A Modular Extension of the User Experience Questionnaire. Repéré à http://ueqplus.ueq-research.org

Senach, B. (1990). *Évaluation ergonomique des interfaces homme-machine : une revue de la littérature* (Rapport - mars 1990). Le Chesney, France : Institut National de Recherche en Informatique et en Automatique (INRIA). Repéré à https://hal.inria.fr/inria-00075378/document

Smutny, P. & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers & Education, 151, 103862.* https://doi.org/10.1016/j.compedu.2020.103862

Tavanapour, N., & Bittner, E. A. (2018). Automated facilitation for idea platforms: design and evaluation of a Chatbot prototype. In *Thirty Ninth International Conference on Information Systems* (San Francisco, 2018). Repéré à https://www.researchgate.net/profile/Navid-Tavanapour/publication/328145863_Automated_Facilitation_for_Idea_Platforms_Design_and_Evaluation_of_a_Chatbot_Prototype/links/5c49ac7f299bf12be3df7e23/Automated-Facilitation-for-Idea-Platforms-Design-and-Evaluation-of-a-Chatbot-Prototype.pdf

Temperman, G., Boumazguida, K., Housni, S., Piret, G., Kumps, A., Laëtitia Dragone, L., Molderez, L., Descamps, S. & De Lièvre, B. (2021). *Évaluation des environnements numériques pour l'apprentissage humain* [MOOC]. UMONS. Repéré sur la plateforme Moodle : https://umooc.umons.ac.be/course/view.php?id=163

Tricot, A. (2001). Interpréter les liens entre utilisabilité et utilité des documents électroniques. Les documents électroniques, méthodes, démarches et techniques cognitives. In M. Mojahid & J. Virbel (Eds.). *Les documents électroniques, méthodes, démarches et techniques cognitives*. Paris : Europia. Repéré à

https://www.researchgate.net/profile/Andre-Tricot/publication/228604876_Interpreter_les_liens_entre_utilisabilite_et_utilite_des_documents_electroniques/links/00b495290634bd8099000000/Interpreter-les-liens-entre-utilisabilite-et-utilite-des-documents-electroniques.pdf

Tricot, A., Plégat-Soutjis, F., Camps, J. F., Amiel, A., Lutz, G., & Morcillo, A. (2003). Utilité, utilisabilité, acceptabilité : interpréter les relations entre trois dimensions de l'évaluation des EIAH. In *Environnements Informatiques pour l'Apprentissage Humain (Strasbourg, avril 2003, pp. 391-402*).

Vallerand, R. J., Blais, M. R., Brière, N. M., & Pelletier, L. G. (1989). Construction et validation de l'échelle de motivation en éducation (EME). *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 21*(3), 323. https://doi.org/10.1037/h0079855

Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly, 27*(3), 425-478. https://doi.org/10.2307/30036540

Ventura, S. (2013, 24 décembre). Presence questionnaire. [Billet de blog]. Repéré à https://marketinginvolvement.files.wordpress.com/2013/12/pq-presence-questionnaire.pdf

Verleger, M. & Pembridge, J. (2018). A pilot study integrating an AI-driven chatbot in an introductory programming course. In *IEEE Frontiers in Education Conference (2018)* https://doi.org/10.1109/FIE.2018.8659282

Ward, D. (2005). Why Users Choose Chat: A Survey of Behavior and Motivations. *Internet Reference Services Quarterly*, *10(1)*, 29-46. https://doi.org/10.1300/J136v10n01_03

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, *7*(3), 225-240. https://doi.org/10.1162/105474698565686

# 10 Annexes

## Annexe 1 - Tableau récapitulatif des thèmes, méta critères, critères, définitions et exemples de questions associées au modèle d'évaluation proposé par Hu, Trigano & Crozat (2001)

| Thème : Ergonomie | | | |
|---|---|---|---|
| **Méta Critère** | **Définition** | **Questions** | **Critères** |
| **Guidage** | « Mesure la qualité des caractéristiques de l'interface permettant d'identifier son fonctionnement. » | • Lorsqu'il faut réaliser une action spécifique, est-ce que le système l'indique à l'utilisateur ?<br>• Vous est-il arrivé de devoir chercher des composants à l'écran, parce qu'il n'avait pas l'aspect ou le format attendu ? | • Incitation<br>• Localisation<br>• Format<br>• Retour |
| **Charge de travail** | « Évalue la pertinence des efforts cognitifs que l'utilisateur doit mettre en œuvre pour accomplir sa tâche. » | • Trouvez-vous que l'écran soit surchargé du point de vue des informations et des possibilités d'actions ?<br>• Avez-vous trouvé que certaines suites d'actions, pour atteindre un but, étaient trop longues ou trop compliquées ? | • Concision<br>• Actions minimales |
| **Manipulation** | « Concerne la qualité de la maîtrise par l'utilisateur du déroulement du logiciel. » | • L'utilisateur a-t-il la possibilité d'interrompre un traitement parce qu'il est trop long par exemple ?<br>• Est-il arrivé que le logiciel entreprenne une action que vous n'aviez pas explicitement demandée ? | • Prévention<br>• Contrôle<br>• Correction |
| **Homogénéité** | « Évalue la cohérence et le maintien tout au long du logiciel d'une charte graphique et fonctionnelle. » | • Est-ce que les caractéristiques de l'interface sont conservées pour des contextes d'utilisation identiques ?<br>• Un même intitulé ou icône conduit-il toujours au même comportement ? | • Graphique<br>• Fonctionnelle |
| **Adaptabilité** | « Concerne l'ensemble des possibilités disponibles pour modifier l'aspect ou les modes d'interaction de l'interface afin d'en améliorer l'utilisation. » | • Le format des textes affichés est-il modifiable ?<br>• L'utilisateur, en fonction de son expérience, peut-il accéder à des fonctions de manipulation avancées ? | • Flexibilité<br>• Expérience |

| Thème : Documents multimédias | | | |
|---|---|---|---|
| **Méta Critère** | **Définition** | **Questions** | **Critères** |
| **Textes** | « Mesure la qualité des choix et des formats des documents basés sur le texte écrit. » | • Le vocabulaire employé est-il familier pour l'utilisateur ?<br>• Les choix de couleurs pour le texte et le fond optimisent-ils la lisibilité ? | • Rédaction<br>• Mise en page<br>• Typographie |
| **Documents visuels** | « Évalue la qualité et la pertinence des documents visuels (ensemble des documents basés sur les images et les représentations graphiques, fixes ou animées). » | • Est-ce que le contenu des images peut entraîner des ambiguïtés ?<br>• Quelle est la qualité de la définition des photographies et vidéos ? | • Images didactiques<br>• Images réelles<br>• Images artificielles |
| **Documents sonores** | « Évalue la qualité et pertinence des voix, des musiques et des bruitages. » | • Le didacticiel a-t-il recours à l'utilisation de la musique dans un but pertinent (détente, transition...) ?<br>• Les bruitages sont-ils mis à profit pour remplir des fonctions de signalisation ou de localisation d'actions ou d'événements ? | • Paroles<br>• Bruitages<br>• Musique |
| **Interdépendance** | « Concerne la qualité d'agencement des différents médias. » | • L'utilisation simultanée de plusieurs médias gêne-t-elle la compréhension ?<br>• Des légendes sont-elles utilisées pour résumer le sens des images ? | • Choix<br>• Concurrence<br>• Complémentarité<br>• Redondance |

| Thème : Scénarisation | | | |
|---|---|---|---|
| **Méta Critère** | **Définition** | **Questions** | **Critères** |
| **Structure** | « Concerne l'organisation des informations dans le didacticiel. » | • La structure des liens qui la composent suit-elle une logique propre ?<br>• Le didacticiel offre-t-il des liens vers Internet ? | • Interne<br>• Externe |
| **Outil de navigation** | « Évalue la qualité des outils permettant à l'utilisateur d'appréhender la structure. » | • Le logiciel propose-t-il à l'utilisateur des outils lui permettant de savoir où il en est et ce qui lui reste à faire ?<br>• L'utilisateur a-t-il la possibilité d'annoter les documents auxquels il accède ? | • Parcours<br>• Construction |
| **Fiction** | « Mesure la qualité et la pertinence de l'utilisation d'une dimension imaginaire dans le parcours du didacticiel ou dans la présentation des contenus. » | • Le récit permet-il de structurer l'apprentissage ?<br>• L'utilisation du didacticiel est-elle génératrice d'émotion ? | • Récit<br>• Ambiance<br>• Personnages |

| Thème : Outils pédagogiques | | | |
|---|---|---|---|
| **Méta Critère et définition** | **Critères** | **Définition** | **Questions** |
| **Situation**<br><br>« Correspond à la description et à la mesure de pertinence des situations techniques (communication, travail de groupe, etc.) prévues par le didacticiel. » | **Communication** | Caractérise et mesure la qualité et la richesse des outils de communication utilisés ou intégrés par le didacticiel. | • Quelles sont les possibilités de communication entre les apprenants ?<br>• Caractériser la synchronisation des outils de communication proposés aux apprenants. |
| | **Encadrement** | Caractérise et évalue la qualité des outils d'encadrement dont dispose le tuteur au sein du didacticiel. | • Quel est le type d'encadrement privilégié par le didacticiel ?<br>• Le tuteur peut-il diriger l'utilisation du logiciel à distance ? |
| | **Temps** | Caractérise et mesure la qualité de prise en compte des contraintes temporelles des apprenants. | • Quel est le temps standard d'une session du didacticiel ?<br>• Le didacticiel permet-il de gérer le temps écoulé entre les sessions ? |
| | **Rapport aux apprenants** | Caractérise les possibilités d'utilisation du logiciel en termes de manipulation de groupe. | • Le produit est-il prévu pour être exploité par plusieurs personnes sur une même interface ?<br>• Le logiciel est-il approprié à l'utilisation coopérative par plusieurs utilisateurs sur des machines différentes ? |
| | **Rapports aux supports** | Mesure la souplesse d'utilisation et la complémentarité du didacticiel avec d'autres supports. | • Le didacticiel peut-il être utilisé de façon autonome, sans autre support de formation ?<br>• Le didacticiel peut-il s'articuler avec d'autres supports de formation ? |
| **Contenu**<br><br>« Mesure la validité et la qualité du contenu. » | **Validité** | Mesure la qualité a priori et la pertinence du contenu vis-à-vis des objectifs affichés par le didacticiel. | • Relève-t-on des erreurs dans le contenu des informations présentées à l'apprenant ?<br>• Le degré d'approfondissement du contenu est-il pertinent par rapport au niveau des apprenants ? |
| | **Impact social** | Concerne le respect des critères socioculturels. | • Le didacticiel respecte-t-il le statut des individus évoqués (âge, sexe, etc.) ? |

| | | | |
|---|---|---|---|
| | | | • Le didacticiel respecte-t-il les cultures et les coutumes de la population cible ? |
| **Personnalisation**<br><br>« Évalue la qualité et la richesse des outils permettant une personnalisation de l'apprentissage. » | **Cadrage** | Concerne la description de la population cible du didacticiel. | • Des informations concernant les prérequis et le niveau des apprenants sont-elles disponibles ?<br>• Des informations concernant les objectifs de formation sont-elles disponibles ? |
| | **Paramétrage** | Mesure la richesse et la pertinence des moyens mis à la disposition des utilisateurs pour adapter le didacticiel à son contexte d'utilisation. | • Le tuteur dispose-t-il de paramètres lui permettant d'adapter le contenu à son enseignement ?<br>• Le didacticiel permet-il de respecter le rythme de chaque apprenant à travers la gestion des sessions ? |
| | **Adaptation** | Concerne la capacité d'adaptation dynamique du didacticiel en fonction du déroulement de l'apprentissage. | • Le didacticiel s'adapte-t-il en fonction du parcours au sein des connaissances ?<br>• Le didacticiel s'adapte-t-il en fonction des résultats d'évaluation ? |
| | **Suivi** | Caractérise et évalue les outils permettant un traitement dans le temps des informations. | • Le didacticiel gère-t-il un historique de l'apprentissage ?<br>• Le didacticiel permet-il d'analyser les données de suivi d'un groupe d'apprenants ? |
| **Activité**<br><br>« Évalue la pertinence et la richesse des activités proposées aux apprenants. » | **Exposition** | Qualité de la présentation des contenus : multiplication des points de vue, approfondissement, mise en scène, etc. | • Est-ce qu'un même contenu est proposé selon des points de vue variés ?<br>• Est-ce que les contenus peuvent être plus ou moins approfondis ? |
| | **Résolution** | Richesse et pertinence des activités de résolution (exercice, problème, etc.) proposées : niveau de difficulté, variété, etc. | • Est-ce que le didacticiel propose des exercices de natures et déformés différentes ?<br>• Le didacticiel propose-t-il des exercices de difficultés variées ? |
| | **Expérimentation** | Richesse et pertinence des activités de manipulation (TP, | • Le didacticiel stimule-t-il la créativité et l'activité des utilisateurs ? |

| | | | |
|---|---|---|---|
| | | expérimentation, création, etc.) proposées : degrés de liberté, variété, etc. | • Le didacticiel permet-il d'assister à des expériences ou à des cas réels ? |
| | **Récréation** | Qualité des activités ludiques ou ludo-éducatives proposées. | • Le didacticiel propose-t-il des activités récréatives ou ludiques aux apprenants ?<br>• Quelle est la qualité des activités récréatives en termes d'apprentissage par le jeu ? |
| **Assistance**<br><br>« Caractérise et évalue la qualité et la richesse des outils d'aide pédagogique proposés à l'apprenant. » | **Accès** | Caractérise et mesure la pertinence des moyens d'accès à l'aide mis en place dans le logiciel. | • Le didacticiel propose-t-il une aide spontanément ?<br>• Le didacticiel permet-il de fournir une aide correspondant à une requête de l'utilisateur ? |
| | **Ressources matérielles** | Concerne la richesse et la qualité des aides méthodologiques apportées : approfondissement pas à pas, description de méthodes, etc. | • Le didacticiel propose-t-il des renvois vers le contenu théorique depuis les exercices ou expérimentations ?<br>• Le didacticiel propose-t-il une ou plusieurs aides méthodologiques pour aider l'apprenant à traiter ses problèmes ? |
| | **Ressources humaines** | Caractérise et évalue la pertinence des moyens d'accès à des personnes ressources. | • Le didacticiel prévoit-il de faire appel à des compétences extérieures ?<br>• Le didacticiel propose-t-il à l'enseignant d'intégrer ses interventions dans le parcours du logiciel ? |
| **Évaluation**<br><br>« Caractérise et évalue la qualité et la bonne adéquation des outils et des dimensions d'évaluation. » | **Technique** | Ce critère caractérise les types d'évaluation mis en place dans le logiciel : auto-évaluation, contrôle continu, etc. | • Le didacticiel permet-il à l'apprenant de s'auto-évaluer ?<br>• Le didacticiel prévoit-il l'intervention du tuteur dans les phases d'évaluation ? |
| | **Critère** | Ce critère décrit et évalue la pertinence des dimensions évaluées par le logiciel : mémoire, rapidité, confiance en soi, etc. | • Quelle est la diversité des caractéristiques évaluées chez l'apprenant ?<br>• Les dimensions évaluées par le didacticiel sont-elles en accord avec les objectifs d'apprentissage ? |

| | | | |
|---|---|---|---|
| | **Positionnement** | Ce critère caractérise les possibilités de positionnement des connaissances de l'apprenant par rapport à lui-même ou par rapport aux niveaux institutionnels. | • Le didacticiel est-il en mesure de fournir un bilan de compétence ?<br>• Le didacticiel donne-t-il la possibilité de se positionner par rapport à un niveau académique ? |

| **Thème : Impressions générales** | | | | | |
|---|---|---|---|---|---|
| Rassurant<br>Déroutant | Luxuriant<br>Dépouillé | Ludique<br>Académique | Actif<br>Passif | Simple<br>Complexe | Innovant<br>Traditionnel |

## Annexe 2 - Items de la version française de l'AttrakDiff selon Lallemand et al. (2015)

| | |
|---|---|
| Qualité pragmatique | Technique – Humain |
| | Compliqué – Simple |
| | Pas pratique – Pratique |
| | Fastidieux – Efficace |
| | Imprévisible – Prévisible |
| | Confus – Clair |
| | Incontrôlable – Maîtrisable |
| Qualité hédonique-identification | M'isole – Me sociabilise |
| | Amateur – Professionnel |
| | De mauvais goût – De bon goût |
| | Bas de gamme – Haut de gamme |
| | M'exclut – M'intègre |
| | Me sépare des autres – Me rapproche des autres |
| | Non présentable – Présentable |
| Qualité hédonique-stimulation | Conventionnel – Original |
| | Sans imagination – Créatif |
| | Prudent – Audacieux |
| | Conservateur – Novateur |
| | Ennuyeux – Captivant |
| | Peu exigeant – Challenging |
| | Commun – Nouveau |
| Attractivité | Déplaisant – Plaisant |
| | Laid – Beau |
| | Désagréable – Agréable |
| | Rebutant – Attirant |
| | Mauvais – Bon |
| | Repoussant – Attrayant |
| | Décourageant – Motivant |

## Annexe 3 - Version française de l'UEQ+ de Schrepp & Thomaschewski (2019)

| **Attractivité** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| De manière générale, je trouve le produit... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Fâcheuse | | | | | | | | Réjouissante |
| Mauvaise | | | | | | | | Bonne |
| Désagréable | | | | | | | | Agréable |
| Pas sympathique | | | | | | | | Sympathique |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

| **Efficacité** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Pour atteindre mes objectifs, je trouve le produit... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Lent | | | | | | | | Rapide |
| Inefficace | | | | | | | | Efficace |
| Pas pratique | | | | | | | | Pratique |
| Surchargé | | | | | | | | Ordonné |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

| **Transparence** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Je trouve l'utilisation du produit... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Incompréhensible | | | | | | | | Compréhensible |
| Difficile à apprendre | | | | | | | | Facile à apprendre |
| Compliquée | | | | | | | | Facile |
| Confuse | | | | | | | | Claire |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

| **Maniabilité** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Je trouve la réaction du produit à mes saisies et instructions... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Imprévisible | | | | | | | | Prévisible |
| Gênante | | | | | | | | Utile |
| Peu sûre | | | | | | | | Sûr |
| Pas conforme aux attentes | | | | | | | | Conforme aux attentes |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

| **Stimulation** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Je trouve l'occupation ou le travail avec le produit... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Pas intéressant | | | | | | | | Intéressant |
| Ennuyeux | | | | | | | | Passionnant |
| De mauvaise qualité | | | | | | | | Précieux |
| Soporifique | | | | | | | | Stimulant |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

| **Originalité** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Je trouve l'idée ou la conception du produit... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Dépourvue d'imagination | | | | | | | | Créative |
| Conventionnelle | | | | | | | | Originale |
| Conventionnelle | | | | | | | | Inédite |
| Conservatrice | | | | | | | | Innovatrice |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

| Confiance | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Concernant l'utilisation de mes informations et données personnelles, le produit est... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Peu sûr | | | | | | | | Sûr |
| Pas sérieux | | | | | | | | Sérieux |
| Non fiable | | | | | | | | Fiable |
| Non transparent | | | | | | | | Transparent |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

| Toucher | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Je trouve la surface du produit... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Instable | | | | | | | | Stable |
| Désagréable au toucher | | | | | | | | Agréable au toucher |
| Rugueuse | | | | | | | | Lisse |
| Glissante | | | | | | | | Antidérapante |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

| Acoustique | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Les bruits émis lors du fonctionnement du produit sont... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Bruyants | | | | | | | | Silencieux |
| Cacophoniques | | | | | | | | Harmonieux |
| Bourdonnants | | | | | | | | Doux et silencieux |
| Stridents | | | | | | | | Doux |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

| **Flexibilité** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Concernant mes exigences et préférences personnelles, le produit est... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Pas personnalisable | | | | | | | | Personnalisable |
| Non modifiable | | | | | | | | Modifiable |
| Rigide | | | | | | | | Flexible |
| Non extensible | | | | | | | | Extensible |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

| **Utilité** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Je trouve la possibilité d'utiliser le produit... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Non bénéfique | | | | | | | | Bénéfique |
| Pas utile | | | | | | | | Utile |
| Non avantageuse | | | | | | | | Avantageuse |
| Non rentable | | | | | | | | Rentable |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

| **Valeur** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| La conception du produit me paraît de manière générale... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| De mauvaise qualité | | | | | | | | Précieuse |
| Non présentable | | | | | | | | Présentable |
| Pas de bon goût | | | | | | | | De bon goût |
| Non élégante | | | | | | | | Élégante |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

| **Esthétique visuelle** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Je trouve la conception visuelle du produit... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Affreuse | | | | | | | | Belle |
| Sans style | | | | | | | | Ayant du style |
| Pas plaisante | | | | | | | | Plaisante |
| Inesthétique | | | | | | | | Esthétique |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

| **Utilisation intuitive** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| L'utilisation du produit me paraît... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Laborieuse | | | | | | | | Aisée |
| Non logique | | | | | | | | Logique |
| Pas évidente | | | | | | | | Évidente |
| Incohérente | | | | | | | | Cohérente |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

| **Crédibilité** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Les informations et données que le produit met à ma disposition sont... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Non bénéfiques | | | | | | | | Bénéfiques |
| Invraisemblables | | | | | | | | Vraisemblables |
| Pas sérieuses | | | | | | | | Sérieuses |
| Imprécises | | | | | | | | Précises |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

| **Qualité du contenu** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Les informations et données que le produit met à ma disposition sont... | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Désuet | | | | | | | | Actuel |
| Pas intéressant | | | | | | | | Intéressant |
| Mal préparé | | | | | | | | Bien préparé |
| Incompréhensible | | | | | | | | Compréhensible |
| Je considère la propriété de produit décrite par ces termes... | | | | | | | | |
| Totalement sans importance | | | | | | | | Très importante |

## Annexe 4 - Items du modèle UTAUT et leur traduction (Venkatesh et al., 2003)

| Items originaux UTAUT de Venkatesh & al. (2003) | | Traduction libre |
|---|---|---|
| Performances attendues | I would find the system useful in my job. | Je trouve l'utilisation du chatbot utile dans mon apprentissage. |
| | Using the system enables me to accomplish tasks more quickly. | Utiliser le chatbot me permet d'accomplir les tâches plus rapidement. |
| | Using the system increases my productivity. | Utiliser le chatbot augmente ma productivité. |
| Efforts attendus | If I use the system, I will increase my chances of getting a raise. | Si j'utilise le chatbot, je vais augmenter mes chances d'avoir une augmentation. |
| | My interaction with the system would be clear and understandable. | Mon interaction avec le chatbot est claire et compréhensible. |
| | It would be easy for me to become skillful at using the system. | C'est facile pour moi de devenir expérimenté en utilisant le chatbot. |
| | I would find the system easy to use. | Je trouve que le chatbot est facile à utiliser. |
| | Learning to operate the system is easy for me. | Apprendre à utiliser le chatbot est facile pour moi. |
| Influence sociale | Using the system is a bad/good idea. | L'utilisation du chatbot est une bonne idée. |
| | The system makes work more interesting. | Le chatbot rend l'apprentissage plus intéressant. |
| | Working with the system is fun. | Travailler avec le chatbot est amusant. |
| | I like working with the system. | J'aime utiliser le chatbot. |
| | People who influence my behavior think that I should use the system. | Des gens qui influencent mon comportement pensent que je devrais utiliser le système. |
| | People who are important to me think that I should use the system. | Les gens qui sont importants pour moi pensent que je devrais utiliser le chatbot. |
| Conditions de facilitation | A specific person (or group) is available for assistance with system difficulties. | Une personne spécifique (ou un groupe) est disponible pour apporter de l'aide en cas de difficultés avec le chatbot. |
| | In general, the organization has supported the use of the system. | En général, l'institution encourage l'utilisation du chatbot. |

| | | |
|---|---|---|
| Efficacité personnels | I have the resources necessary to use the system. | J'ai les ressources nécessaires pour utiliser le chatbot. |
| | I have the knowledge necessary to use the system. | J'ai les connaissances nécessaires pour utiliser le chatbot. |
| | The system is not compatible with other systems I use. | Le chatbot n'est pas compatible avec d'autres systèmes que j'utilise. |
| Anxiété | I could complete a job or task using the system... | Je pourrais réaliser une tâche en utilisant le chatbot... |
| | If there was no one around to tell me what to do as I go. | S'il n'y a personne autour de moi pour me dire quoi faire. |
| | If I could call someone for help if I got stuck. | Si je peux appeler quelqu'un à l'aide si je suis bloqué. |
| | If I had a lot of time to complete the job for which the software was provided. | Si j'ai beaucoup de temps à disposition pour réaliser la tâche à l'aide du chatbot. |
| | If I had just the built-in help facility for assistance. | Si je n'ai que l'assistance prédéfinie à disposition. |
| | I feel apprehensive about using the system. | J'appréhende l'utilisation du chatbot. |
| | It scares me to think that I could lose a lot of information using the system by hitting the wrong key. | J'ai peur de penser que je peux perdre un grand nombre d'informations en utilisant le système ou en cliquant sur un mauvais bouton. |
| | I hesitate to use the system for fear of making mistakes I cannot correct. | J'hésite à utiliser le système par peur de faire des erreurs que je ne peux pas corriger. |
| | The system is somewhat intimidating to me. | L'utilisation du chatbot m'intimide un peu. |
| Intention d'utiliser le système | I intend to use the system in the next <n> months. | J'ai l'intention d'utiliser le chatbot ces prochains mois. |
| | I predict I would use the system in the next <n> months. | Je pense que je vais utiliser le chatbot ces prochains mois. |
| | I plan to use the system in the next <n> months. | Je planifie d'utiliser le chatbot ces prochains mois. |

## Annexe 5 - Échelle utilisée avec les apprenants (Caron & Heutte, 2017)

| Échelle utilisée avec les apprenants (Caron et Heutte, 2017) | | | |
|---|---|---|---|
| 01 | a_UTG_1 | L'utilisation du [dispositif instrumenté] dans mes études me permet d'accomplir des tâches plus rapidement. | 1 2 3 4 5 6 7 |
| 02 | a_UTF_2 | Je pense que le [dispositif instrumenté] est utile pour former. | 1 2 3 4 5 6 7 |
| 03 | a_UTI_1 | J'utilise le [dispositif instrumenté] parce que l'on me le demande. | 1 2 3 4 5 6 7 |
| 04 | a_UTE_3 | Le [dispositif instrumenté] améliore ma capacité à appliquer la formation reçue. | 1 2 3 4 5 6 7 |
| 05 | a_UTF_1 | Avec le [dispositif instrumenté] je pense que c'est plus facile pour former. | 1 2 3 4 5 6 7 |
| 06 | a_UTB_2 | C'est facile d'utiliser le [dispositif instrumenté] pour faire ce que je veux faire. | 1 2 3 4 5 6 7 |
| 07 | a_UTE_1 | Le [dispositif instrumenté] me permet d'apprendre plus rapidement. | 1 2 3 4 5 6 7 |
| 08 | a_UTI_3 | J'utilise le [dispositif instrumenté] parce que j'espère de son usage une certaine reconnaissance au sein de mon établissement. | 1 2 3 4 5 6 7 |
| 09 | a_UTB_1 | C'est facile d'apprendre à utiliser le [dispositif instrumenté] | 1 2 3 4 5 6 7 |
| 10 | a_UTE_4 | Avec le [dispositif instrumenté] c'est plus facile d'apprendre. | 1 2 3 4 5 6 7 |
| 11 | a_UTG_3 | Je trouve le [dispositif instrumenté] utile pour mon projet d'étude. | 1 2 3 4 5 6 7 |
| 12 | a_UPM_2 | Le [dispositif instrumenté] m'a permis de mieux mettre en œuvre mon projet d'étude. | 1 2 3 4 5 6 7 |
| 13 | a_UTB_3 | Je trouve le [dispositif instrumenté] facile à utiliser. | 1 2 3 4 5 6 7 |
| 14 | a_UTE_2 | Avec le [dispositif instrumenté] j'ai plus de facilité pour accéder à la formation reçue. | 1 2 3 4 5 6 7 |
| 15 | a_UTI_2 | J'utilise le [dispositif instrumenté] parce que les autres apprenants l'utilisent. | 1 2 3 4 5 6 7 |
| 16 | a_UTG_2 | L'utilisation du [dispositif instrumenté] me rend plus efficace. | 1 2 3 4 5 6 7 |
| 17 | a_UTE_5 | Le [dispositif instrumenté] est utile pour mon apprentissage. | 1 2 3 4 5 6 7 |
| 18 | a_UPM_1 | Avec le [dispositif instrumenté] je comprends mieux mon projet d'étude. | 1 2 3 4 5 6 7 |

**Dimensions :**

- Utilité de façon générale : a_UTG 1, a_UTG_2, a_UTG_3
- Utilité projetée sur le formateur pour enseigner : a_UTF_1, a_UTF_2
- Utilité apprenant : a_UTE_1, a_UTE_2, a_UTE_3, a_UTE 4, a_UTE 5
- Utilisabilité : a_UTB_1, a_UTB_2, a_UTB_3
- Injonction : a_UTI_1, a_UTI_2, a_UTI_3
- Utilité pour moi même : a_UPM_1, a_UPM_2

## Annexe 6 - Échelle de motivation en formation des adultes (EMFA) de Fenouillet, Heutte & Vallerand (2015)

| Pourquoi suivez-vous cette formation ? | |
|---|---|
| **Items** | **Dimension de la motivation** |
| 1. Parce qu'elle va me permettre de gagner davantage | Motivation extrinsèque - régulation externe |
| 2. Parce que ce type de formation fait partie intégrante de moi. | Motivation extrinsèque – Intégrée |
| 3. Parce que j'éprouve du plaisir et de la satisfaction à apprendre de nouvelles choses. | Motivation intrinsèque à la connaissance |
| 4. Parce que selon moi elle va m'aider dans la poursuite de ma carrière (ou pour ma future carrière). | Motivation extrinsèque – identifiée |
| 5. Honnêtement, je ne le sais pas ; j'ai vraiment l'impression de perdre mon temps en formation. | Amotivation |
| 6. Pour me prouver à moi-même que je suis capable de suivre cette formation. | Motivation extrinsèque – introjectée |
| 7. Pour pouvoir décrocher un emploi plus important. | Motivation extrinsèque - régulation externe |
| 8. Pour le plaisir que j'ai à découvrir de nouvelles choses jamais vues auparavant. | Motivation intrinsèque à la connaissance |
| 9. Parce qu'éventuellement cela va me permettre de travailler dans un domaine que j'aime. | Motivation extrinsèque – identifiée |
| 10. J'ai déjà eu de bonnes raisons de suivre cette formation, mais maintenant je me demande si je devrais continuer à y aller. | Amotivation |
| 11. Parce que cette formation me ressemble. | Motivation extrinsèque – Intégrée |
| 12. Parce que le fait de réussir cette formation me permet de me sentir important à mes propres yeux. | Motivation extrinsèque – introjectée |
| 13. Parce que je veux pouvoir faire «la belle vie » plus tard. | Motivation extrinsèque - régulation externe |
| 14. Pour le plaisir d'en savoir plus long sur des sujets qui m'intéressent. | Motivation intrinsèque à la connaissance |
| 15. Parce que cela va m'aider dans le développement de ma carrière professionnelle. | Motivation extrinsèque – identifiée |
| 16. Je ne parviens pas à voir pourquoi je suis cette formation et franchement je m'en fous pas mal. | Amotivation |
| 17. Parce que cette formation est une partie de ce que je suis. | Motivation extrinsèque – Intégrée |
| 18. Pour me prouver que je suis une personne intelligente. | Motivation extrinsèque – introjectée |

| | |
|---|---|
| 19. Pour avoir un meilleur salaire plus tard. | Motivation extrinsèque - régulation externe |
| 20. Parce que cette formation va me permettent de continuer à en apprendre sur une foule de choses qui m'intéressent. | Motivation intrinsèque à la connaissance |
| 21. Parce que je crois que cette formation va augmenter mon potentiel au travail. | Motivation extrinsèque – identifiée |
| 22. Je ne le sais pas ; je ne parviens pas à comprendre pourquoi. | Amotivation |
| 23. Parce que je veux me prouver à moi-même que je suis capable de réussir cette formation. | Motivation extrinsèque – introjectée |
| 24. Parce que cette formation est une expression de moi-même | Motivation extrinsèque – Intégrée |

## Annexe 7 - Items originaux du SUS de Brooke (1996) et leur traduction française (Gronier & Baudet, 2021)

| Items originaux du SUS de Brooke (1996) tirés de (Bangor et al., 2008) | Traduction (Gronier & Baudet, 2021) |
|---|---|
| I think that I would like to use this system frequently | Je voudrais utiliser ce système fréquemment. |
| I found the system unnecessarily complex | Ce système est inutilement complexe. |
| I thought the system was easy to use | Ce système est facile à utiliser. |
| I think that I would need the support of a technical person to be able to use this system | J'aurais besoin du soutien d'un technicien pour être capable d'utiliser ce système. |
| I found that the various functions in this system were well integrated | Les différentes fonctionnalités de ce système sont bien intégrées. |
| I thought that there was too much inconsistency in this system | Il y a trop d'incohérences dans ce système. |
| I would imagine that most people would learn to use this system very quickly | La plupart des gens apprendront à utiliser ce système très rapidement. |
| I found the system very cumbersome to use | Ce système est très lourd à utiliser. |
| I felt very confident using the system | Je me suis senti·e très en confiance en utilisant ce système. |
| I need to learn a lot of things before I could get going with this system | J'ai eu besoin d'apprendre beaucoup de choses avant de pouvoir utiliser ce système. |

## Annexe 8 - Items originaux du TAM (Davis, 1989) et leur traduction par Huu Binh (2014)

| Items originaux du TAM de Davis (1989) | | Traduction (Huu Binh, 2014) |
|---|---|---|
| *Utilité perçue* | Using [the chatbot] in my job would enable me to accomplish tasks more quickly. | Utiliser [le chatbot]me permettra d'accomplir les tâches plus rapidement. |
| | Using [the chatbot] would improve my job performance | L'utilisation [du chatbot] augmentera mes performances. |
| | Using [the chatbot] at work would increase my productivity. | L'utilisation [du chatbot] augmentera ma productivité. |
| | Using [the chatbot] would enhance my effectiveness in my job. | L'utilisation [du chatbot] améliorera mon efficacité. |
| | Using [the chatbot] would make it easier to do my job. | Utiliser [le chatbot] faciliterait mon travail. |
| | I would find [the chatbot] useful in my job. | Je trouve que [le chatbot] est un outil utile dans mon travail. |
| *Facilité d'utilisation perçue* | Learning to operate [the chatbot] would be easy for me. | Apprendre à utiliser [le chatbot] sera facile pour moi. |
| | I would find it easy to get [the chatbot] to do what I want it to do. | Je trouve qu'il est facile de demander [au chatbot] de faire ce que je veux. |
| | My interaction with [the chatbot] would be clear and understable. | Mon interaction avec [le chatbot] est claire et compréhensible. |
| | I would find [the chatbot] to be flexible to interact with. | Je trouve [le chatbot] flexible dans nos interactions. |
| | It would be easy for me to become skillful at using [the chatbot]. | Il me sera facile de devenir habile dans l'utilisation [du chatbot]. |
| | I would find the technology easy to use. | Je trouve que [le chatbot] est facile à utiliser. |

## Annexe 9 - Critères et items du questionnaire USE (Lund, 2001) et leur traduction par (Monfort et al., 2018)

| Critères et items (Lund, 2001) | | Traduction (Monfort & al., 2018) | |
|---|---|---|---|
| Usefulness | ● It helps me be more effective.<br>● It helps me be more productive.<br>● It is useful.<br>● It gives me more control over the activities in my life.<br>● It makes the things I want to accomplish easier to get done.<br>● *It saves me time when I use it.*<br>● *It meets my needs.*<br>● *It does everything I would expect it to do.* | Utilité | ● Il m'aide à être plus efficace.<br>● Il m'aide à être plus productif.<br>● Il est utile.<br>● Il me permet d'avoir plus de contrôle sur les activités de ma vie quotidienne.<br>● Il me permet de faire plus facilement ce que je veux faire.<br>● *Il me fait gagner du temps quand je l'utilise.*<br>● *Il répond à mes besoins.*<br>● *Il fait tout ce que je veux qu'il fasse.* |
| Ease of Use | ● It is easy to use.<br>● It is simple to use.<br>● It is user friendly.<br>● It requires the fewest steps possible to accomplish what I want to do with it.<br>● *It is flexible.*<br>● *Using it is effortless.*<br>● *I can use it without written instructions.*<br>● *I don't notice any inconsistencies as I use it.*<br>● *Both occasional and regular users would like it.*<br>● *I can recover from mistakes quickly and easily.*<br>● *I can use it successfully every time.* | Utilisabilité | ● Il est facile à utiliser.<br>● Il est simple à utiliser.<br>● Il est convivial.<br>● Il nécessite le moins d'étapes possibles pour accomplir ce que je veux faire avec lui.<br>● *Il est flexible.*<br>● *L'utiliser ne nécessite pas d'effort.*<br>● *Je peux l'utiliser sans instruction écrite.*<br>● *Je ne rencontre pas d'incohérence quand je l'utilise.*<br>● *Il convient à la fois aux utilisateurs réguliers et aux utilisateurs occasionnels.*<br>● *Je peux corriger mes erreurs rapidement et facilement.*<br>● *Je peux l'utiliser avec succès à chaque fois.* |
| Ease of Learning | ● I learned to use it quickly.<br>● I easily remember how to use it.<br>● It is easy to learn to use it.<br>● *I quickly became skillful with it.* | Facilité d'apprentissage | ● J'ai appris à l'utiliser facilement<br>● Je me rappelle facilement comment l'utiliser.<br>● Il est facile d'apprendre à l'utiliser.<br>● *Je suis rapidement devenu habile avec lui.* |
| Satisfaction | ● I am satisfied with it.<br>● I would recommend it to a friend.<br>● It is fun to use.<br>● It works the way I want it to work.<br>● It is wonderful.<br>● *I feel I need to have it.*<br>● *It is pleasant to use.* | Satisfaction | ● J'en suis satisfait.<br>● Je le recommanderais à un ami.<br>● Il est amusant à utiliser.<br>● Il fonctionne comme je veux qu'il fonctionne.<br>● Il est merveilleux.<br>● *J'ai l'impression que j'ai besoin d'en avoir un.*<br>● *Il est agréable à utiliser.* |

## Annexe 10 - Échelle SAM de Bradley & Lang (1994)

| Facteurs | Items | Traduction libre |
|---|---|---|
| **Pleasure** | Unhappy - Happy | Malheureux - Heureux |
| | Unsatisfied - Satisfied | Insatisfait - Satisfait |
| **Arousal** | Sluggish - Frenzied | Lenteur/Léthargique - Frénétique/Frénésie/Rapidité |
| **Dominance** | Guided - Autonomous | Guidé - Autonome |

## Annexe 11 - Computer Emotion Scale (Kay & Loverock, 2008)

| In general, when I am learning how to use a new software package, I feel | | | | | |
|---|---|---|---|---|---|
| | | None of the time | Some of the time | Most of the time | All of the time |
| 1 | Satisfied *a* | 0 | 1 | 2 | 3 |
| 2 | Disheartened *b* | 0 | 1 | 2 | 3 |
| 3 | Anxious *c* | 0 | 1 | 2 | 3 |
| 4 | Irritable *d* | 0 | 1 | 2 | 3 |
| 5 | Excited *a* | 0 | 1 | 2 | 3 |
| 6 | Dispirited *b* | 0 | 1 | 2 | 3 |
| 7 | Insecure *c* | 0 | 1 | 2 | 3 |
| 8 | Frustrated *d* | 0 | 1 | 2 | 3 |
| 9 | Curious *a* | 0 | 1 | 2 | 3 |
| 10 | Helpless *c* | 0 | 1 | 2 | 3 |
| 11 | Nervous *c* | 0 | 1 | 2 | 3 |
| 12 | Angry *d* | 0 | 1 | 2 | 3 |
| *a Happiness construct* / *b Sadness construct* / *c Anxiety construct* / *d Anger construct* | | | | | |

| Traduction libre des items et transformation sous forme de paires de mots pour l'évaluation du chatbot |
|---|
| Satisfait - Insatisfait |
| Découragé.e - Confiant/encouragé/Stimulé |
| Anxieux.se/Stressé - Serein/calme/tranquille |
| Irritable/Impatient - Calme |
| Excité.e/Agité - Calme |
| Déptié.e/Déçu - Satisfait/Ravi/Impressionné.e |
| Peu sûr.e de moi |
| Frustré.e - Satisfait |
| Curieux.se |
| Impuissant.e - Capable/Efficace |
| Nerveux.se - Calme/Patient |
| En colère - Impassible/Calme |

## Annexe 12 - Questionnaire sur l'état de présence (Witmer & Singer, 1998)

| Questionnaire sur l'état de présence | | |
|---|---|---|
| **Item** | **Version originale de Witmer & Singer (1998)** | **Traduction française du Laboratoire de Cyberpsychologie de l'Université du Québec (2002)** |
| 1 | How much were you able to control events? | Dans quelle mesure étiez-vous capable de contrôler les événements ? |
| 2 | How responsive was the environment to actions that you initiated (or performed)? | Dans quelle mesure l'environnement était-il réactif (sensible) aux actions que vous y faisiez ? |
| 3 | How natural did your interactions with the environment seem? | Dans quelle mesure vos interactions avec l'environnement vous semblaient-elles naturelles ? |
| 4 | How much did the visual aspects of the environment involve you? | Dans quelle mesure les aspects visuels de l'environnement vous invitaient-ils à vous y impliquer ? |
| 5 | How natural was the mechanism which controlled movement through the environment? | Dans quelle mesure les mécanismes permettant votre mouvement dans l'environnement vous semblaient-ils naturels ? |
| 6 | How compelling was your sense of objects moving through space? | Dans quelle mesure vos sens étaient-ils trompés par le réalisme du mouvement des objets à travers l'espace ? |
| 7 | How much did your experiences in the virtual environment seem consistent with your real-world experiences? | Dans quelle mesure les expériences que vous avez vécues dans l'environnement virtuel ressemblaient-elles à celles de l'environnement réel ? |
| 8 | Were you able to anticipate what would happen next in response to the actions that you performed? | Étiez-vous capable d'anticiper les conséquences des mouvements que vous faisiez ? |
| 9 | How completely were you able to actively survey or search the environment using vision? | Jusqu'à quel point étiez-vous en mesure d'explorer activement l'environnement de façon visuelle ? |
| 10 | How compelling was your sense of moving around inside the virtual environment? | Jusqu'à quel point la sensation de déplacement à l'intérieur de l'environnement virtuel était-elle confondante (réaliste) ? |
| 11 | How closely were you able to examine objects? | À quelle distance pouviez-vous examiner les objets ? |
| 12 | How well could you examine objects from multiple viewpoints? | Jusqu'à quel point pouviez-vous examiner les objets sous différents angles ? |
| 13 | How involved were you in the virtual environment experience? | Jusqu'à quel point étiez-vous impliqué(e) dans l'expérience vécue dans l'environnement virtuel ? |

| | | |
|---|---|---|
| 14 | How much delay did you experience between your actions and expected outcomes? | Jusqu'à quel point avez-vous ressenti un délai séparant vos actions de leurs conséquences ? |
| 15 | How quickly did you adjust to the virtual environment experience? | À quel rythme vous êtes-vous adapté(e) à l'expérience vécue dans l'environnement virtuel ? |
| 16 | How proficient in moving and interacting with the virtual environment did you feel at the end of the experience? | En termes d'interactions et de déplacements dans l'environnement virtuel, jusqu'à quel point vous sentiez-vous compétent(e) à la fin de l'expérience ? |
| 17 | How much did the visual display quality interfere or distract you from performing assigned tasks or required activities? | Jusqu'à quel point la qualité visuelle de l'appareillage graphique vous a-t-elle incommodé(e) dans l'exécution des tâches requises ? |
| 18 | How much did the control devices interfere with the performance of assigned tasks or with other activities? | Dans quelle mesure les mécanismes de contrôle de votre mouvement ont-ils interféré avec l'exécution des tâches requises ? |
| 19 | How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities? | Jusqu'à quel point êtes-vous parvenu(e) à vous concentrer sur les tâches requises plutôt que sur les mécanismes utilisés pour effectuer lesdites tâches ? |
| 20 | How much did the auditory aspects of the environment involve you? | Dans quelle mesure les aspects auditifs de l'environnement vous invitaient-ils à vous y impliquer ? |
| 21 | How well could you identify sounds? | Dans quelle mesure arriviez-vous à identifier correctement les sons produits dans l'environnement ? |
| 22 | How well could you localize sounds? | Dans quelle mesure arriviez-vous à localiser correctement les sons produits dans l'environnement ? |
| 23 | How well could you actively survey or search the virtual environment using touch? | Dans quelle mesure pouviez-vous explorer activement et de façon tactile (par le toucher) l'environnement ? |
| 24 | How well could you move or manipulate objects in the virtual environment? | Jusqu'à quel point pouviez-vous déplacer ou manipuler les objets dans l'environnement virtuel ? |
| Sous-échelle « Réalisme » : Items 3 + 4 + 5 + 6 + 7 + 10 + 13<br>Sous-échelle « Possibilité d'agir » : Items 1 + 2 + 8 + 9<br>Sous-échelle « Qualité de l'interface » : Items (tous inversés) 14 + 17 + 18<br>Sous-échelle « Possibilité d'examiner » : Items 11 + 12 + 19<br>Sous-échelle « Auto-évaluation de la performance » : Items 15 + 16<br>Sous-échelle « Auditif » : Items 20 + 21 + 22<br>Sous-échelle « Haptique » : Items 23 + 24 | | |

## Annexe 13 - Échelle UES (O'Brien, Cairns & Hall, 2018)

| Sous-échelle | Items originaux User Engagement Scale Short Form (UES-SF) | Traduction des items (Fontaine & al., 2019) |
|---|---|---|
| Attention soutenue | I lost myself in this experience. | 1. Cette expérience m'a tellement absorbé que j'en ai perdu la notion du temps. |
| | The time I spent using Application X just slipped away. | 2. Le temps a filé lorsque j'ai utilisé l'Application X. |
| | I was absorbed in this experience. | 3. J'ai été absorbé par cette expérience. |
| Convivialité perçue | I felt frustrated while using this Application X. | 4. J'ai éprouvé de la frustration lors de l'utilisation de l'Application X. |
| | I found this Application X confusing to use. | 5. J'ai trouvé que l'Application X était difficile à utiliser. |
| | Using this Application X was taxing. | 6. L'utilisation de l'Application X était pénible. |
| Attrait esthétique | This Application X was attractive | 7. L'Application X était attrayante. |
| | This Application X was aesthetically appealing | 8. L'Application X était visuellement agréable. |
| | This Application X appealed to my senses. | 9. La présentation visuelle de l'Application X était plaisante à regarder. |
| Gratification | Using Application X was worthwhile | 10. L'utilisation de l'Application X en valait la peine. |
| | My experience was rewarding. | 11. Mon expérience s'est avérée gratifiante. |
| | I felt interested in this experience. | 12. J'ai été intéressé par cette expérience. |

## Annexe 14 - Révision des items par les experts

| Item | Auteur | Sous-critère | Critère | Raison | Commentaire |
|---|---|---|---|---|---|
| Rassurant / Déroutant | Hu, Trigano & Crozat (2001) | Impressions générales | Expérience utilisateur | ok | |
| Luxuriant / Dépouillé | Hu, Trigano & Crozat (2001) | Impressions générales | Expérience utilisateur | Trop vague | |
| Ludique / Académique | Hu, Trigano & Crozat (2001) | Impressions générales | Expérience utilisateur | Ok | |
| Actif / Passif | Hu, Trigano & Crozat (2001) | Impressions générales | Expérience utilisateur | ok | |
| Simple / Complexe | Hu, Trigano & Crozat (2001) | Impressions générales | Expérience utilisateur | Similaire | Similaire à l'item : « Je trouve le chatbot facile à utiliser » de Caron & Heutte (2017) |
| Innovant / Traditionnel | Hu, Trigano & Crozat (2001) | Impressions générales | Expérience utilisateur | ok | |
| Technique – Humain | Hassenzahl et al. (2003) | Qualité pragmatique | Expérience utilisateur | Similaire | Similaire à Naturel / Robotisé |
| Compliqué – Simple | Hassenzahl et al. (2003) | Qualité pragmatique | Expérience utilisateur | Similaire | Similaire à l'item : « Je trouve le chatbot facile à utiliser » de Caron & Heutte (2017) |
| Pas pratique – Pratique | Hassenzahl et al. (2003) | Qualité pragmatique | Expérience utilisateur | Identique | Même item dans UEQ+ |
| Fastidieux – Efficace | Hassenzahl et al. (2003) | Qualité pragmatique | Expérience utilisateur | Similaire | Similaire à « Inefficace - Efficace » du critère Efficacité et performances de Schrepp & Thomaschewski (2019) |
| Imprévisible – Prévisible | Hassenzahl et al. (2003) | Qualité pragmatique | Expérience utilisateur | Trop vague | |
| Confus – Clair | Hassenzahl et al. (2003) | Qualité pragmatique | Expérience utilisateur | Similaire | Similaire à « Surchargé - Ordonné » dans UEQ+ |
| Incontrôlable – Maîtrisable | Hassenzahl et al. (2003) | Qualité pragmatique | Expérience utilisateur | Trop vague | |
| M'isole – Me sociabilise | Hassenzahl et al. (2003) | Qualité hédonique-identification | Expérience utilisateur | Trop vague | |
| Amateur – Professionnel | Hassenzahl et al. (2003) | Qualité hédonique-identification | Expérience utilisateur | Similaire | « Ludique - Académique » semble plus approprié |

| | | | | | |
|---|---|---|---|---|---|
| De mauvais goût – De bon goût | Hassenzahl et al. (2003) | Qualité hédonique-identification | Expérience utilisateur | Trop vague | |
| Bas de gamme – Haut de gamme | Hassenzahl et al. (2003) | Qualité hédonique-identification | Expérience utilisateur | Trop vague | |
| M'exclut – M'intègre | Hassenzahl et al. (2003) | Qualité hédonique-identification | Expérience utilisateur | Trop vague | |
| Me sépare des autres – Me rapproche des autres | Hassenzahl et al. (2003) | Qualité hédonique-identification | Expérience utilisateur | Trop vague | |
| Non présentable – Présentable | Hassenzahl et al. (2003) | Qualité hédonique-identification | Expérience utilisateur | Trop vague | |
| Conventionnel – Original | Hassenzahl et al. (2003) | Qualité hédonique-stimulation | Expérience utilisateur | Identique | Similaire à « Innovant - Traditionnel » dans EMPI<br>Même item chez Schrepp & Thomaschewski (2019) |
| Sans imagination – Créatif | Hassenzahl et al. (2003) | Qualité hédonique-stimulation | Expérience utilisateur | Reformuler | Reformuler en « Dépourvu d'imagination - Créatif » |
| Prudent – Audacieux | Hassenzahl et al. (2003) | Qualité hédonique-stimulation | Expérience utilisateur | Trop vague | |
| Conservateur – Novateur | Hassenzahl et al. (2003) | Qualité hédonique-stimulation | Expérience utilisateur | Similaire | Similaire à « Innovant - Traditionnel » |
| Ennuyeux – Captivant | Hassenzahl et al. (2003) | Qualité hédonique-stimulation | Expérience utilisateur | Similaire | Similaire à « Ennuyeux - Passionnant » |
| Peu exigeant – Challenging | Hassenzahl et al. (2003) | Qualité hédonique-stimulation | Expérience utilisateur | Similaire | Similaire à « Actif - Passif » |
| Commun – Nouveau | Hassenzahl et al. (2003) | Qualité hédonique-stimulation | Expérience utilisateur | Similaire | Similaire à « Innovant - Traditionnel » |
| Déplaisant – Plaisant | Hassenzahl et al. (2003) | Attractivité | Expérience utilisateur | Non évalué | Aspect esthétique pas évalué |
| Laid – Beau | Hassenzahl et al. (2003) | Attractivité | Expérience utilisateur | Non évalué | Aspect esthétique pas évalué |
| Désagréable – Agréable | Hassenzahl et al. (2003) | Attractivité | Expérience utilisateur | Ok | |
| Rebutant – Attirant | Hassenzahl et al. (2003) | Attractivité | Expérience utilisateur | Non évalué | Aspect esthétique pas évalué |
| Mauvais – Bon | Hassenzahl et al. (2003) | Attractivité | Expérience utilisateur | Trop vague | |
| Repoussant – Attrayant | Hassenzahl et al. (2003) | Attractivité | Expérience utilisateur | Non évalué | Aspect esthétique pas évalué |

| | | | | | |
|---|---|---|---|---|---|
| Décourageant – Motivant | Hassenzahl et al. (2003) | Attractivité | Expérience utilisateur | Trop vague | |
| Fâcheuse / Réjouissante | Schrepp & Thomaschewski (2019) | Attractivité | Expérience utilisateur | Reformuler | au masculin |
| Mauvaise / Bonne | Schrepp & Thomaschewski (2019) | Attractivité | Expérience utilisateur | ok | |
| Désagréable / Agréable | Schrepp & Thomaschewski (2019) | Attractivité | Expérience utilisateur | Similaire | Item similaire chez Hassenzahl et al. (2003) |
| Pas sympathique / Sympathique | Schrepp & Thomaschewski (2019) | Attractivité | Expérience utilisateur | Ok | |
| Incompréhensible / Compréhensible | Schrepp & Thomaschewski (2019) | Transparence | Expérience utilisateur | Ok | |
| Difficile à apprendre / Facile à apprendre | Schrepp & Thomaschewski (2019) | Transparence | Expérience utilisateur | Similaire | Similaire à « Avec le chatbot c'est plus facile d'apprendre » |
| Compliquée / Facile | Schrepp & Thomaschewski (2019) | Transparence | Expérience utilisateur | Similaire | Similaire à l'item : « Je trouve le chatbot facile à utiliser » de Caron & Heutte (2017) |
| Confuse / Claire | Schrepp & Thomaschewski (2019) | Transparence | Expérience utilisateur | Similaire | Similaire à « Surchargé - Ordonné » dans UEQ+ |
| Pas intéressant / Intéressant | Schrepp & Thomaschewski (2019) | Stimulation | Expérience utilisateur | Ok | |
| Ennuyeux / Passionnant | Schrepp & Thomaschewski (2019) | Stimulation | Expérience utilisateur | ok | |
| De mauvaise qualité / Précieux | Schrepp & Thomaschewski (2019) | Valeur | Expérience utilisateur | Ok | Remplacer « précieux » par «de bonne qualité » Évalue l'efficacité |
| Soporifique / Stimulant | Schrepp & Thomaschewski (2019) | Stimulation | Expérience utilisateur | ok | |
| Dépourvue d'imagination / Créative | Schrepp & Thomaschewski (2019) | Originalité | Expérience utilisateur | Reformuler | au masculin |
| Conventionnelle / Originale | Schrepp & Thomaschewski (2019) | Originalité | Expérience utilisateur | Reformuler | au masculin |
| Conventionnelle / Inédite | Schrepp & Thomaschewski (2019) | Originalité | Expérience utilisateur | Similaire | Idem « Conventionnelle - Originale » |

| | | | | | |
|---|---|---|---|---|---|
| Conservatrice / Innovatrice | Schrepp & Thomaschewski (2019) | Originalité | Expérience utilisateur | Reformuler | au masculin |
| Peu sûr / Sûr | Schrepp & Thomaschewski (2019) | Confiance | Expérience utilisateur | Trop vague | |
| Pas sérieux / Sérieux | Schrepp & Thomaschewski (2019) | Confiance | Expérience utilisateur | ok | |
| Non fiable / Fiable | Schrepp & Thomaschewski (2019) | Confiance | Expérience utilisateur | Ok | |
| Non transparent / Transparent | Schrepp & Thomaschewski (2019) | Confiance | Expérience utilisateur | Trop vague | |
| Non présentable / Présentable | Schrepp & Thomaschewski (2019) | Valeur | Expérience utilisateur | Trop vague | |
| Pas de bon goût / De bon goût | Schrepp & Thomaschewski (2019) | Valeur | Expérience utilisateur | Trop vague | |
| Non élégante / Élégantes | Schrepp & Thomaschewski (2019) | Valeur | Expérience utilisateur | Trop vague | |
| Invraisemblables / Vraisemblable | Schrepp & Thomaschewski (2019) | Crédibilité | Expérience utilisateur | Trop vague | |
| Pas sérieuses / Sérieuses | Schrepp & Thomaschewski (2019) | Crédibilité | Expérience utilisateur | Trop vague | |
| Imprécises / Précises | Schrepp & Thomaschewski (2019) | Crédibilité | Expérience utilisateur | Ok | |
| L'utilisation du chatbot est une bonne idée. | Venkatesh et al. (2003) | Influence sociale | Acceptabilité | Trop vague | |
| Le chatbot rend l'apprentissage plus intéressant. | Venkatesh et al. (2003) | Influence sociale | Acceptabilité | Ok | |
| Travailler avec le chatbot est amusant. | Venkatesh et al. (2003) | Influence sociale | Acceptabilité | Ok | |
| J'aime utiliser le chatbot. | Venkatesh et al. (2003) | Influence sociale | Acceptabilité | Ok | |
| Des gens qui influencent mon comportement pensent que je devrais utiliser le système. | Venkatesh et al. (2003) | Influence sociale | Acceptabilité | Trop vague | |
| Les gens qui sont importants pour moi pensent que je | Venkatesh & al. (2003) | Influence sociale | Acceptabilité | Trop vague | |

| | | | | | |
|---|---|---|---|---|---|
| devrais utiliser le chatbot. | | | | | |
| Une personne spécifique (ou un groupe) est disponible pour apporter de l'aide en cas de difficultés avec le chatbot. | Venkatesh et al. (2003) | Conditions de facilitation | Acceptabilité | Trop vague | |
| En général, l'institution encourage l'utilisation du chatbot. | Venkatesh et al. (2003) | Conditions de facilitation | Acceptabilité | Non évalué | |
| J'ai les ressources nécessaires pour utiliser le chatbot. | Venkatesh et al. (2003) | Conditions de facilitation | Acceptabilité | Trop vague | |
| J'ai les connaissances nécessaires pour utiliser le chatbot. | Venkatesh et al. (2003) | Conditions de facilitation | Acceptabilité | Non évalué | |
| Le chatbot n'est pas compatible avec d'autres systèmes que j'utilise. | Venkatesh et al. (2003) | Conditions de facilitation | Acceptabilité | Non évalué | |
| Je pourrais réaliser une tâche en utilisant le chatbot...<br>S'il n'y a personne autour de moi pour me dire quoi faire. | Venkatesh et al. (2003) | Anxiété | Acceptabilité | Non évalué | |
| Si je peux appeler quelqu'un à l'aide si je suis bloqué. | Venkatesh et al. (2003) | Anxiété | Acceptabilité | Non évalué | |
| Si j'ai beaucoup de temps à disposition pour réaliser la tâche à l'aide du chatbot. | Venkatesh et al. (2003) | Anxiété | Acceptabilité | Trop vague | |
| Si je n'ai que l'assistance prédéfinie à disposition. | Venkatesh et al. (2003) | Anxiété | Acceptabilité | Trop vague | |
| J'appréhende l'utilisation du chatbot. | Venkatesh et al. (2003) | Anxiété | Acceptabilité | ok | |
| J'ai peur de penser que je peux perdre un grand nombre d'informations en utilisant le système ou en cliquant sur un mauvais bouton. | Venkatesh et al. (2003) | Anxiété | Acceptabilité | Trop vague | |

| | | | | | |
|---|---|---|---|---|---|
| J'hésite à utiliser le système par peur de faire des erreurs que je ne peux pas corriger. | Venkatesh et al. (2003) | Anxiété | Acceptabilité | Trop vague | |
| L'utilisation du chatbot m'intimide un peu. | Venkatesh et al. (2003) | Anxiété | Acceptabilité | Similaire | Similaire à « J'appréhende l'utilisation du chatbot » |
| J'utilise le chatbot parce que l'on me le demande. | Caron & Heutte (2017) | Injonction | Acceptabilité | ok | |
| J'utilise le chatbot parce que les autres apprenants l'utilisent. | Caron & Heutte (2017) | Injonction | Acceptabilité | ok | |
| J'utilise le chatbot parce que j'espère de son usage une certaine reconnaissance au sein de mon établissement. | Caron & Heutte (2017) | Injonction | Acceptabilité | Non évalué | Évaluation ne concerne pas la formation mais un soutien à l'apprentissage. |
| Je trouve l'utilisation du chatbot utile dans mon apprentissage. | Venkatesh et al. (2003) | Performances attendues | Utilisabilité | Ok | Reformuler en « utile pour » |
| Utiliser le chatbot me permet d'accomplir les tâches plus rapidement. | Venkatesh et al. (2003) | Performances attendues | Utilisabilité | Ok | |
| Utiliser le chatbot augmente ma productivité. | Venkatesh et al. (2003) | Performances attendues | Utilisabilité | Non évalué | Productivité non évaluée car accompagnant |
| Mon interaction avec le chatbot est claire et compréhensible. | Venkatesh et al. (2003) | Efforts attendus | Utilisabilité | Ok | |
| C'est facile pour moi de devenir expérimenté en utilisant le chatbot. | Venkatesh et al. (2003) | Efforts attendus | Utilisabilité | Trop vague | |
| Je trouve que le chatbot est facile à utiliser. | Venkatesh et al. (2003) | Efforts attendus | Utilisabilité | Similaire | Idem que « Je trouve que le chatbot est facile à utiliser » de Davis (1989). Item retenu de Carron & Heutte (2017) |
| Apprendre à utiliser le chatbot est facile pour moi. | Venkatesh et al. (2003) | Efforts attendus | Utilisabilité | Similaire | Similaire à « Apprendre à utiliser le chatbot sera facile pour moi » de Davis (1989) |
| J'ai l'intention d'utiliser le chatbot ces prochains mois. | Venkatesh et al. (2003) | Intention d'utiliser le système | Utilisabilité | Ok | |

| | | | | | |
|---|---|---|---|---|---|
| Je pense que je vais utiliser le chatbot ces prochains mois. | Venkatesh et al. (2003) | Intention d'utiliser le système | Utilisabilité | Similaire | Idem que « J'ai l'intention d'utiliser le chatbot ces prochains mois » de Venkatesh et al. (2003) |
| Je planifie d'utiliser le chatbot ces prochains mois. | Venkatesh et al. (2003) | Intention d'utiliser le système | Utilisabilité | Similaire | Idem que « J'ai l'intention d'utiliser le chatbot ces prochains mois » de Venkatesh et al. (2003) |
| Apprendre à utiliser le chatbot sera facile pour moi. | Davis (1989) | Facilité d'utilisation perçue | Utilisabilité | ok | |
| Je trouve qu'il est facile de demander au chatbot de faire ce que je veux. | Davis (1989) | Facilité d'utilisation perçue | Utilisabilité | Trop vague | |
| Mon interaction avec le chatbot est claire et compréhensible. | Davis (1989) | Facilité d'utilisation perçue | Utilisabilité | Identique | Même item chez Venkatesh et al. (2003) |
| Je trouve le chatbot flexible dans nos interactions. | Davis (1989) | Facilité d'utilisation perçue | Utilisabilité | Trop vague | |
| Il me sera facile de devenir habile dans l'utilisation du chatbot. | Davis (1989) | Facilité d'utilisation perçue | Utilisabilité | Trop vague | |
| Je trouve que le chatbot est facile à utiliser. | Davis (1989) | Facilité d'utilisation perçue | Utilisabilité | Similaire | Idem que « Je trouve que le chatbot est facile à utiliser » de Venkatesh et al. (2003).<br>Item retenu de Carron & Heutte 2017) |
| Je voudrais utiliser ce système fréquemment. | Brooke (1996) | Utilisabilité | Utilisabilité | ok | |
| Ce système est inutilement complexe. | Brooke (1996) | Utilisabilité | Utilisabilité | Similaire | Similaire à « Je trouve le chatbot facile à utiliser » de Caron & Heutte (2017) |
| Ce système est facile à utiliser. | Brooke (1996) | Utilisabilité | Utilisabilité | Ok | Remplacer « système » par « chatbot » |
| J'aurais besoin du soutien d'un technicien pour être capable d'utiliser ce système. | Brooke (1996) | Utilisabilité | Utilisabilité | Similaire | Similaire à « Je trouve le chatbot facile à utiliser » de Caron & Heutte (2017) |

| | | | | | |
|---|---|---|---|---|---|
| Les différentes fonctionnalités de ce système sont bien intégrées. | Brooke (1996) | Utilisabilité | Utilisabilité | ok | |
| Il y a trop d'incohérences dans ce système. | Brooke (1996) | Utilisabilité | Utilisabilité | Reformuler | Remplacer « système » par « chatbot » |
| La plupart des gens apprendront à utiliser ce système très rapidement. | Brooke (1996) | Utilisabilité | Utilisabilité | Trop vague | |
| Ce système est très lourd à utiliser. | Brooke (1996) | Utilisabilité | Utilisabilité | Similaire | Similaire à « Je trouve le chatbot facile à utiliser » de Caron & Heutte (2017) |
| Je me suis senti·e très en confiance en utilisant ce système. | Brooke (1996) | Utilisabilité | Utilisabilité | Reformuler | Remplacer « système » par « chatbot » |
| J'ai eu besoin d'apprendre beaucoup de choses avant de pouvoir utiliser ce système. | Brooke (1996) | Utilisabilité | Utilisabilité | Trop vague | |
| C'est facile d'apprendre à utiliser le chatbot. | Caron & Heutte (2017) | Utilisabilité | Utilisabilité | Similaire | Similaire à « Je trouve le chatbot facile à utiliser » de Caron & Heutte (2017) |
| C'est facile d'utiliser le chatbot pour faire ce que je veux faire. | Caron & Heutte (2017) | Utilisabilité | Utilisabilité | Similaire | Similaire à « Je trouve le chatbot facile à utiliser » de Caron & Heutte (2017) |
| Je trouve le chatbot facile à utiliser. | Caron & Heutte (2017) | Utilisabilité | Utilisabilité | Ok | |
| Le chatbot est facile à utiliser. | Lund (2001) | Utilisabilité | Utilisabilité | Similaire | Similaire à « Je trouve le chatbot facile à utiliser » de Caron & Heutte (2017) |
| Le chatbot est simple à utiliser. | Lund (2001) | Utilisabilité | Utilisabilité | Similaire | Similaire à « Je trouve le chatbot facile à utiliser » de Caron & Heutte (2017) |
| Le chatbot est convivial. | Lund (2001) | Utilisabilité | Utilisabilité | Similaire | Similaire à « Je trouve le chatbot facile à utiliser » de Caron & Heutte (2017) |
| Le chatbot nécessite le moins d'étapes possibles pour | Lund (2001) | Utilisabilité | Utilisabilité | Reformuler | Reformuler en « Le chatbot permet de réaliser des |

| | | | | | |
|---|---|---|---|---|---|
| accomplir ce que je veux faire avec lui. | | | | | tâches en peu d'étapes ». |
| Le chatbot est flexible. | Lund (2001) | Utilisabilité | Utilisabilité | Trop vague | |
| Utiliser le chatbot ne nécessite pas d'effort. | Lund (2001) | Utilisabilité | Utilisabilité | Trop vague | |
| Je peux utiliser le chatbot sans instruction écrite. | Lund (2001) | Utilisabilité | Utilisabilité | Trop vague | |
| Je ne rencontre pas d'incohérence quand j'utilise le chatbot. | Lund (2001) | Utilisabilité | Utilisabilité | Similaire | Similaire à « Il y a trop d'incohérences dans ce système » de Brooke (1996) |
| Le chatbot convient à la fois aux utilisateurs réguliers et aux utilisateurs occasionnels. | Lund (2001) | Utilisabilité | Utilisabilité | Trop vague | |
| Je peux corriger mes erreurs rapidement et facilement. | Lund (2001) | Utilisabilité | Utilisabilité | Trop vague | |
| Je peux utiliser le chatbot avec succès à chaque fois. | Lund (2001) | Utilisabilité | Utilisabilité | Trop vague | |
| J'ai appris à utiliser le chatbot facilement. | Lund (2001) | Facilité d'apprentissage | Utilisabilité | Similaire | Similaire à « Je trouve le chatbot facile à utiliser » de Caron & Heutte (2017) |
| Je me rappelle facilement comment utiliser le chatbot. | Lund (2001) | Facilité d'apprentissage | Utilisabilité | Similaire | Similaire à « Je trouve le chatbot facile à utiliser » de Caron & Heutte (2017) |
| Il est facile d'apprendre à utiliser le chatbot. | Lund (2001) | Facilité d'apprentissage | Utilisabilité | Similaire | Similaire à « Ce chatbot est facile à utiliser ». « J'aurai besoin du soutien d'un technicien pour être capable d'utiliser ce chatbot ». « La plupart des gens apprendront à utiliser ce chatbot très rapidement ». « Ce chatbot est lourd à utiliser ». |
| Je suis rapidement devenu habile avec le chatbot. | Lund (2001) | Facilité d'apprentissage | Utilisabilité | Trop vague | |
| Je suis satisfait du chatbot. | Lund (2001) | Satisfaction | Utilisabilité | Ok | |

| | | | | | |
|---|---|---|---|---|---|
| Je recommanderais le chatbot à un ami. | Lund (2001) | Satisfaction | Utilisabilité | Reformuler | Reformuler en « Je recommanderais l'utilisation du chatbot à un ami » |
| Le chatbot est amusant à utiliser. | Lund (2001) | Satisfaction | Utilisabilité | Similaire | Similaire à « Travailler avec le chatbot est amusant » de Venkatesh et al. (2003) |
| Le chatbot fonctionne comme je veux qu'il fonctionne. | Lund (2001) | Satisfaction | Utilisabilité | Similaire | Similaire à « Les différentes fonctionnalités de ce chatbot sont bien intégrées » de Brooke (1996) |
| Le chatbot est merveilleux. | Lund (2001) | Satisfaction | Utilisabilité | Trop vague | |
| J'ai l'impression que j'ai besoin d'avoir un chatbot. | Lund (2001) | Satisfaction | Utilisabilité | Similaire | Similaire à « Le chatbot répond à mes besoins » |
| Le chatbot est agréable à utiliser. | Lund (2001) | Satisfaction | Utilisabilité | Ok | Reformuler « agréable » en « plaisant » |
| Laborieuse / Aisée | Schrepp & Thomaschewski (2019) | Utilisation intuitive | Utilisabilité | ok | |
| Non logique / Logique | Schrepp & Thomaschewski (2019) | Utilisation intuitive | Utilisabilité | Ok | |
| Pas évidente / Évidente | Schrepp & Thomaschewski (2019) | Utilisation intuitive | Utilisabilité | Trop vague | |
| Incohérente / Cohérente | Schrepp & Thomaschewski (2019) | Utilisation intuitive | Utilisabilité | Similaire | Similaire à « Il y a trop d'incohérences dans ce chatbot » |
| Imprévisible / Prévisible | Schrepp & Thomaschewski (2019) | Maniabilité | Utilisabilité | Ok | Identique à Hassenzahl et al. (2003) |
| Gênante / Utile | Schrepp & Thomaschewski (2019) | Maniabilité | Utilisabilité | Trop vague | |
| Peu sûre / Sûr | Schrepp & Thomaschewski (2019) | Maniabilité | Utilisabilité | Trop vague | |
| Pas conforme aux attentes / Conforme aux attentes | Schrepp & Thomaschewski (2019) | Maniabilité | Utilisabilité | Trop vague | |
| Non bénéfique / Bénéfique | Schrepp & Thomaschewski (2019) | Utilité et Crédibilité | Utilité | Ok | |
| Pas utile / Utile | Schrepp & Thomaschewski (2019) | Utilité | Utilité | Similaire | Similaire à « Le chatbot est utile » de Lund (2001) |

| | | | | | |
|---|---|---|---|---|---|
| Non avantageuse / Avantageuse | Schrepp & Thomaschewski (2019) | Utilité | Utilité | Trop vague | |
| Non rentable / Rentable | Schrepp & Thomaschewski (2019) | Utilité | Utilité | Non évalué | |
| Utiliser le chatbot me permettra d'accomplir les tâches plus rapidement. | Davis (1989) | Utilité perçue | Utilité | Similaire | Idem que « Utiliser le chatbot me permet d'accomplir les tâches plus rapidement » de Venkatesh et al. (2003) |
| L'utilisation du chatbot augmentera mes performances. | Davis (1989) | Utilité perçue | Utilité | Non évalué | Performances non évaluées car accompagnant |
| L'utilisation du chatbot augmentera ma productivité. | Davis (1989) | Utilité perçue | Utilité | Non évalué | Productivité non évaluée car accompagnant |
| L'utilisation du chatbot améliorera mon efficacité. | Davis (1989) | Utilité perçue | Utilité | Similaire | Similaire à « L'utilisation du chatbot me rend plus efficace » de Caron & Heutte (2017) |
| Utiliser le chatbot faciliterait mon travail. | Davis (1989) | Utilité perçue | Utilité | Reformuler | Reformuler en « Le chatbot me permet d'accomplir les tâches plus facilement » |
| Je trouve que le chatbot est un outil utile dans mon travail. | Davis (1989) | Utilité perçue | Utilité | Similaire | Similaire à « Je trouve le chatbot utile pour mon apprentissage » de Venkatesh et al. (2003) |
| L'utilisation du chatbot dans mes études me permet d'accomplir des tâches plus rapidement. | Caron & Heutte (2017) | Utilité générale | Utilité | Similaire | Idem que « Utiliser le chatbot me permet d'accomplir les tâches plus rapidement » de Venkatesh et al. (2003) |
| L'utilisation du chatbot me rend plus efficace. | Caron & Heutte (2017) | Utilité générale | Utilité | Ok | |
| Je trouve le chatbot utile pour mon projet d'étude. | Caron & Heutte (2017) | Utilité générale | Utilité | Similaire | Similaire à « Je trouve le chatbot utile » de Lund (2001) |
| Le chatbot me permet d'apprendre plus rapidement. | Caron & Heutte (2017) | Utilité apprenant | Utilité | Similaire | Similaire à « Utiliser le chatbot me permet d'accomplir les tâches plus rapidement » de |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | Venkatesh et al. (2003) |
| Avec le chatbot j'ai plus de facilité pour accéder à la formation reçue. | Caron & Heutte (2017) | Utilité apprenant | Utilité | Non évalué | Évaluation ne concerne pas l'accès à la formation. |
| Le chatbot améliore ma capacité à appliquer la formation reçue. | Caron & Heutte (2017) | Utilité apprenant | Utilité | Non évalué | Évaluation ne concerne pas la formation mais un soutien à l'apprentissage. |
| Avec le chatbot c'est plus facile d'apprendre. | Caron & Heutte (2017) | Utilité apprenant | Utilité | Ok | |
| Le chatbot est utile pour mon apprentissage. | Caron & Heutte (2017) | Utilité apprenant | Utilité | Similaire | Similaire à « Je trouve l'utilisation du chatbot utile dans mon apprentissage » de Venkatesh et al. (2003) |
| Avec le chatbot, je comprends mieux mon projet d'étude. | Caron & Heutte (2017) | Utilité pour moi-même | Utilité | Non évalué | Évaluation ne concerne pas un projet d'étude, seulement le cours ADID. |
| Le chatbot m'a permis de mieux mettre en œuvre mon projet d'étude. | Caron & Heutte (2017) | Utilité pour moi-même | Utilité | Non évalué | Évaluation ne concerne pas un projet d'étude, seulement le cours ADID. |
| Le chatbot m'aide à être plus efficace. | Lund (2001) | Utilité | Utilité | Similaire | Idem que « L'utilisation du chatbot me rend plus efficace » de Caron & Heutte (2017) |
| Le chatbot m'aide à être plus productif | Lund (2001) | Utilité | Utilité | Non évalué | Productivité non évaluée car accompagnant |
| Le chatbot est utile. | Lund (2001) | Utilité | Utilité | Ok | |
| Le chatbot me permet d'avoir plus de contrôle sur les activités de ma vie quotidienne. | Lund (2001) | Utilité | Utilité | Trop vague | |
| Le chatbot me permet de faire plus facilement ce que je veux faire. | Lund (2001) | Utilité | Utilité | Similaire | Idem que « Le chatbot me permet d'accomplir les tâches plus facilement » Davis (1989) |
| Le chatbot me fait gagner du temps quand je l'utilise. | Lund (2001) | Utilité | Utilité | Ok | Reformuler en « Le chatbot me fait gagner du temps » |

| | | | | | |
|---|---|---|---|---|---|
| Le chatbot répond à mes besoins. | Lund (2001) | Utilité | Utilité | Ok | Reformuler en « Le chatbot correspond à ce que j'ai besoin » |
| Le chatbot fait tout ce que je veux qu'il fasse. | Lund (2001) | Utilité | Utilité | Reformuler | Reformuler en « Le chatbot fait tout ce que j'attends de lui » |
| Qu'est-ce qui vous motive à utiliser un chatbot ? | Brandtzaeg et al. (2017) | Motivation | Motivation | Ok | |
| Quelle est votre principale raison d'utiliser un chatbot ? | Brandtzaeg et al. (2017) | Motivation | Motivation | Ok | |
| En quoi ce chatbot est le plus performant ? | Kirakowski, O'Donnell & Yiu (2009) | Humanité | Comportemen t | Ok | |
| Malheureux - Heureux | Bradley & Lang (1994) | Pleasure | Émotions | ok | |
| Insatisfait - Satisfait | Bradley & Lang (1994) | Pleasure | Émotions | Ok | |
| Lenteur - Rapidité | Bradley & Lang (1994) | Arousal | Émotions | Autre critère évalué | Évalue le critère Efficacité Idem « Lent - Rapide » de Schrepp & Thomaschewski (2019) |
| Guidé - Autonome | Bradley & Lang (1994) | Dominance | Émotions | ok | |
| Satisfait | Kay & Loverock (2008) | Happiness | Émotions | Similaire | Idem à « Insatisfait - Satisfait » de Bradley & Lang (1994) |
| Découragé | Kay & Loverock (2008) | Sadness | Émotions | Reformuler | Trop négatif et/ou à transformer en paire de mots |
| Anxieux | Kay & Loverock (2008) | Anxiety | Émotions | Reformuler | Trop négatif et/ou à transformer en paire de mots |
| Irritable | Kay & Loverock (2008) | Anger | Émotions | Reformuler | Trop négatif et/ou à transformer en paire de mots |
| Excité | Kay & Loverock (2008) | Happiness | Émotions | Reformuler | Transformer en paire de mots |
| Dépité | Kay & Loverock (2008) | Sadness | Émotions | Reformuler | Trop négatif et/ou à transformer en paire de mots |
| Peu sûr de moi | Kay & Loverock (2008) | Anxiety | Émotions | Reformuler | Trop négatif et/ou à transformer en paire de mots |
| Frustré | Kay & Loverock (2008) | Anger | Émotions | Reformuler | Trop négatif et/ou à transformer en paire de mots |

| | | | | | |
|---|---|---|---|---|---|
| Curieux | Kay & Loverock (2008) | Happiness | Émotions | Reformuler | Transformer en paire de mots |
| Impuissant | Kay & Loverock (2008) | Anxiety | Émotions | Reformuler | Trop négatif et/ou à transformer en paire de mots |
| Nerveux | Kay & Loverock (2008) | Anxiety | Émotions | Reformuler | Trop négatif et/ou à transformer en paire de mots |
| En colère | Kay & Loverock (2008) | Anger | Émotions | Reformuler | Trop négatif et/ou à transformer en paire de mots |
| Lent / Rapide | Schrepp & Thomaschewski (2019) | Efficacité | Efficacité et performances | Ok | |
| Inefficace / Efficace | Schrepp & Thomaschewski (2019) | Efficacité | Efficacité et performances | Ok | |
| Pas pratique / Pratique | Schrepp & Thomaschewski (2019) | Efficacité | Efficacité et performances | Ok | |
| Surchargé / Ordonné | Schrepp & Thomaschewski (2019) | Efficacité | Efficacité et performances | ok | |
| Cette expérience m'a tellement absorbé que j'en ai perdu la notion du temps. | O'Brien, Cairns & Hall (2018) | Attention soutenue | Engagement | Trop vague | |
| Le temps a filé lorsque j'ai utilisé le chatbot. | O'Brien, Cairns & Hall (2018) | Attention soutenue | Engagement | Trop vague | |
| J'ai été absorbé par cette expérience. | O'Brien, Cairns & Hall (2018) | Attention soutenue | Engagement | Trop vague | |
| J'ai éprouvé de la frustration lors de l'utilisation du chatbot. | O'Brien, Cairns & Hall (2018) | Convivialité perçue | Engagement | Similaire | Similaire à Émotions |
| J'ai trouvé que le chatbot était difficile à utiliser. | O'Brien, Cairns & Hall (2018) | Convivialité perçue | Engagement | Similaire | Similaire à Utilisabilité |
| L'utilisation du chatbot était pénible. | O'Brien, Cairns & Hall (2018) | Convivialité perçue | Engagement | Similaire | Similaire à Utilisabilité |
| L'utilisation du chatbot en valait la peine. | O'Brien, Cairns & Hall (2018) | Gratification | Engagement | Similaire | Similaire à Utilité |
| Mon expérience s'est avérée gratifiante | O'Brien, Cairns & Hall (2018) | Gratification | Engagement | Trop vague | |
| J'ai été intéressé par cette expérience. | O'Brien, Cairns & Hall (2018) | Gratification | Engagement | Similaire | Similaire à Utilité |
| Dans quelle mesure vos interactions avec l'environnement vous | Witmer & Singer (1998) | Réalisme | Présence | Ok | Former une paire de mot : Naturelles - Robotisées |

| | | | | | |
|---|---|---|---|---|---|
| semblaient-elles naturelles ? | | | | | |
| Dans quelle mesure les expériences que vous avez vécues dans l'environnement virtuel ressemblaient-elles à celles de l'environnement réel ? | Witmer & Singer (1998) | Réalisme | Présence | Reformuler | Transformer en paire de mots |

## Annexe 15 - Résultat du tri de cartes des items sous formes de phrases

| Critères | Items | Participants | | | | | | | | Total | Utile | Non utile | Incompréh |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Alpha | Beta | Delta | Epsilon | Zeta | Êta | Thêta | Iota | | | | |
| Acceptabilité | Utiliser le chatbot me permet d'accomplir les tâches plus rapidement. | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |
| Acceptabilité | Mon interaction avec le chatbot est claire et compréhensible. | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |
| Acceptabilité | Le chatbot rend l'apprentissage plus intéressant. | Utile | Inco | Utile | Non | Non | Utile | Utile | Utile | **5/8** | 5 | 2 | 1 |
| Acceptabilité | Travailler avec le chatbot est amusant. | Non | Utile | Non | Utile | Utile | Utile | Utile | Utile | **6/8** | 6 | 2 | 0 |
| Acceptabilité | J'aime utiliser le chatbot. | Utile | Non | Utile | Utile | Utile | Non | Utile | Utile | **6/8** | 6 | 2 | 0 |
| Acceptabilité | J'utilise le chatbot parce qu'on me le demande. | Inco | Non | Non | Non | Inco | Non | Non | Non | **0/8** | 0 | 6 | 2 |
| Acceptabilité | J'utilise le chatbot parce que les autres apprenants l'utilisent. | Non | Non | Non | Non | Non | Utile | Non | Non | **1/8** | 1 | 7 | 0 |
| Acceptabilité | J'appréhende l'utilisation du chatbot. | Inco | Utile | Non | Utile | Non | Non | Utile | Utile | **4/8** | 4 | 3 | 1 |
| Acceptabilité | J'ai l'intention d'utiliser le chatbot ces prochains mois. | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Inco | **7/8** | 7 | 0 | 1 |
| Acceptabilité | L'utilisation du chatbot me rend plus efficace. | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |
| Acceptabilité | Avec le chatbot, c'est plus facile d'apprendre. | Utile | Utile | Utile | Non | Utile | Utile | Utile | Utile | **7/8** | 7 | 1 | 0 |
| Acceptabilité | Je trouve le chatbot utile pour mon apprentissage. | Utile | Utile | Utile | Non | Utile | Utile | Utile | Inco | **6/8** | 6 | 1 | 1 |
| Acceptabilité | Je trouve le chatbot facile à utiliser. | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |
| Utilité | Le chatbot est utile. | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Inco | **7/8** | 7 | 0 | 1 |
| Utilité | Le chatbot me permet d'accomplir les tâches plus facilement. | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Utilité | Le chatbot me fait gagner du temps. | Utile | Non | Utile | Utile | Utile | Utile | Utile | Utile | **7/8** | 7 | 1 | 0 |
| Utilité | Le chatbot correspond à ce que j'ai besoin. | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |
| Utilité | Le chatbot fait tout ce que j'attends de lui. | Utile | Utile | Non | Non | Non | Utile | Utile | Non | **4/8** | 4 | 4 | 0 |
| Utilité | Le chatbot permet de réaliser des tâches en peu d'étapes. | Utile | Utile | Utile | Utile | Utile | Non | Utile | Non | **6/8** | 6 | 2 | 0 |
| Utilité | Je suis satisfait.e du chatbot. | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |
| Utilité | Je recommanderais l'utilisation du chatbot à un ami. | Utile | Non | Utile | Utile | Utile | Non | Utile | Non | **5/8** | 5 | 3 | 0 |
| Utilisabilité | Je voudrais utiliser ce chatbot fréquemment. | Utile | Non | Utile | Utile | Non | Non | Utile | Non | **4/8** | 4 | 4 | 0 |
| Utilisabilité | Ce chatbot est facile à utiliser. | Non | Inco | Utile | Utile | Utile | Utile | Utile | Utile | **6/8** | 6 | 1 | 1 |
| Utilisabilité | Les différentes fonctionnalités de ce chatbot sont bien intégrées. | Inco | Utile | Utile | Inco | Utile | Utile | Utile | Utile | **6/8** | 6 | 0 | 2 |
| Utilisabilité | Il y a trop d'incohérences dans ce chatbot. | Inco | Utile | Utile | Utile | Inco | Non | Utile | Utile | **5/8** | 5 | 1 | 2 |
| Utilisabilité | Je me suis senti.e très en confiance en utilisant ce chatbot. | Utile | Utile | Non | Utile | Non | Utile | Utile | Non | **5/8** | 5 | 3 | 0 |
| Utilisabilité | Le chatbot est plaisant à utiliser. | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |

## Annexe 16 - Résultats du tri de cartes des items sous forme de paires de mots

| Critères | Paire de mots | Participants | | | | | | | | Total | Utile | Non utile | Incompréhe |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Alpha | Beta | Delta | Epsilon | Zeta | Êta | Thêta | Iota | | | | |
| Qualité de l'environnement en ligne | Rassurant - Déroutant | Utile | Utile | Utile | Utile | Non | Utile | Utile | Utile | **7/8** | 7 | 1 | 0 |
| Qualité de l'environnement en ligne | Ludique - Académique | Utile | Non | Non | Utile | Utile | Utile | Utile | Utile | **6/8** | 6 | 2 | 0 |
| Qualité de l'environnement en ligne | Actif - Passif | Utile | Non | Utile | Non | Non | Non | Non | Utile | **3/8** | 3 | 5 | 0 |
| Qualité de l'environnement en ligne | Innovant - Traditionnel | Utile | Non | Non | Inco | Utile | Non | Utile | Utile | **4/8** | 4 | 3 | 1 |
| Expérience utilisateur (Attractivité et transparence) | Fâcheux - Réjouissant | Inco | Utile | Inco | Inco | Non | Inco | Utile | Inco | **2/8** | 2 | 1 | 5 |
| Expérience utilisateur (Attractivité et transparence) | Désagréable - Agréable | Utile | Utile | Utile | Non | Utile | Utile | Utile | Non | **6/8** | 6 | 2 | 0 |
| Expérience utilisateur (Attractivité et transparence) | Pas sympathique - Sympathique | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |
| Expérience utilisateur (Attractivité et transparence) + Crédibilité | Incompréhensible - Compréhensible | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Inco | **7/8** | 7 | 0 | 1 |
| Efficience et simulation | Lent - Rapide | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |
| Efficience et simulation | Inefficace - Efficace | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |
| Efficience et simulation | Pas pratique - Pratique | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |
| Efficience et simulation | Surchargé - Ordonné | Utile | Utile | Utile | Utile | Utile | Utile | Non | Inco | **6/8** | 6 | 1 | 1 |
| Efficience et simulation | Pas intéressant - Intéressant | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Non | **7/8** | 7 | 1 | 0 |
| Efficience et simulation | Ennuyeux - Passionnant | Utile | Utile | Utile | Non | Non | Utile | Utile | Non | **5/8** | 5 | 3 | 0 |
| Efficience et simulation + Originalité et valeur | De mauvaise qualité - De bonne qualité | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Non | **7/8** | 7 | 1 | 0 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Efficience et simulation | Soporifique - Stimulant | Utile | Utile | Utile | Non | Non | Utile | Utile | Non | **5/8** | 5 | 3 | 0 |
| Utilisation intuitive | Laborieux - Aisé | Inco | Inco | Utile | Utile | Utile | Utile | Utile | Utile | **6/8** | 6 | 0 | 2 |
| Efficience et simulation | Non logique - Logique | Utile | Utile | Utile | Utile | Utile | Non | Utile | Utile | **7/8** | 7 | 1 | 0 |
| Originalité et valeur | Dépourvu d'imagination - Créatif | Non | Non | Non | Inco | Non | Non | Utile | Non | **1/8** | 1 | 6 | 1 |
| Originalité et valeur | Conventionnel - Original | Utile | Non | Non | Inco | Utile | Non | Utile | Utile | **4/8** | 4 | 3 | 1 |
| Originalité et valeur | Conservateur - Innovateur | Utile | Non | Non | Inco | Utile | Non | Utile | Utile | **4/8** | 4 | 3 | 1 |
| Confiance | Non fiable - Fiable | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |
| Crédibilité | Non bénéfique - Bénéfique | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Non | **7/8** | 7 | 1 | 0 |
| Crédibilité | Pas sérieux - Sérieux | Utile | Inco | Utile | Utile | Utile | Non | Utile | Non | **5/8** | 5 | 2 | 1 |
| Crédibilité | Imprécis - Précis | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |
| Crédibilité | Mal préparé - Bien préparé | Inco | Utile | Utile | Inco | Utile | Utile | Non | Inco | **4/8** | 4 | 1 | 3 |
| Présence | Naturel - Robotisé | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |
| Présence | Environnement réel - Environnement fictif | Inco | Inco | Utile | Utile | Inco | Inco | Non | Inco | **2/8** | 2 | 1 | 5 |
| Émotions | Satisfait.e - Insatisfait.e | Utile | Utile | Utile | Utile | Utile | Utile | Utile | Utile | **8/8** | 8 | 0 | 0 |
| Émotions | Déçu.e - Stimulé.e | Inco | Utile | Inco | Utile | Non | Utile | Utile | Utile | **5/8** | 5 | 1 | 2 |
| Émotions | Stressé.e - Serein.e | Non | Utile | Non | Utile | Non | Non | Utile | Utile | **4/8** | 4 | 4 | 0 |
| Émotions | Impatient.e - Calme | Non | Utile | Non | Inco | Non | Non | Non | Inco | **1/8** | 1 | 5 | 2 |

## Annexe 17 - Réponses au questionnaire

| Item | Critère<br>*Sous-critère* | Nb item du critère | Réponse 1 | Réponse 2 | Réponse 3 | Réponse 4 | Réponse 5 | Réponse 6 | Score global | Nb item |
|---|---|---|---|---|---|---|---|---|---|---|
| Qu'est-ce qui vous motive à utiliser un chatbot ? | Motivation | 2 | Obligation | Vous aider à récolter des données :). Sinon, de manière générale pour les chatbots, c'est pour avoir une réponse. | il est sympa | Curiosité, appui pédagogique | Le fait de recevoir une réponse d'une entité | Le travail de mémoire d'Aurélie et Simon | | |
| Quelle est votre principale raison d'utiliser un chatbot ? | Motivation | 2 | Pour tester | Vous aider à récolter des données :) sinon de manière générale parce que je ne peux pas obtenir de réponse autrement | participer à son évaluation | Lorsque j'ai besoin d'un renseignement | L'étude dans laquelle il est utilisé | Aider Aurélie et Simon | | |
| Le chatbot est utile. | Utilité | 6 | 3 | 4 | 7 | 3 | 3 | 1 | 21 | 6 |
| Ce chatbot est facile à utiliser. | Utilisabilité | 10 | 7 | 7 | 7 | 5 | 6 | 7 | 39 | 10 |
| J'aime utiliser le chatbot. | Acceptabilité<br>*Influence sociale* | 3 | 3 | 7 | 7 | 4 | 6 | 4 | 31 | 3 |
| Pas pratique – Pratique | Efficacité | 4 | 3 | 7 | 7 | 5 | 3 | 4 | 29 | 4 |
| De mauvaise qualité – De bonne qualité | Efficacité<br>*Valeur* | 4 | 6 | 5 | 7 | 5 | 6 | 5 | 34 | 4 |
| Lent – Rapide | Efficacité | 4 | 6 | 7 | 7 | 5 | 6 | 3 | 34 | 4 |
| Inefficace – Efficace | Efficacité | 4 | 3 | 6 | 7 | 4 | 6 | 4 | 30 | 4 |
| Non fiable – Fiable | Expérience utilisateur<br>*Confiance* | 6 | 7 | 7 | 6 | 5 | 6 | 5 | 36 | 6 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Naturel – Robotisé | Présence *Réalisme* | 1 | 6 | 6 | 7 | 3 | 7 | 7 | 36 | 1 |
| Non bénéfique – Bénéfique | Utilité *Crédibilité* | 6 | 2 | 6 | 6 | 2 | 3 | 2 | 21 | 6 |
| Imprécis – Précis | Expérience utilisateur *Crédibilité* | 6 | 5 | 6 | 6 | 2 | 5 | 4 | 28 | 6 |
| Non logique – Logique | Utilisabilité *Utilisation intuitive* | 10 | 4 | 7 | 7 | 5 | 7 | 6 | 36 | 10 |
| Incompréhensible – Compréhensible | Expérience utilisateur *Transparence* | 6 | 6 | 7 | 7 | 6 | 7 | 6 | 39 | 6 |
| Pas intéressant – Intéressant | Expérience utilisateur *Stimulation* | 6 | 2 | 6 | 6 | 2 | 5 | 4 | 25 | 6 |
| En quoi ce chatbot est le plus performant ? | Comportement *Humanité* | 1 | Humanité des réponses | Il est rapide et logique dans ses réponses. | Sa proximité dans la relation est vraiment sympa. | Ludique | Offrir une expérience conviviale | L'humour | | 1 |
| Utiliser le chatbot me permet d'accomplir les tâches plus rapidement. | Utilisabilité *Performances attendues* | 10 | 4 | 4 | 6 | 2 | 3 | 4 | 23 | 10 |
| Le chatbot me permet d'accomplir les tâches plus facilement. | Utilité | 6 | 4 | 4 | 6 | 2 | 6 | 4 | 26 | 6 |
| Le chatbot me fait gagner du temps. | Utilisabilité | 10 | 4 | 4 | 6 | 3 | 3 | 4 | 24 | 10 |
| L'utilisation du chatbot me rend plus efficace. | Utilité | 6 | 4 | 4 | 6 | 2 | 6 | 4 | 26 | 6 |
| Je trouve l'utilisation du chatbot utile pour mon apprentissage. | Utilisabilité *Performances attendues* | 10 | 2 | 4 | 6 | 3 | 5 | 4 | 24 | 10 |
| Avec le chatbot c'est plus facile d'apprendre. | Utilité | 6 | 2 | 4 | 6 | 3 | 4 | 4 | 23 | 6 |
| Ludique – Académique | Expérience utilisateur *Impressions générales* | 1 | 1 | 1 | 1 | 2 | 3 | 1 | 9 | 1 |
| Insatisfait – Satisfait | Émotions | 1 | 2 | 6 | 7 | 3 | 2 | 1 | 21 | 1 |
| Désagréable – Agréable | Expérience utilisateur *Attractivité* | 6 | 7 | 7 | 7 | 6 | 7 | 7 | 41 | 6 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Pas sympathique – Sympathique | Expérience utilisateur *Attractivité* | 6 | 7 | 7 | 7 | 6 | 7 | 7 | 41 | 6 |
| Travailler avec le chatbot est amusant. | Acceptabilité *Influence sociale* | 3 | 4 | 5 | 7 | 6 | 6 | 4 | 32 | 3 |
| Le chatbot est plaisant à utiliser. | Utilisabilité *Satisfaction* | 10 | 6 | 7 | 7 | 6 | 7 | 4 | 37 | 10 |
| Je trouve le chatbot facile à utiliser. | Utilisabilité | 10 | 7 | 7 | 7 | 6 | 7 | 7 | 41 | 10 |
| Je suis satisfait du chatbot. | Utilisabilité *Satisfaction* | 10 | 3 | 5 | 7 | 3 | 3 | 3 | 24 | 10 |
| Le chatbot correspond à ce que j'ai besoin. | Utilité | 6 | 2 | 2 | 7 | 2 | 3 | 3 | 19 | 6 |
| Mon interaction avec le chatbot est claire et compréhensible. | Utilisabilité *Efforts attendus* | 10 | 6 | 7 | 7 | 4 | 7 | 6 | 37 | 10 |
| Le chatbot rend l'apprentissage plus intéressant. | Acceptabilité *Influence sociale* | 3 | 2 | 2 | 6 | 3 | 5 | 4 | 22 | 3 |
| J'ai l'intention d'utiliser le chatbot ces prochains mois. | Utilisabilité *Intention d'utiliser le système* | 10 | 1 | 2 | 6 | 5 | 2 | 1 | 17 | 10 |

# Annexe 18 - Formulaire de consentement

Aurélie Arnoldi
Simon Jöhr

Expérience dans le cadre du Mémoire
du Master MALTT 2021-2022

UNIVERSITÉ DE GENÈVE
FACULTÉ DE PSYCHOLOGIE ET DES SCIENCES DE L'ÉDUCATION
Technologies de formation et d'apprentissage (TECFA)

## Formulaire de consentement

Bonjour,

Merci d'avoir accepté de participer à cette expérience qui vise à :

- **Tester un chatbot.**
- **Évaluer la clarté et la pertinence d'un questionnaire d'évaluation d'un chatbot.**

Dans l'expérience à laquelle vous participez, vous allez utiliser des outils étant en phase de développement. Sur la base des tests effectués et de vos retours, différentes versions pourront être produites. À travers cette expérience, nous estimons pouvoir acquérir des informations importantes afin de pouvoir améliorer le chatbot ainsi que son outil d'évaluation.

### Déroulement de l'expérience

L'expérience prévoit plusieurs phases successives pour une durée totale d'environ 30-60 minutes.

Lors de la première phase, vous allez tester le chatbot et interagir avec lui au travers d'une discussion écrite.

Dans la deuxième phase, vous allez évaluer la pertinence et la clarté d'un questionnaire qui permettra ensuite d'améliorer le chatbot précédemment utilisé.

À la fin de ces deux parties, dans la troisième phase de l'étude, vous serez interrogé.e.s durant un court entretien ayant pour but, d'une part, de récolter vos différents critères d'utilisation, besoins et fonctionnalités concernant le chatbot et d'autre part, de vérifier la formulation et la pertinence des questions du questionnaire précédemment transmis.

Nous attirons votre attention sur le fait que, durant l'expérience, votre identité est pseudonymisée, ce qui veut dire qu'un nom d'emprunt vous est attribué, et que votre réelle identité ne sera pas divulguée ou enregistrée. Néanmoins, afin de pouvoir analyser les données récoltées, l'expérience sera enregistrée (enregistrement visuel et sonore).

Aurélie Arnoldi
Simon Jöhr

Expérience dans le cadre du Mémoire
du Master MALTT 2021-2022

UNIVERSITÉ DE GENÈVE
FACULTÉ DE PSYCHOLOGIE ET DES SCIENCES DE L'ÉDUCATION
Technologies de formation et d'apprentissage (TECFA)

## Respect de la sphère privée, conservation des données

Votre anonymat est garanti à travers les plateformes *Botsify*, *Qualtrics* et *Provenbyusers* où les données seront recueillies.

Pendant l'expérience, nous allons recueillir ces informations :

- Votre tranche d'âge
- Votre sexe
- Vos interactions avec le chatbot
- Vos réponses aux questionnaires avant et après le test du chatbot.

Même lors de l'entretien, aucune donnée personnelle ne sera enregistrée et vous y participerez ainsi de manière anonyme, en désactivant votre caméra. Par ailleurs, les enregistrements réalisés (expérience et entretien) tout comme la retranscription de votre entretien seront étiquetés uniquement avec un pseudonyme que l'on va vous attribuer. Les enregistrements seront stockés dans une clé USB protégée par un mot de passe, gardée dans un tiroir fermé à clé. Les enregistrements de l'expérience et de l'entretien seront immédiatement détruits, respectivement, après exploitation et retranscription des échanges sur un fichier numérique, avant le 31 janvier 2022. Ledit fichier, tout comme l'ensemble des données collectées, sera conservé et archivé sous la responsabilité de Madame Gaëlle Molinari, Professeure à l'Université de Genève et directrice du Master MALTT. De cette manière, nous ne posséderons plus vos enregistrements et ne serons plus en mesure d'apparier vos réponses à votre identité. Par conséquent, après cette date, nous ne serons plus en mesure de détruire vos données si vous en faites la demande.

Les données pourront être réutilisées dans le cadre d'autres recherches menées par les responsables cités dans l'encadré au début de ce document. Les données pourront être déposées sur des plateformes « open science » (outil de travail de recherche collaboratif) afin de les partager avec d'autres chercheurs.

En cas de besoin, vous pouvez contacter Madame Aurélie Arnoldi ou Monsieur Simon Jöhr, les responsables de cette étude, aux numéros de téléphone :

+41 76 510 13 13
+41 76 533 52 15

ou par courriel aurelie.arnoldi@etu.unige.ch et simon.johr@etu.unige.ch.

## Publication des résultats et contacts

Les résultats de cette recherche seront disponibles dans les travaux de Mémoire de Madame Aurélie Arnoldi et Monsieur Simon Jöhr, dont la publication est prévue pour 2021-2022. Vous pouvez consulter le site officiel de l'Université de Genève et son Archive ouverte pour plus d'informations sur la publication de ces travaux.

En cas d'intérêt de votre part quant aux résultats de la recherche vous pouvez contacter les responsables de l'étude, cités précédemment. Au vu de l'anonymisation des données, aucun résultat individuel ne pourra être transmis.

## Condition d'arrêt de l'expérience

Vous pouvez arrêter l'expérience de votre propre chef, à n'importe quel moment, en cliquant sur le bouton « annuler l'expérience », si présent, ou en fermant la fenêtre de votre navigateur, ou en quittant la salle d'entretien virtuelle. Vous n'êtes pas tenus d'indiquer la raison qui vous a conduit à cette décision. L'expérience peut s'arrêter également pendant la tâche à cause de problèmes techniques. Dans ce cas, une deuxième connexion vous sera proposée ou vous passerez directement à la phase suivante en fonction de l'état d'avancement dans l'expérience.

## Compensation

L'expérience n'est pas rémunérée.

## Consentement de participation à la recherche

Sur la base des informations qui précèdent, vous confirmez votre accord pour participer à la recherche nommée « Utilisation d'un chatbot et d'un questionnaire d'évaluation » et vous autorisez :

- L'utilisation des données anonymes à des fins scientifiques et la publication des résultats de la recherche dans des revues ou livres scientifiques, étant entendu que les données resteront anonymes et qu'aucune information ne sera donnée sur votre identité ;

- L'utilisation des données anonymes à des fins pédagogiques (cours et séminaires de formation d'étudiants ou de professionnels soumis au secret professionnel) ;

- Que l'entretien soit enregistré

Aurélie Arnoldi
Simon Jöhr

Expérience dans le cadre du Mémoire
du Master MALTT 2021-2022

UNIVERSITÉ
DE GENÈVE
FACULTÉ DE PSYCHOLOGIE
ET DES SCIENCES DE L'ÉDUCATION
Technologies de formation
et d'apprentissage (TECFA)

Vous confirmez avoir choisi volontairement de participer à cette recherche et avoir été informé-e-s du fait que vous pouvez vous retirer en tout temps, sans fournir de justification.

Ce consentement ne décharge pas les organisateurs de la recherche de leurs responsabilités. Vous conservez tous vos droits garantis par la loi.

**Acceptez-vous de participer à cette étude ?**

- ☐ Oui
- ☐ Non

**Date**

.....

**Signature:**

.....

## Annexe 19 - Guide d'entretien

Aurélie Arnoldi
Simon Jöhr

Expérience dans le cadre du Mémoire
du Master MALTT 2021-2022

UNIVERSITÉ DE GENÈVE
FACULTÉ DE PSYCHOLOGIE ET DES SCIENCES DE L'ÉDUCATION
Technologies de formation et d'apprentissage (TECFA)

# Guide d'entretien (version 2.2)

Timing estimé de la passation : 45-60 minutes max.

Structure souhaitée de l'entretien
(adapté de : Goodman et al., (2012), Baxter et al., (2015) :

1. Accueil – introduction – présentation – formulaire de consentement
2. Échauffement – contexte général – explication du déroulement de l'entretien

3. Phase 1 : *activité de test du chatbot*
   a. Questions prétest
   b. Explication du déroulement de la phase 1
   c. Test de l'outil et observation

4. Phase 2 : *entretien post test chatbot*
   a. Explication du déroulement de la phase 2
   b. Corps de l'entretien avec questions génériques puis spécifiques

5. Phase 3 : *évaluation du chatbot*
   a. Explication du déroulement de la phase 3
   b. Définition des critères pertinents d'évaluation
   c. Tris de cartes des items du questionnaire
   d. Tris de cartes des paires de mots du questionnaire
   e. Questions ouvertes sur les items du questionnaire

6. Débriefing et clôture – ajouts supplémentaires – remerciements

Aurélie Arnoldi
Simon Jöhr

Expérience dans le cadre du Mémoire
du Master MALTT 2021-2022

UNIVERSITÉ DE GENÈVE
FACULTÉ DE PSYCHOLOGIE ET DES SCIENCES DE L'ÉDUCATION
Technologies de formation et d'apprentissage (TECFA)

## Formulaire de consentement

Le formulaire de consentement complet est disponible ici.

Renommer participant.e en pseudo dans Zoom.
Attention aux onglets ouverts dans le navigateur des participant.e.s.
Partager l'écran du participant.e.

**Début de l'enregistrement de la session**

L'entièreté du questionnaire est disponible à cette adresse :
https://fpse.qualtrics.com/jfe/form/SV_bORNI181h9kkeeG

## Consignes

L'objectif de la présente étude est d'évaluer la version beta d'un chatbot créé ainsi que son questionnaire d'évaluation.

Vous allez utiliser un outil étant en phase de développement. Sur la base des tests effectués et de vos retours, différentes versions vont pouvoir être produites. À travers cette expérience, nous estimons pouvoir acquérir des informations importantes afin de pouvoir améliorer le chatbot ainsi que son outil d'évaluation.

L'étude se compose de 3 phases :

**1. Utilisation du chatbot**
**2. Entretien post test du chatbot**
**3. Évaluation du chatbot**

**Merci d'avance pour votre participation à notre travail de mémoire !**

Est-ce que vous avez des questions sur le déroulement de l'expérience?
Oui / Non

**(Ne pas expliquer le tout car cela sera expliqué au début de chaque phase)**
**Phase 1 :**

Aurélie Arnoldi
Simon Jöhr

Expérience dans le cadre du Mémoire
du Master MALTT 2021-2022

UNIVERSITÉ DE GENÈVE
FACULTÉ DE PSYCHOLOGIE ET DES SCIENCES DE L'ÉDUCATION
Technologies de formation et d'apprentissage (TECFA)

Dans le but d'évaluer le fonctionnement d'un chatbot, nous vous demandons d'interagir par texte avec le robot conversationnel durant 2-3 minutes.

Durant cette conversation, vous allez répondre à plusieurs questions du robot. Pour certaines demandes, il va vous répondre et instaurer alors un dialogue. Il se peut que, pour certaines demandes, le robot n'ait pas de réponses à fournir. Dans cette situation, n'y faites pas attention et passez directement à la suite.
Durant ce test, nous ne cherchons pas à évaluer vos capacités ou votre rapidité à exécuter les tâches, mais à savoir si le chatbot est fonctionnel.
Pour compléter cela, il est important pour nous que vous **communiquiez à haute voix vos pensées** et que vous expliquiez le mieux possible vos interactions et ceci le plus fréquemment.

Comme expliqué dans le formulaire de consentement, sachez que vos manipulations et vos commentaires seront enregistrés de manière anonyme.

**Phase 2 :**
Nous allons analyser ensemble votre ressenti sur le chatbot.

**Phase 3 :**
Vous allez travailler sur les différents items du questionnaire d'évaluation concernant le chatbot. Durant l'expérience, il est important pour nous que vous **communiquiez à haute voix vos pensées** et que vous expliquiez le mieux possible votre compréhension des items et votre ressenti sur ce que vous êtes en train de faire, et ceci le plus souvent possible.

## 3. a. Questions démographiques/statistiques :

Nous vous remercions de bien vouloir répondre aux questions démographiques :

- Indiquez votre genre :
  *Homme / Femme / Non-binaire ou transgenre / Ne souhaite pas répondre*

- Quel âge avez-vous ?

## 3. b. Consignes test utilisateur :

**Phase 1:**
Dans le but d'évaluer le fonctionnement d'un chatbot, nous vous demandons de converser par texte avec le robot durant environ 2-3 minutes.

Durant cette conversation, vous allez interagir avec le robot et répondre à des questions. Pour certaines demandes, il va vous répondre et instaurer alors un dialogue. Il se peut que, à certaines occasions, le robot n'ait pas de réponses à fournir. Dans cette situation, n'y faites pas attention et passer directement à la suite.

Durant ce test, nous ne cherchons pas à évaluer vos capacités ou votre rapidité à exécuter les tâches, mais à savoir si le chatbot est fonctionnel.

Pour compléter cela, il est important pour nous que vous communiquiez à haute voix vos pensées et que vous expliquiez le mieux possible vos interactions et ceci le plus souvent possible. Comme expliqué dans le formulaire de consentement, sachez que vos manipulations et vos commentaires seront enregistrés de manière anonyme.

Est-ce que vous avez des questions sur le déroulement de l'expérience?
Oui / Non

## 3. c. Questions envisagées **avant** la phase de test du chatbot :

- Avant votre utilisation dans le cadre du cours ADID, pensez-vous avoir déjà utilisé un Chatbot ? Oui / Non / Je ne sais pas

  Si oui, dans quel contexte ?

- Avant votre utilisation dans le cadre du cours ADID, saviez-vous ce qu'était un chatbot (ou aussi appelé « agent conversationnel ») ?

## 3. d. Test utilisateur :

Nous vous invitons à présent, à utiliser le chatbot durant 2-3 minutes au maximum. Merci de bien vouloir communiquer à haute voix vos pensées et expliquez vos interactions aussi souvent que possible.

Afin de démarrer la conversation, cliquez sur le bouton "Commencer" de la fenêtre de dialogue du chatbot (en bas à droite) puis sélectionnez ensuite : "Période à distance (3e utilisation et +)"

Si le bloc de discussion n'apparaît pas, n'hésitez pas à le signaler.

Lien d'accès au chatbot : *http://tecfaetu.unige.ch/etu-maltt/zelda/johrsim0/adid/*

## 4. Questions envisagées **après** la phase de test du chatbot :

- Quelle est votre impression globale du chatbot ?
  *Très mauvaise / Mauvaise / Ni bonne ni mauvaise / Bonne / Très bonne*

Nous allons analyser ensemble votre ressenti sur le chatbot. Les réponses aux questions seront à fournir oralement durant l'entretien.

- J'aimerais que vous me parliez de votre test du chatbot que vous venez de réaliser, qu'en avez-vous pensé ?

- Dans l'éventualité où l'on vous dit qu'un chatbot sera utilisé pour vous soutenir dans le cadre de vos cours du Master MALTT, qu'aimeriez-vous qu'il soit capable de faire (des fonctionnalités en particulier) ?

- Selon vous, que pourrait-il vous apporter d'utile dans vos apprentissages ?

- Qu'attendriez-vous précisément de cet outil ? Avez-vous des besoins spécifiques ?

## 5. Evaluation du chatbot

### a. Explication du déroulement de la phase 3

Tout en gardant à l'esprit votre expérience avec le chatbot, nous allons à présent nous intéresser à comment l'évaluer.

L'expérience est divisée en plusieurs étapes. Les instructions vous seront données par oral et seront également affichées à l'écran.

*Les consignes pourront être affichées durant toute l'expérience en cliquant sur "Voir les instructions" en haut de la page.*

Est-ce que vous avez des questions sur le déroulement de l'expérience?
Oui / Non

### b. Critères d'évaluation

Afin de démarrer l'expérience, entrez votre pseudo puis cliquez sur "Commencer".

Attendez avant de cliquer sur le bouton "Continuer".

Question 1: Selon vous, quels critères d'évaluation devraient être utilisés pour évaluer le chatbot?

Relance: Quelles sont les questions qui pourraient être posées pour évaluer le chabot?

Est-ce que vous avez des questions sur le déroulement de l'expérience?
Oui / Non
Vous pouvez à présent cliquer sur "Continuer".

Voici une liste de critères d'évaluation. Il est possible que vous ayez déjà cité certains des critères affichés ou qu'il en manque. Lisez les critères une première fois à haute voix et donnez-nous vos pensées, impressions, ressenti et compréhension de ces critères. Il est important que vous communiquiez à voix haute le plus souvent possible durant tout l'exercice.

Question 2: Quels critères vous semblent pertinents pour évaluer le chatbot?

Cochez la/les cases des critères qui vous semblent pertinents.

Est-ce que vous avez des questions sur le déroulement de l'expérience?
Oui / Non

Relance: Quelles questions pourraient être posées pour évaluer tel critère?

- Acceptabilité
- Utilité
- Utilisabilité
- Qualité de l'environnement en ligne
- Expérience utilisateur
- Attractivité
- Transparence
- Efficacité
- Stimulation
- Originalité
- ~~Valeur~~
- Confiance
- Crédibilité
- Langage naturel
- Humanité
- Comportement
- Présence
- Emotions
- Motivation

Une fois terminé, cliquer sur "Continuer"

## 5. c. Evaluation du chatbot - Tri de cartes - Etape 1 :

**Vous trouverez à gauche de votre écran des items qui ont pour objectif d'évaluer le chatbot que vous avez utilisé.**

**1) Lisez les items *(Liste de cartes)* et énoncez à haute voix le critère évalué par l'item.**

**2) Classez ensuite l'item dans l'une des 3 catégories qui se trouvent à droite de l'écran selon si vous les trouvez:**

**- Utiles pour évaluer le chatbot**

**- Non pertinents pour évaluer le chatbot**

**- Incompréhensibles**

**N'oubliez pas de nous donner votre compréhension des items. Il est important que vous communiquiez à voix haute le plus souvent possible durant tout l'exercice.**

Est-ce que vous avez des questions sur le déroulement de l'expérience?
Oui / Non

***Les consignes sont disponibles en tout temps en cliquant sur "Voir les instructions" en haut de la page.***

**----> Cliquez sur "Ok" pour commencer l'expérience**

Question: Pour chaque item que vous avez classé dans "Utiles pour l'évaluation du chatbot", selon vous, quelle dimension est évaluée grâce à cet item? (Plusieurs dimensions/critères peuvent être évaluées par un même item)

**Question: est-ce que vous voyez un ou plusieurs autres items qui vous semble pertinent pour évaluer le chatbot à rajouter? (Vous pouvez ajouter un item qui vous semble pertinent à évaluer en cliquant sur l'icône de la bulle).**

***Une fois cette étape terminée, cliquez sur "J'ai terminé". Pour passer à l'étape 2, l'ensemble des items doivent avoir été classés dans les différentes catégories.***

-> Donner le lien Proven By Users (attention max 3 participants par tri de cartes)
**Fonctionnalités Proven By Users:**

- **Vous pouvez dupliquer les items se trouvant dans la partie de gauche de l'écran en double-cliquant dessus. Par exemple, si vous souhaitez classer cet item dans plusieurs catégories.**
- **Vous pouvez créer des sous-groupes à l'intérieur des blocs en cliquant sur (+)**
- **Vous pouvez ajouter un commentaire en cliquant sur la bulle. Par exemple pour rajouter un item qui vous semble pertinent à évaluer.**

- **Vous pouvez créer une nouvelle catégorie en déplaçant un item sur la partie de droite de l'écran, hors des box "Catégorie"**
- **Vous pouvez supprimer un groupe en cliquant sur la corbeille.**

## 5. d. Evaluation du chatbot - Etape 2 :

1) Vous trouverez à gauche de votre écran des paires de mots qui ont pour objectif d'évaluer le chatbot que vous avez utilisé.

2) Classez les paires de mots dans l'une des 3 catégories qui se trouvent à droite de l'écran selon si vous les trouvez:

- Utiles pour évaluer le chatbot

- Non pertinents pour l'évaluation du chatbot

- Incompréhensibles

Est-ce que vous avez des questions sur le déroulement de l'expérience?
Oui / Non

*Les consignes sont disponibles en tout temps en cliquant sur "Voir les instructions" en haut de la page.*

**----> Cliquez sur "Ok" pour continuer l'expérience**

Question: Pour chaque paires de mots que vous avez classé dans "Utiles pour l'évaluation du chatbot", selon vous, quelle dimension est évaluée grâce à cet item? (Plusieurs dimensions/critères peuvent être évaluées par un même item)

**Question: est-ce que vous voyez une ou plusieurs autres paires de mots qui vous semble pertinent pour évaluer le chatbot à rajouter? (Vous pouvez ajouter une paire de mots qui vous semble pertinent à évaluer en cliquant sur l'icône de la bulle).**

***Une fois cette étape terminée, cliquez sur "J'ai terminé". Pour clôturer l'expérience, l'ensemble des items doivent avoir été classés dans les différentes catégories.***

## 5. e. Questions envisagées **après** le tri de cartes – évaluation des items et paires de mots du questionnaire d'évaluation :

**Selon vous, est-ce que l'utilisation de phrases permet d'évaluer le chatbot de façon optimale ?**

*Ne permet pas du tout d'évaluer le chatbot*
*Ne permet pas tout à fait d'évaluer le chatbot*
*Évalue plus ou moins le chatbot*
*Évalue bien le chatbot*
*Évalue tout à fait le chatbot*

**Selon vous, est-ce que l'utilisation de paires de mots permet d'évaluer le chatbot de façon optimale ?**

*Ne permet pas du tout d'évaluer le chatbot*
*Ne permet pas tout à fait d'évaluer le chatbot*
*Évalue plus ou moins le chatbot*
*Évalue bien le chatbot*
*Évalue tout à fait le chatbot*

Est-ce que vous voyez d'autres moyens ou méthodes pour évaluer le chabot?

Quelles pourraient être les pistes d'amélioration pour évaluer le chatbot?

**Relances:**

- Que pensez-vous de l'utilisation de phrases pour évaluer le chabot?
- Que pensez-vous de l'utilisation de paires de mots pour évaluer le chatbot?
- En quoi les items/paires de mots permettent d'évaluer le chatbot?
- En quoi les items/paires de mots NE permettent PAS d'évaluer le chatbot?

## 6. Débriefing et clôture – ajouts supplémentaires – remerciements

L'expérience arrive à sa fin.
Discussion, si nécessaire.

Retour sur certains éléments de l'entretien, si nécessaire.

Remerciements aux participants.

**Fin de l'enregistrement de la session**

# Annexe 20 - Retranscriptions entretiens

## Retranscription des entretiens semi-structurés

**Entretien Beta**

SJ : Est-ce que tu es de retour par hasard ? Je ne t'ai pas entendu...

Bêta : Oui, je suis de retour.

SJ: Ah! Cool! Ok. Super. Alors écoute, je vais t'envoyer un lien dans la conversation, qu'on va te demander de cliquer sur ce lien et du coup de partager ton écran, s'il te plaît. Après on va tout t'expliquer comme ça va se passer.

Bêta : Hmm, voilà.

SJ : Cool ! Super. Je me fais mon affichage à moi. Ok. Alors, est-ce que tu nous entends bien ? Y a pas de souci ? Y a pas d'écho ?

Bêta : Hmm, non, y a pas de souci.

SJ : Ok. Cool. Comme ça on est sûr qu'on peut démarrer dans les meilleures dispositions. Alors, on t'explique. Donc, tu fais, participes à notre étude. On réalise donc dans le cadre de notre mémoire, notre mémoire respectif à AA et à moi. Nous, on s'intéresse sur le développement et l'évaluation de chatbot et aux effets qu'il peut avoir sur les apprentissages. Moi je m'intéresse surtout au développement du chatbot, et AA, elle, développe un questionnaire, enfin une évaluation de ce chatbot. Aujourd'hui, dans l'entretien qu'on va avoir, tu vas...Il va y avoir deux phases. La première, tu vas tester le chatbot, comme tu as déjà fait à distance. Simplement, aujourd'hui, on va te redemander de le tester, mais alors plus au niveau du fonctionnement, de comment tu interagis avec lui, plus qu'au niveau du contenu. Lorsque tu les as...l'as testé les dernières semaines, c'était plus au niveau du contenu, là c'est vraiment voir comment tu l'utilises. Ça c'est la première phase, ce sera plutôt ma partie. Et ensuite on aura une deuxième phase, plutôt la partie d'AA, où tu vas devoir un peu évaluer la clarté, la pertinence d'un questionnaire d'évaluation de chatbot. Donc, c'est ces deux phases que tu vas avoir aujourd'hui avec nous. Ça dure environ 30 à 60 minutes, ça dépendra beaucoup de tes réponses et de comment tu fonctionnes, comment on interagit. Je pense ce sera plutôt dans les 60 minutes. A mon avis, on va plutôt tabler là-dessus comme durée grand max. Déjà première question, t'es bien sur ton ordinateur, donc ça c'est bon. T'es pas sur un appareil mobile. C'est juste par souci d'affichage. Tu...Je vois que tu es sur quel navigateur, tu es sur Chrome.

Bêta : Ouais, je suis sur Chrome.

SJ : Ouais, normalement c'est tout bon. Chrome y a tout qui va marcher. Donc, pour...comme tu l'as peut-être vu dans le questionnaire, enfin le formulaire de consentement, tu peux à tout moment interrompre cet entretien si tu le souhaites, y a vraiment pas de souci. Si tu le sens pas, y a aucune obligation, tu es libre à tout moment d'interrompre, sans donner de raison forcément. Tu peux aussi accéder aux données de recherche ; sache qu'elles seront plus accessibles à partir de janvier l'année prochaine. Donc voilà, si à tout moment y a un problème, tu peux nous alerter, y a vraiment pas de souci. Alors si tu descends, le formulaire tu l'as déjà accepté, donc du coup tu peux passer effectivement à la suite. Alors, pour t'expliquer, on va simplement, le lien qu'on t'a envoyé là sur Qualtrics, ça va nous servir un peu de fil rouge tout au long de notre entretien. ça nous permet d'avoir un endroit où on peut tout avoir les réponses et t'orienter vers les différentes parties de notre entretien. Donc, comme je t'ai dit tout à l'heure, ça va être surtout l'objectif de cet entretien, c'est d'évaluer un chatbot ainsi qu'un

questionnaire d'évaluation. Ils sont tous deux en phase de développement, alors ça peut arriver qu'il y ait encore des petits bugs, des petits problèmes, mais nous, en tout cas, ça va grandement nous être utile, ta participation dans notre phase de recueil de données. Donc, l'entretien va être en trois parties ; on va d'abord utiliser ensemble le chatbot. Tu vas le tester. Après, on va un petit peu discuter comment ça s'est passé. Et après, on va passer vraiment à la partie de l'évaluation avec le questionnaire d'AA. Voilà...Est-ce qu'on a quelque chose encore à rajouter. Cette partie-là, est-ce que déjà tout est clair pour toi ?

Bêta : Euh oui oui.

SJ : Ouais, c'est bon. Hésite pas alors, à tout moment à nous interrompre si t'as des questions, c'est vraiment pas un entretien fermé fermé, t'as le droit de te manifester et de nous poser des questions. C'est vraiment...Faut que tu te sentes libre de nous interrompre, y a vraiment pas de problème. Alors, bah du coup, la première partie c'est juste quelques...deux questions un peu démographiques. Je te laisse répondre. Euh...

[Silence pendant que Bêta répond aux questions sur l'écran]

SJ : Alors, donc, la première phase, comme tu le vois à l'écran, je vais te demander, on va te demander bientôt d'aller retester le chatbot. Je sais que toi tu as déjà testé la troisiè- le troisième scénario à distance. Tu as tout bien fait tes devoirs comme une bonne étudiante. Du coup, malheureusement, on va te demander de le tester, mais cette fois-ci, comme je te l'ai dit tout à l'heure, ce sera pas vraiment au niveau du contenu, mais plus la façon dont tu interagis avec lui. Et il est important, dans cette phase-là de test, que à tout moment tu expliques un petit peu ce que, ce qui se passe à l'écran, que tu communiques un peu tout ce qui te passe par la tête. Alors ça va te paraître un peu bizarre de devoir tout commenter, euh, ce que tu fais avec le chatbot, mais pour nous, ça nous est grandement utile. Euh, donc hésite pas à tout moment de dire, euh là je fais ça, je lui réponds, il me dit ceci, 'fin voilà, c'est vraiment le but de cette recherche. Alors oui, à tout moment il peut, ça peut arriver qu'il bug, comme je pense que tu l'as déjà vu dans le passé, alors à ce moment-là, faut pas faire attention, on passe directement à la suite. Et vraiment le but de cette phase de test, c'est pas de voir si tu, si t'es rapide, si tu réponds correctement, mais c'est vraiment de voir comment tu utilises le chatbot. Y a pas de mauvais ou de bonne réponse, voilà. Est-ce que c'est tout bon pour toi ?

Bêta : Euh ouais.

SJ : Ok. Alors tu peux passer à la suite. Alors de petites questions euh, avant d'utiliser le chatbot. Voilà, alors est-ce tu avais déjà utilisé un chatbot. Ok. Est-ce que tu te souviens dans quel contexte tu avais utilisé ça, c'était sur une plateforme, une application...Si tu t'en souviens, si tu t'en souviens pas c'est pas grave non plus. Ah ouais.

Bêta : Ouais, comment dire. Hmm. [Bêta tape une réponse à l'écran] ...de l'aide ? Je sais pas...

SJ : Ouais. Ouais ouais, mais très bien. Ok. Et du coup, est-ce qu'avant d'utiliser bah euh, un chatbot, je pense que oui, vu que tu avais déjà utilisé avec les CFF Est-ce que tu savais ce que c'était ? [Bêta : Hmm affirmatif] Je pense...Super. Du coup, on va passer à notre test. Je vais te demander d'aller sur le lien, d'ouvrir ce lien, il...normalement, il va t'ouvrir un nouvel onglet, voilà. Et, dans le chatbot, de faire « Commencer », de choisir le scénario numéro trois.

[Silence pendant que Bêta ouvre le chatbot]

Bêta : [à voix basse] Voilà.

SJ : Voilà. Alors là, je vais juste te demander de supprimer, donc d'aller tout en haut dans les, les trois petits...juste...voilà. Ici. Supprimer la conversion, comme ça on est sûrs de partir dès le début. ça peut arriver qu'il reprenne les données de ce tu avais dit précédemment. Alors, « Commencer ». Tu peux effectivement commencer. Et là je vais te demander de choisir, après, le troisième scénario...d'utilisation.

[Bêta choisit le scénario demandé]

SJ : Voilà. Et là du coup, je vais te demander, tout simplement, d'utiliser le chatbot comme tu l'utiliserais à distance et de pas hésiter à communiquer à voix haute ce que tu fais, ce que tu lis, et voilà. Un petit peu tout ce qui te passe par la tête, tout ce qui est utile pour notre entretien.

Bêta : Ok. Alors…

AA : J'ai ton pseudo.

Bêta : Oui ! Mon pseudo.

SJ : Oui, tout à fait. Bien vu AA, ouais. On va te demander d'utiliser aujourd'hui ton pseudo, c'est Bêta, effectivement.

Bêta : Alors, « Pourrais-tu me rappeler ton prénom » Bêta.

SJ : Tu es la participante Bêta aujourd'hui.

Bêta : Hmm... Merci, ma mémoire vive me joue vraiment des tours. (petit rire)

SJ : Donc là si t'arrives à un petit peu communiquer tout simplement...Dire ce tu fais ou...Tout ce qui te passe par la tête, même si des fois, je sais pas ça, tu comprends pas quelque chose, ou il y a quelque chose qui ne va pas, hésite pas à le dire.

Bêta : Mais du coup la troisième, enfin le troisième scénario, ça veut dire que j'ai déjà testé deux fois, je suppose.

SJ : Alors oui, mais la première fois que tu l'as testé c'était vraiment au niveau du contenu et tes réponses étaient importantes, aujourd'hui c'est plus vraiment l'interaction avec le chatbot, comment ça se passe, [Bêta : Ouais enfin- Ok.] l'aspect technique. Si à des moments tu veux répondre, je ne sais pas, par exemple, «je ne sais pas » ou « je n'ai pas envie de répondre », libre à toi, il y a vraiment pas de...Il n'y a pas de...Tu es libre de répondre ce que tu veux, il n'y a vraiment pas de souci.

Bêta : Ok. C'est juste, parce qu'il dit « rappeler mon prénom » alors que...Je sais pas, si c'était la première...

SJ : Effectivement, oui, c'est comme on est dans une troisième utilisation, il est censé t'avoir déjà rencontré auparavant, c'est pour ça. (rire)

Bêta : (lit ce qui est écrit) « Ok. Voilà, je crois que je dois me répéter mais j'ai toujours de la peine à analyser plus de deux réponses qui se suivent. Je suis vraiment désolé. » Ok. « Dites-moi Bêta, comment ça va durant cette période à distance ?» (lit et tape au clavier en même temps) « Ça va bien et toi ?» Hop. Ah mince. Ok. Um... » Comment vis-tu le travail en groupe ?» Euh... » Ça se passe bien, mais... » Qu'est-ce que je pourrais dire ? Hmm... » Mais il y a beaucoup à faire. »

SJ : Donc, je te rappelle si jamais, les réponses que tu donnes aujourd'hui au chatbot ne seront pas du tout prises en compte en tant que telles dans le contenu des réponses. T'inquiète pas, tu peux dire ce que tu veux, ce ne sera pas retenu contre toi. (rire)

Bêta : (lit ce qui est écrit) « Qu'est-ce que tu penses avoir appris durant cette période à distance ?» Euh...j'ai appris...Hmm « Oui ». « Ok je note ! Cette période de travail à distance, justement comment tu la vis ?» (relit le chat) « Qu'est-ce que tu penses avoir appris durant cette période à distance ?», « Ok je note ! Cette période de travail à distance, justement comment tu la vis ?» ...Attends, est-ce que c'est la même ? (relit le chat) « Comment vis-tu le travail en groupe ?» Ah, c'est différent. Um...Comment ? Euh...(écrit) « Bien. »

Bêta : (lit ce qui est écrit) « Alors, pour cette prochaine question, si tu n'as pas envie de répondre, il n'y a pas de souci, je comprendrai. Il s'agit d'évaluer ton humeur. » Bah pourquoi pas. Ah. Hmm... (lit ce qui est écrit) « Je n'ai pas envie de répondre à cela », « Je peux te répondre »

SJ : Tu es libre de choisir ce que tu veux, il n'y a vraiment pas d'obligation.

Bêta : Ok. (lit ce qui est écrit) « Ok cool ! Alors qu'elle est ton humeur actuelle ?» Hmm...Hmm... » Une autre » (lit ce qui est écrit) « Si aucun... » ...Ouais « n'a réussi à exprimer ton humeur, tu peux l'écrire ci-dessous. » (écrit) « Pensive ».

Bêta : (lit ce qui est écrit) « Alors ok. Passons à notre jauge de motivation. » Hmm. Je mets...(écrit) « Ma motivation à quoi ?» Ok. Mouais, ne pleure pas ! Bon, « bye bye ».

SJ : Super. Très bien. Bah écoute, je vais te demander de fermer l'onglet du chatbot s'il te plaît, et de retourner sur Qualtrics.

Bêta : Fermer c'est...là.

SJ : Ouais, et de fermer l'onglet là, ouais, ou d'aller directement sur Qualtrics, si tu veux garder l'onglet derrière, y a pas de souci.

Bêta : Attends-

SJ : Sinon, retourner à l'enquête. Le « Online Survey... » Ouais, voilà, celui-là. Ok ! Et du coup de passer à l'étape suivante.

Bêta : Alors... (lit ce qui est écrit) « Quelle est votre impression globale du chatbot ?» Hmm...

SJ : Alors là, tu as mis un système d'étoiles pour la conversation que t'as eu, mais là ce serait plus au niveau général du chatbot, de ton impression.

Bêta : Hmm. Bah j'ai pas trop d'avis. Enfin, j'ai l'impression- ouais, je ne sais pas.

SJ : Ouais bah mets, tu réponds ce que tu penses, y a vraiment pas de souci. Alors là dans cette deuxième phase, ça va être plutôt des questions un peu ouvertes, on va- il va y avoir des questions qui vont s'afficher à l'écran, et on va juste simplement discuter tranquillement, avoir ton ressenti un petit peu, donc tu peux passer à la suite, et on aura la première question qui va s'afficher.

Bêta : Ok.

SJ : Donc la première question c'est, on aimerait savoir un peu, que tu nous parles de ce test que tu viens de faire du chatbot, et qu'est-ce que t'en as pensé, vraiment comme ça, le feeling que t'as eu. Et tu peux vraiment dire tout ce qui te passe par la tête, et même si t'as rien à dire, c'est pas grave, c'est aussi une réponse de ne rien dire.

Bêta : Ok.

SJ : Qu'est-ce que t'en as pensé, de ce test ?

Bêta : Bah, je...j'ai pas trop compris, je pense, ce qu'il veut vraiment. Enfin, j'ai l'impression de parler pour parler, mais je vois pas trop en quoi il va m'aider. Bah il me demande comment c'est, mais c'est- Enfin, il a l'air d'être pour récolter des informations sur notre ressenti pendant le cours d'ADID, mais...

SJ : Donc c'est ton impression, c'est ça ? C'est la recherche du but, de ne pas vraiment avoir compris à.…Ouais, à quoi il servait, c'est ça ? Si je résume un petit peu (Bêta : Hmm. (acquiesce)) Ok.

Bêta : Est-ce qu'il servait à répondre à nos questions ou pas, enfin, je ne sais pas trop. Je ne savais plus trop.

SJ : D'accord. Ouais, un doute dans son objectif du chatbot. Ok ! Très bien. Bah écoute, c'est noté. Est-ce que par hasard tu aurais d'autres choses, tu aurais pensé à autre chose ?

Bêta : Hmm...

SJ : Dans le test, l'impression que t'as eue...Si tu n'as pas d'autre élément, y a pas de souci, on passe à la suite, hein. Tu sais, tu n'es pas obligée de trouver un élément en plus.

Bêta : Euh...Bah des fois c'était des questions un peu...évidentes, je dirais ? Genre, je ne sais pas quoi, comment dire, juste « oui », peut-être. Je ne sais pas s'il voulait quelque chose de très, d'assez...Il a dit qu'il fallait que ce soit concis, mais est-ce qu'il voulait juste un « oui », ou euh...Justement un « oui, ça va bien ». Enfin une phrase, ou juste des mots.

SJ : D'accord. Donc aussi, tu...ouais, au niveau de ses intentions, on va dire que c'était pas très clair, c'est ça ? Si je résume un petit peu dans les interactions avec lui, c'était un peu...un peu flou. (Bêta : Hmm. (acquiesce)) Est-ce que c'est correct ? Ok, très bien. Super. Bah écoute, on peut passer à la question suivante, s'il-te-plaît.

SJ : Du coup, si je te dis qu'il y aura un chatbot qui pourra être utilisé dans l'année prochaine, par exemple, dans tes cours que tu auras dans ta deuxième année du Master MALTT, qu'est-ce que tu aimerais qu'il soit capable de faire ? Est-ce que tu penses à des fonctionnalités en particulier ? Des choses qui te passent par la tête, si on te dit que tu pourrais avoir vraiment un chatbot qui t'aide au quotidien l'année prochaine. Est-ce que t'as une idée ?

Bêta : Hmm...Euh bah peut-être pour les petites questions un peu.

SJ : Qu'est-ce que tu entends par petites questions ? Ce serait quoi par exemple ?

Bêta : Mouais...Souvent des questions qu'on a peut-être pour les dates de rendus. Des rappels de dates de rendus.

SJ : Ok. Des questions que tu te dis peut-être, je vais pas interagir avec les enseignants ou avec d'autres...Enfin voilà, des questions un peu d'ordre...Je résume, les petites questions pour toi, c'est ça, c'est des questions d'organisation, de-

Bêta : Oui, oui ! Tout à fait.

SJ : Ok, d'accord. C'est pour être bien sûr que je note les bonnes choses. Ok. Est-ce que par hasard tu penses à une autre fonctionnalité, ou à une autre capacité que t'aimerais bien le voir faire, à part de rappeler des dates ou de t'aider à t'organiser ?

Bêta : Peut-être un rappel des objectifs, aussi.

SJ : Yep, super. Une bonne idée. Ok, bah écoute, si tu as pas forcément d'autres pensées, on peut passer à la suite.

Bêta : Ouais c'est bon.

SJ : Et selon toi, qu'est-ce qu'il pourrait t'apporter d'utile dans vraiment tes apprentissages ? Alors, tu me disais là, à propos de l'organisation, pour des rappels de dates de rendu, mais ça c'est pas vraiment l'ordre de l'apprentissage, c'est plus ton organisation vis-à-vis des cours. Est-ce que vraiment, au niveau des apprentissages, tu aurais une idée de ce qu'il pourrait t'apporter ?

Bêta : Bah peut-être pour les concepts, ou les...Ouais, les concepts, surtout.

SJ : Hmm. (acquiesce) Qu'est-ce que tu entends, il viendrait faire quoi avec ces concepts, il viendrait te...t'aider si tu lui poses une question très précise ? Ou ce serait quoi exactement, tu le verrais comment ?

Bêta : Alors, pourquoi pas proposer des, de la littérature à lire sur tels concepts, peut-être ?

SJ : Hmm. (acquiesce) Ou tu lui demanderais, tu lui poserais une question, et puis lui-même t'orienterait vers les choses à lire, par exemple, c'est ça que tu... (Bêta : Hmm ! (acquiesce)) Ok, d'accord.

Bêta : Ouais, plutôt que le forum directement. (rire)

SJ : D'accord. Il viendrait remplacer temporai- Enfin, ok, remplacer, par exemple, le forum. Ok. Très bien, c'est noté. Est-ce que tu as d'autres éléments ? Ou sinon, tu peux passer à la suite, si tu n'en n'as pas.

Bêta : Hmm, non, suivant.

SJ : Et du coup, qu'est-ce que tu attendrais vraiment précisément de cet outil ? Est-ce que toi, spécifiquement pour toi, est-ce que tu aurais des besoins ? Alors c'est vrai que tu disais avant, t'organiser, peut-être te rappeler les dates de rendus, ça devient un peu un besoin spécifique. Est-ce que tu penses à d'autres choses ? Ou là tu parlais aussi d'explications sur certains concepts, vraiment un besoin que tu aurais identifié, ces derniers temps, dans un domaine, dans un cours...Est-ce que tu penses à quelque chose ?

Bêta : Hmm...

SJ : Et pareil, si tu n'as pas forcément de réponse, c'est pas grave. A ce moment-là, on passe à la suite.

Bêta : Ouais, j'en ai pas trop, là.

SJ : Ok, ben pas de souci, tu peux passer, du coup, à la suite. Et bah là du coup, je vais passer la parole à AA, parce qu'on va attaquer sa partie. Je vais couper mon micro pour le moment, mais je reste là, à l'affût.

AA : Ok, alors dans cette partie, on va s'intéresser en fait à évaluer le chatbot, ou du moins, tenter de l'évaluer. Alors, faut bien garder à l'esprit l'expérience que t'as eu de manière générale avec ce chatbot. Ce que j'ai fait, en fait, pour cette évaluation, c'est qu'en fait, j'ai créé...Enfin j'ai «créé», oui. J'ai créé un questionnaire, qu'en fait j'ai déstructuré. Et puis donc on va en fait le structurer ensemble, si tu veux bien, parce qu'en fait j'ai divisé ça en plusieurs étapes. Donc pareil que pour SJ, on a toujours les consignes à l'écran, et puis je les donne aussi par oral, donc hésite pas si t'as des questions. Donc pour y accéder, tu peux cliquer sur le lien.

(La page ne fonctionne pas)

AA : Ah. Ma faute...Pourtant j'avais publié. Alors...

SJ : Ça va marcher. (intelligible)

AA : Arrête, je l'ai dupliqué. (rire)

SJ : Non mais pas de souci.

AA : J'ai eu peur. Ah voilà.

SJ : Ouais mais autrement, si tu veux revenir en arrière, mais du coup je vais- peut-être qu'on va mettre le lien du coup dans le chat de Zoom, comme ça ce sera peut-être plus simple, si tu l'as AA?

AA : Ouais, je l'ai, c'est le même.

SJ : Alors c'est le même. Bon bah du coup, tu peux recliquer sur...ouais, ou ça, bonne idée. Et bah voilà, super. (AA : Ok.) Donc tu es pseudo «Bêta», du coup, pour aujourd'hui.

AA : Ok, alors, à ton avis, sur quels critères on pourrait évaluer le chatbot que t'as utilisé? Alors si tu veux, une piste un peu de réflexion, un chatbot, on peut dire que c'est un outil numérique. Donc voilà, qu'est-ce que tu verrais comme critères qui te semblent importants pour évaluer ce type d'outil?

Bêta : Hmm...Est-ce que je peux dire l'utilité? Peut-être, est-ce que c'est utile ou pas?

AA : Hmm (acquiesce). D'autres ou...?

Bêta : Hmm...Non, pas trop quoi. Genre le seul (intelligible).

AA : Ok, alors sur «Continuer». Donc là, je te propose toute une liste de critères d'évaluation, donc on voit par exemple «l'utilité», que t'as cité, elle est dedans. Donc l'idée en fait, c'est de regarder un peu ces critères, et puis de me dire point par point si tu penses que c'est utile ou pas pour évaluer le chatbot.

Bêta : Ok. Usabilité bien sûr...Euh...Hmm (acquiesce). Efficacité...Hmm...Confiance, oui. Crédibilité....Hmm...Non ça va, ça va. Hmm...Comportement? Comportement, ce serait quoi? Enfin ce serait...

AA : Le comportement du chatbot, vu que c'est un robot, comparé à un humain, enfin, dans ce sens-là

Bêta : Mouais bah...Présence, peut-être. Um...C'est acceptabilité? Hmm...Peut-être ça aussi oui. Et motiv- aussi...Là, ce serait «motivation» des gens qui utilisent, ou du chatbot?

AA : Alors ça peut être les deux.

Bêta : Hmm.

AA : Il y en a une qui serait bien d'évaluer et pas l'autre, ou les deux, ou aucune?

Bêta : Hmm. Ouais. Je pense ouais.

AA : Toutes, ou juste une motivation en particulier?

Bêta : Je pense plutôt la motivation aux gens à utiliser...Après, ça reviendrait à «l'acceptabilité» et «l'utilité», et...»l'utilisabilité».

AA : Oui, il y a un peu tout qui est lié. (rire)

Bêta : Hmm...Voilà, je pense que c'est tout.

AA : Ok.

Bêta : Attractivité...Ouais ça va…Ok.

AA : Ok. Alors, ici on a, sur la partie de gauche, la liste des cartes, des items qui ont pour objectif d'évaluer le chatbot. Et puis, l'idée c'est de classer ces items dans la partie de droite, où il y a trois catégories. «Utile pour évaluer le chatbot», «Non pertinent pour évaluer le chatbot», ou des «Items incompréhensibles». Et c'est important, en fait, que tu nous donnes ta compréhension de l'item, comment tu le comprends, et puis après que tu dises pourquoi tu le classes dans l'une des trois catégories.

Bêta : Ok. Bon. Alors...

SJ : Et si jamais, les instructions, tu peux à tout moment les revoir en haut à droite, si jamais tu as un doute sur ce que tu viens de lire, ok?

Bêta : Oki. Alors «Je suis satisfaite du chatbot», super utile, je p- Ah mince, ça va pas dedans, attends, voilà. Voilà. C'est sur la satisfaction, si t'es satisfait c'est utile. Um... »Le chatbot est plaisant à utiliser». Bah, c'est pareil.

AA : Au niveau de la satisfaction? Ok.

Bêta : Moui voilà, au niveau de la satisfaction et l'utilité. S'il est pas plaisant, pourquoi on va- pourquoi l'utiliser en fait?

AA : Hmm (acquiesce).

Bêta : «J'ai l'intention d'utiliser le chatbot ces prochains mois». Aussi, parce que si personne ne veut l'utiliser, bah...Ça veut dire qu'il n'est pas utile. (rire) «Utiliser le chatbot me permet d'accomplir les tâches plus rapidement» C'est intéressant aussi, hop. Parce que comme ça, on peut évaluer s'il est utile dans les apprentissages des gens. «J'appréhende l'utilisation du chatbot» Um...Bah si, aussi c'est utile, parce qu'il faudrait voir justement qu'est-ce qui freine les gens à l'utiliser. «...fait tout ce que j'attends de lui.» «Le chatbot fait tout ce que j'attends de lui.» Bah c'est pareil. Attends, je pense, ça revient au même que «J'appréhende l'utilisation du chatbot», enfin, c'est, ça fait partie des attentes des apprenants. «Il y a trop d'incohérences dans ce chatbot» Bah c'est encore pareil (rire), ça revient pour les trois. Um... »Je trouve le chatbot facile à utiliser», bah c'est bien ça, c'est...Est-ce que c'est pareil que dans les attentes et l'utilisation? Oui, voilà, il faut qu'il soit facile en fait.

Bêta : «Le chatbot est utile», oui, ben tout à fait, c'est le principal, je le mettrais tout en haut. «Je recommanderais l'utilisation du chatbot à un ami», euh...Est-ce que je...j'hésite. Hmm...Bon, je pense que ça peut être utile, mais ça peut...C'est pas si pertinent que ça, je pense, parce que, ouais, ça veut dire qu'il trouve que le chatbot, il est bien fait, du coup il peut le recommander à un ami, mais je pense que ça a pas la peine pour une évaluation dessus.

AA : Ok.

Bêta : «Ce chatbot est facile à utiliser», bon...Attends, je crois qu'il y a déjà, ah, un c'est «Je trouve», et l'autre c'est «Ce chatbot» «...est facile à utiliser». Euh...

AA : Quelle formulation te semble la meilleure?

Bêta : Hmm...(murmure) «Ce chatbot est facile à utiliser», «Je trouve...» Hmm...Alors là, «Je trouve», il n'y a que moi, peut-être ça, il peut être difficile pour les autres, alors que ce chatbot, il est plutôt en général...Peut-être? Du coup, je dois en éliminer un, c'est ça? Enfin, je dois...

AA : C'est égal. Si tu devais en éliminer un, t'éliminerais lequel, plutôt?

Bêta : Peut-être «Ce chatbot est facile à utiliser», parce que je trouve que c'est plus orienté utilisateur, le «Je trouve».

AA : Ok.

Bêta : «Ce chatbot est facile à utiliser» Ouais, je le mettrais là alors...Mince, je ne sais pas. Hmm...

AA : Oui, tu peux mettre là, pas de souci.

Bêta : (murmure) «Le chatbot est facile à....» (volume normal) «Le chatbot permet de réaliser des tâches en peu d'étapes.» Ça, ça revient à la même chose que...il est où, «Utiliser les tâches- réaliser les tâches plus rapidement». (AA : Hmm (acquiesce)) «...en peu d'étapes», ouais. «J'utilise le chatbot parce qu'on me le demande». Hmm...Bof. Voilà. Voilà, ça évalue pas trop le chatbot, on va dire. Quoique. Ouais, on va dire ça. «Travailler avec le chatbot est amusant», oui, ça fait partie de la satisfaction, je dirais. Oh, «Le chatbot est plaisant à utiliser, travailler avec le chatbot est amusant». Hmm. Ok.

Bêta : «L'utilisation du chatbot me rend plus efficace». Oui. Est-ce que... (fredonne) tu~tu~tu...Est-ce qu'il y a quelque chose de similaire? Hmm...Efficace? L'efficacité, est-ce qu'il y a quelque chose déjà dedans, ou pas? Non, bah alors du coup c'est pour évaluer l'efficacité. Bah quoique, en fait, ça rejoindrait ça, en fait aussi, ouais, et l'efficacité.

Bêta : «J'utilise le chatbot parce que les autres apprenants l'utilisent». Hmm...Je dirais que non. Euh...»Avec le chatbot c'est plus facile d'apprendre». Oui, ça, ça rejoint...les tâches...à réaliser (AA : Hmm (acquiesce)). C'est plus facile, c'est de la facilité. Voilà, pour évaluer la facilité. «Le chatbot me fait gagner du temps». Du coup voilà, ça rejoint aussi...Mais est-ce que c'est utile à évaluer? Attends. «me fait gagner du temps»...Mouais, en fait, je le mettrais plutôt dans «Non pertinent», parce que ça dit pas vraiment comment ça te fait gagner du temps...Enfin...

AA : Trop vague, quoi.

Bêta : Ouais, c'est un peu vague, voilà.

Bêta : (a du mal à lire l'intitulé) «Je trouve que le chatbot utile...» «est utile pour...» Attends...Ouais, c'est ça. «Je trouve que le chatbot est utile pour mon apprentissage». Ouais, mais du coup ça remplace ça...Hmm...»Le chatbot est utile» et «Je trouve que le chatbot est utile pour mon apprentissage». Hmm....Ouais, du coup je remettrais «Le chatbot est utile» là.

AA : Est-ce qu'il serait trop général, c'est ça?

Bêta : Hmm (acquiesce). (AA : Ok.) Bah ça, pour l'évaluer...Ouais, parce que je pense, pour...Hmm...Ouais, en fait, non je pense...en fait ouais, faudrait que je le remette...Je pense que ça, ce serait...Euh...l'évaluation de l'utilité, enfin, comment moi je trouve si c'est utile ou pas, et l'autre c'est, si...Est-ce qu'il donne les informations assez utiles ou pas. Enfin, est-ce que...Je sais pas si je me suis bien...(AA : Ok, ouais ouais).

AA : De manière comme l'autre, en fait. Le premier c'est général, et dessous c'est plus un avis personnel.

Bêta : Ouais, c'est ça, ouais. Alors...»Mon interaction avec ce chatbot est claire et compréhensible». Bah oui, sinon...Sinon il n'y en a plus...Enfin, je pense que je le mettrais, «clair et compréhensible», ouais, je le mettrais tout en haut, parce que s'il n'y aucune compréhension, bah il n'y aura pas d'utilité, il n'y aura rien, en fait.

Bêta : «Les différentes fonctionnalités de ce chatbot sont bien intégrées». Hmm...Est-ce que ça rejoint ça? Hmm...»...de ce chatbot sont bien intégrées». Bah oui, ça fait partie du fonctionnement. «Le chatbot correspond à ce que j'ai besoin». Du coup ça, c'est pour revenir à la satisfaction.

Bêta : «Le chatbot est plaisant à utiliser». Euh...»Je me suis senti très en confiance en utilisant ce chatbot». La confiance...Oui. Mais comment on évaluerait ça? Hmm...Oui, c'est important, ça dépend des informations qu'il donne en fait, c'est vrai. C'est vrai parce que s'il nous des informations...euh...Pour prendre l'exemple de la littérature qu'il pourrait proposer, s'il nous propose des choses étranges à lire, c'est pas...C'est pas super.

Bêta : Alors...»Je voudrais utiliser ce chatbot fréquemment» Bah...bof. Enfin, ouais. Enfin, je l'utilise...Enfin, que l'apprenant utilise quand il veut, quoi, quand il a besoin. «J'aime utiliser le chatbot». Hmm, ça j'hésite. Bah je dirais que «non». Enfin, «Non pertinent», parce que...Hmm...Pourquoi, pourquoi. «J'aime utiliser le chatbot». Bah, je trouve que l'utilité est plus importante...Enfin, est-ce que tu en as besoin, et du coup tu l'utilises, plutôt que «j'aime» ou «j'aime pas», en fait.

AA : Ok.

Bêta : «Le chatbot rend l'apprentissage plus intéressant». Hmm...»plus intéressant», non...ouais, est-ce que c'est intéressant, l'apprentissage est intéressant? Umm...Hmm...Mouais, je regarderai après. Euh...»Le chatbot me permet d'accomplir les tâches plus facilement». Mais attends, ça revient à «en peu d'étapes plus facilement»...Hmm, ça revient à ici mais...Attends, je ne vois pas la différence entre les trois. Euh...»Le chatbot me permet d'accomplir les tâches plus facilement», «me permet d'accomplir les tâches plus rapidement», «tâches en peu d'étapes»...Bon, du coup ça rejoint les trois, je dirais alors. (AA : Ok.) La facilité, la rapidité et...

Bêta : Du coup, «Le chatbot rend l'apprentissage plus intéressant». Ça évaluerait...Hmm...Qu'est-ce que ça évaluerait...»rend l'apprentissage plus intéressant»...Hmm...Mouais, je dirais «Incompréhensible». Je ne sais pas qu'est-ce qu'on pourrait évaluer sur ça. Mouais. (AA : Ok.) Du coup voilà. Je pense.

AA : Ok. Alors tu peux cliquer sur «J'ai terminé». C'est parfait, parce que du coup t'as bien expliqué à chaque fois ce que tu veux penser que ça concernait, donc c'est parfait.

Bêta : Oki.

AA : Alors, là, il y aura exactement le même exercice, mais cette fois avec des paires de mots. Donc pareil, chaque paire de mot, elle devrait en principe évaluer le chatbot. Donc bien lire à chaque fois la

carte et puis de dire qu'est-ce que ça pourrait évaluer. Toujours en gardant aussi un peu à l'esprit les autres...les critères qu'on a vu juste avant et puis encore dans la liste. Parce que des fois, les mots, ils ne sont peut-être pas clairs. Donc quand c'est pas clair, les mettre dans «Pas compréhensible» et puis après définir si c'est utile ou pas utile pour évaluer le chatbot.

Bêta : Ok ok. Alors...»Naturel - robotisé» Ouais. Mais du coup ça, bah pour moi, le fait qu'il soit un robot, c'est pas perturbant, mais pour certains ça pourrait être pertinent. Hmm...Bon, je dirais que majoritairement, peut-être que les gens sont, ils sont...Ça leur bloque un peu si c'est trop robotisé, je dirais, alors je mettrais dans utile.

Bêta : Um...Bah ça, super utile. «Pratique» ou «Pas pratique», mais est-ce que...bah ouais. C'est tout ce qui...Enfin, ce que je disais aussi. Alors...»Imprécis» et «Précis». Hmm, oui, parce que plus il donne des réponses précises et complètes, plus c'est mieux. Bah c'est ce qui manquait un peu, je pense, à la version d'ADID bêta, là.

AA : Et ça, ça évaluerait du coup...l'aspect contenu, plutôt, c'est ça?

Bêta : Ouais, le contenu des réponses par exemple. Peut-être quand je disais par exemple, bah «Ça va et toi?», bah il avait pas pris en compte le «et toi?», peut-être, du coup.

AA : Ok.

Bêta : «Innovant - traditionnel». Dans tous les cas, c'est innovant, du coup je dirais que c'est «Non pertinent». Hmm, quoique, peut-être «traditionnel» pour revenir si le chatbot parle comme les enseignants, mais ça reste innovant, du coup non. «Laborieux - aisé». Du coup ça, ça serait la facilité ou pas? «Laborieux - aisé»...»aisé», c'est que ce soit...Hmm. Je pense que j'ai pas compris les termes, du coup...

AA : Ok.

Bêta : «Stressé - serein». Euh...Peut-être «Non pertinent»...Quoique. Hmm...Hmm...Mouais, c'est vrai que si le robot nous stresse à répondre, c'est pas super. Faudrait qu'on ait le temps de bien formuler les phrases.

AA : Du coup ça évaluerait quoi?

Bêta : Si la...Est-ce que le robot arrive à mettre à l'aise le...les apprenants. Du coup, ça, ce serait pareil en fait, du coup. Hmm...»Impatient» et «Calme». Mais du coup, ça, ce serait au niveau de l'humeur du robot...Je mettrais ça, ouais, évaluer pour...l'humeur du robot. (AA : Ok). Et ça, ce serait au niveau de l'apprenant.

Bêta : Euh...»Soporifique - stimulant». Du coup ça, bah c'est ce...Bah justement, les phrases du chatbot. «Mal préparé - bien préparé» Bah ça aussi. Ça, ce serait au niveau du contenu donné par le chatbot. «Conservateur - innovateur» ? Ouais, ce serait encore...Ouais. Bah ce serait «innovateur» encore, enfin ça resterait...(AA : Hmm (acquiesce)) Du coup, «Non pertinent». Euh...»Surchargé - ordonné», oui. Bah par exemple, à la fin, quand il m'a donné plein de messages, c'était un peu surchargé.

Bêta : Euh...alors «Non logique - logique», oui, c'est important. Faut que ça suive les réponses des apprenants. «Dépourvu d'imagination» et «créatif» ? Hmm...Ça va, «Non pertinent». C'est juste le caractère du robot quoi. Genre, enfin, justement. «Inefficace» et «efficace». Du coup oui, je le mettrais en haut aussi. Avec «Efficace - pratique». Voilà. «Satisfait - insatisfait». Ça, c'est au niveau de l'apprenant, s'il est satisfait ou pas du robot. Et «bénéfique», pour l'apprenant aussi, si ça lui apporte quelque chose.

Bêta : «Actif - passif». Hmm...Hmm...Je dirais «Non pertinent», parce que...Enfin, une conversation, je pense que ça reste actif, je suppose. A moins que les réponses du robot soient très soporifiques, justement. «Rassurant - déroutant». Ouais, ça, ça va avec…(fredonne) tu~tu~tu...ça va avec...le caractère du robot, s'il est impatient et calme. Enfin, s'il arrive à mettre à l'aise les apprenants. «Conventionnel» et «original». Hmm...Bah, je dirais, ça va justement avec l'innovation et tout. «Lent -

rapide». Oui, je pense que c'est important. S'il prend trop de temps à charger ses réponses, bah on va s'endormir aussi.

Bêta : Euh...»Pas sympathique - sympathique» Bah...oui, j'ai pas envie d'un robot pas sympathique et méchant, qui me répond mal. Enfin, parce que je pense que je suis une...enfin, ouais. Oui oui, c'est important, parce qu'on veut pas parler avec des gens pas sympathiques un peu. Hop là... »Pas intéressant» et «intéressant». Hmm...Hmm. Est-ce qu'il serait pas intéressant? Qu'est-ce qu'on pourrait évaluer, le contenu ou le robot? Hmm...Bah du coup, je dirais que pour le contenu du robot, faudrait qu'il, que ce soit intéressant.

Bêta : «Sérieux - pas sérieux». Ça je sais pas. Je mettrais dans «Non compréhensible». Ah...oui, d'accord, je vois. Si...ouais. Ça serait peut-être...Comment le robot...Le langage qu'utilise le robot, et moi je dirais...Bah en fait, ouais. «Non pertinent», parce que je pense que l'évalua- enfin, le chatbot sera fait en fonction d'un public, et puis voilà. Ah. Quoique. Hmm. Ouais. Ouais, c'est vrai, parce que peut-être qu'il y a des gens qui n'aiment pas ne pas se faire vouvoyer. Hmm. Bon, j'hésite, du coup je le mettrais là. (rire)

AA : Ok.

Bêta : «Désagréable - agréable». Bah ça revient à «Pas sympathique - sympathique». (AA : Hmm (acquiesce)). «Environnement réel» et «Environnement fictif». Hmm...Hmm...Bah ça reste, ouais, en fait c'est relié avec le réel, mais c'est un environnement fictif, le robot, il est sur une page, c'est fictif. Mais en même temps, il dit du contenu qui est en rapport avec le cours. Euh, bon, moi je mettrais «Non pertinent». Non, je dirais même «Non compréhensible», parce que je comprends vraiment pas «L'environnement réel».

Bêta : «De mauvaise qualité - de bonne qualité». Oui, bah il faut...Enfin, on espère que ce soit de bonne qualité. Enfin, voilà. (prononce faux la paire suivante) «Incompréhensif - compréhensif», bah il faut que ce soit compréhensif, euh, compréhensible. «Fâcheux - réjouissant». Ouais, du coup, ça revient au «Sympathique - Pas sympathique». «De mauvaise...» Attends, ça...Je crois qu'elle est déjà passée, celle-là, non? Attends...

AA : Ah ouais, elle est à double. On l'a trouvée! (rire)

Bêta : (rire) Oui ! Alors...»Ludique» et «Académique». Hmm...Est-ce que c'est important d'évaluer ça? Non, je ne pense pas que c'est important. Ça, ça revient à.…à voilà, la caractéristique du robot. (AA : Hmm (acquiesce)). «Non fiable» et «Fiable», oui. Je mettrais en haut. Hop, voilà. «Ennuyeux» et «Passionnant». Euh...(fredonne) tu~tu~tu~tu~tu...Hmm. Bah je dirais que oui, ,mais après je ne sais pas comment on pourrait l'évaluer. Enfin, parce que...Je pense pas que je m'attends que le robot donne un cours ou...Enfin, du contenu totalement passionnant. Enfin, je veux juste qu'il réponde à mes questions. Du coup, je vais revenir à.…Ouais, je vais l'associer avec «Naturel - robotisé». (AA : Ok.) Ouais, selon les attentes des apprenants, s'ils veulent vraiment quelque chose de très naturel, je pense pas qu'ils veulent quelque chose d'ennuyeux et ils veulent quelque chose de passionnant.

Bêta : «Déçu» et «stimulé». Mais du coup ça reviendrait à «Ennuyeux - passionnant». Non, quoique. «Déçu»...Bah oui, je pense. Je pense...Bah c'est ça...voilà. C'est pour évaluer les attentes, bah, justement, s'ils s'attendaient à quelque chose d'autre, les apprenants.

AA : Ok.

Bêta : Voilà!

AA : Alors tu peux mettre «J'ai terminé».

Bêta : (répond au message «Merci» affiché sur l'écran) De rien!

Bêta : Du coup j'appuie sur «Suivant» ?

AA : Oui, tu peux. Alors, un peu le même style de la question de SJ tout à l'heure. On a vu que...Enfin, j'ai fait exprès de séparer en deux parties, l'utilisation de phrases d'une part. Est-ce que tu penses que

les phrases, elles permettent d'évaluer le chatbot optimal, et du coup, j'anticipe déjà, la deuxième question, ça sera en rapport avec les paires de mots. Comme ça, tu peux peut-être nuancer tes ré-...ta réponse. Je le mettrais la prochaine fois sur les deux pages. Enfin, les deux questions sur la page, si jamais, SJ, pour la prochaine fois.

Bêta : Hmm (acquiesce).

AA : Voilà. Est-ce que l'utilisation des phrases, elle permet d'évaluer le chatbot de façon optimale?

SJ : Donc le premier tri de cartes que t'as fait.

Bêta : Ouais. Euh...Je dirais que oui. Mais après, il y avait les deux...la façon justement, en «je», ou la façon justement «le chatbot est...». Du coup...ouais, faudrait, je pense choisir un des deux ou, je ne sais pas...C'est...Parce que ouais, c'est après...Je ne sais plus trop ce que...Enfin, c'est trop subtil. (rire) (intelligible) ...pas un peu ce que j'évalue. Est-ce que j'évalue ma perception du robot, ou justement, j'évalue justement le robot en général. (AA : Ok.) Mais les phrases, je trouve que ça fait...ça évalue bien. (AA : Ok.) Du coup j'appuie sur «Suivant», c'est ça?

AA : Oui. Et puis du coup même question. Bah voilà, au niveau de l'utilisation des paires de mots. (Bêta : Hmm (acquiesce)). Est-ce que ça permet d'évaluer optimalement le chatbot?

Bêta : Hmm...Bah c'était souvent vrai, mais justement, si on évaluait plutôt, bah justement le chatbot, qu'est-ce qu'il faisait, ou qu'est-ce que les apprenants faisaient sur le chatbot, ou qu'est-ce que...Enfin, c'est beaucoup plus vague. Du coup, je mettrais plutôt là. Mais après...Mouais. Ouais, c'est assez libre, et...ouais. Et du coup ouais, après ce serait plus dur à interpréter et à analyser la pensée de celui qui fait l'évaluation.

AA : Alors, est-ce que tu vois d'autres moyens ou d'autres méthodes que les phrases et les paires de mots pour évaluer le chatbot?

Bêta : Hmm...Hmm...Peut-être… (fredonne) tu~tu~tu. Peut-être la fréquence de l'utilisation du chatbot, peut-être? Ou non, ou peut-être, est-ce que le chatbot répond aux qu...ouais. Peut-être l'évaluation dans le chatbot en lui-même, est-ce que justement il...(bruit de notification) Oups! Il...Il....gnignigni. Désolée, le bruit m'a perturbée! Ah, pourtant j'avais bien éteint, mais...Euh...

SJ : En fait, juste pour reprendre ce que tu dis, si quand tu converses avec le chatbot, s'il répond à tes demandes, c'est ça? Que tu voulais, s'il répond correctement à tes questions? (Bêta dit quelque chose en même temps, intelligible)

Bêta : Hmm (acquiesce). Et si on...ouais, bah du coup ça. Puis voir dans le chatbot s'il répond, et voilà. C'est une bonne évaluation, je pense.

AA : De faire, comme en fait, comme un test de...Enfin, d'utilisabilité, dans le sens où tu poses une question, et d'évaluer si la réponse est bonne ou pas bonne? C'est ça? En gros?

Bêta : Hmm (acquiesce). (AA : Ok.) Je pense, ouais. Enfin, j'arrive pas à me projeter, vraiment. (rire)

AA : Ok. Tu peux mettre «Suivant», et puis, dernière question, c'est, est-ce que...Donc là on a vu des paires de mots et des phrases, est-ce que tu vois des pistes d'amélioration pour ce questionnaire? Enfin, t'as déjà dit peut-être de bien comprendre qu'est-ce que qu'on évalue, si on évalue le chatbot ou est-ce qu'on évalue l'appre...Enfin, comment se sent l'apprenant, est-ce qu'il y a d'autres choses, peut-être, qui te viennent, ou...?

Bêta : Um, bah peut-être le design...Enfin, je savais pas trop, justement, parce que moi, ça me dérange pas...Mais...Je crois qu'il y avait...Ah oui, je crois que c'est apparu un peu, parce que...Euh...Je ne sais plus si c'était dans les paires de mots, je crois que c'était «est-ce que c'était visuel ou pas»...Non, je ne sais plus c'était quoi les termes. Mais...

AA : D'évaluer le design, c'est ça?

Bêta : Ouais, le design. Ah, est-ce que c'était «Conventionnel», ou (marmonne) «pas conve...» (voix normale) Ah je sais plus si c'était quelle paire de mots, mais ouais, juste...juste...

AA : Il n'y avait pas assez d'éléments là-dessus, en fait, c'est ça?

Bêta : Ouais, voilà, sur le design, peut-être. Bah est-ce que...Est-ce que...ouais, est-ce que ça plaît.

AA : Donc plaire au niveau de l'utilisateur? (Bêta : Hmm! (acquiesce)) L'esthétique?

Bêta : Oui, l'esthétique, plutôt. (rire)

AA : D'autres choses qui auraient manqué, ou...?

Bêta : Hmm...Non, je crois c'est tout.

AA : Ok. Alors tu peux cliquer sur «Suivant», puis je laisse SJ conclure. Il aime bien!

SJ : Effectivement, bah du coup, on arrive à la fin de cette expérience. Euh...Avant de se quitter ou de couper l'enregistrement et discuter encore un petit peu entre nous, est-ce que tu as des points sur lesquels tu as des incompréhensions ou sur lesquels t'aimerais qu'on revienne un petit peu...il y a des choses qui t'ont durant tout l'entretien, des choses que t'aimerais qu'on ré-aborde, que t'aurais peut-être pas compris sur le moment, ou...Tout ce qui te vient à l'esprit, est-ce que t'as quelque chose? Comme ça.

Bêta : Hmm...Hmm...Enfin, je pourrais juste faire un petit commentaire sur les tris de carte, et que je pense, au début j'étais assez claire dans les tris, puis...Je pense vers les fins de cartes, je...Il y a tout qui commençait à se mélanger dans ma tête.

SJ : (rire)

AA : Bah bizarrement, j'ai trouvé...j'ai...Enfin, j'ai cru que le...les paires de mots te plaisaient plus...Enfin, te parlaient plus, parce que j'ai trouv...Alors je ne sais pas si c'est l'exercice qui a fait que le premier bah...On s'entraîne, et puis le deuxième on a l'habitude, mais je croyais que t'as beaucoup plus parlé…Enfin, c'est...Je sais pas, SJ, après tu me diras mais, t'as plus parlé sur les paires de mots que sur les phrases. J'ai trouvé personnellement.

SJ : Hmm (acquiesce). C'est vrai, y a de ça...Parce que les phrases, tu voyais tout de suite le mot...la compréhension, et du coup tu disais tout de suite ce que c'était, et tu les classais, ouais. Tandis que sur les mots, t'avais plus de...de...Enfin, je vais pas dire, plus de choses à dire, mais t'avais effectivement plus de ressenti, tu le faisais plus...Tu le communiquais un peu plus, ouais, effectivement. Il y avait de ça.

Bêta : Ouais. Non, en fait je voulais parler de la longueur des cartes.

AA : Ah, la longueur?

SJ : Ouais, et puis la quantité des cartes? (Bêta : Ouais la quantité. (intelligible)) C'est ça ce que tu dis, la quantité- ok ouais.

AA : Mais finalement, du coup, t'en n'as pas tellement, enfin, enlevé, vu que t'as beau- Il y en a quand même pas mal que t'as jugé utile, c'était aussi ça que c'était intéressant à regarder. (SJ : Hmm (acquiesce)) C'est qu'il y a beaucoup, mais en même temps il y a beaucoup de choses que tu dis (rire).

Bêta : Ouais (rire).

SJ : C'est vrai qu'il y avait beaucoup de choses finalement-

AA : C'est mon dilemme quotidien (rire).

SJ : Non, bah écoute, très bien. Est-ce que...Voilà, je pense que c'est tout bon. On a... En tout cas, nous, on a toutes les réponses qu'on voulait, donc on va te remercier pour ta participation, c'était cool. Je vais arrêter l'enregistrement, si c'est ok pour AA. Je vais faire «Stop», tac.

**Entretien Zeta**

SJ : Bienvenue, merci de ta participation pour notre formulaire, pardon pour notre entretien. Tu as donc déjà rempli le formulaire de consentement que tu nous as déjà envoyé donc c'est tout bon.

À un moment donné je vais t'envoyer un lien et à la première page il va te demander, pour des raisons administratives, d'à nouveau de valider si oui ou non tu veux participer. Saches qu'en tout temps tu peux, tu peux décider d'arrêter cet entretien sans donner de raison, ça y a vraiment pas de soucis. Hum… C'est la règle, t'as le droit, à tout moment, comme si t'as un problème technique ou quelque chose qui se passe chez toi tu peux arrêter à tout moment. Euh faut que tu te sentes libre de quitter ce Zoom si t'as vraiment besoin, tu peux, y a vraiment pas de soucis.

Zeta : Ok.

SJ : Hum du coup je vais t'expliquer un peu comment ça va se passer, tu vas retrouver un lien dans la conversation de Zoom. Je vais te demander de cliquer dessus, ça va t'ouvrir normalement une fenêtre.

Voilà. Donc en fait on va, tout le long de l'entretien, on va se baser sur un formulaire dans Qualtrics. Qui va nous con-, nous servir un peu de fil rouge pour cette, pour cette étude. Et du coup, durant cette étude, comme tu le sais sans doute, on l'a déjà expliqué auparavant, c'est donc dans le cadre de nos mémoires de master avec AA. Moi je m'intéresse au développement du chatbot, et elle principalement à l'évaluation de ce chatbot ainsi qu'aux effets sur, qui peut y avoir sur les apprentissages.

Donc aujourd'hui, on va de nouveau tester ensemble le chatbot. Je sais que tu l'as déjà testé à distance, mais là on va le tester simplement à l'écran, et ce sera principalement au niveau de l'interaction avec lui, de la façon dont tu l'utilises, plus qu'au niveau des contenus. Si au niveau des contenus tu veux répondre à chaque fois « non », « je ne sais pas », ou « je n'ai rien à dire » y a vraiment pas de soucis c'est pas ça qu'on va regarder aujourd'hui. C'est vraiment comment tu interagis avec lui. Et dans une deuxième partie, qui est la partie qui est un peu plus liée à, au travail d'AA qui est dans l'évaluation, tu vas avec elle, hum, évaluer quelques items d'un questionnaire et la pertinence de l'évaluation.

Zeta : D'accord.

SJ : Hmm en moyenne, euh pour te dire, on a mis entre 30 et 60 minutes. C'est plus dans les 45 minutes en ces dernier euh ces derniers, ces dernières passations, pour te donner un peu un ordre d'idées. Mais après ça dépend aussi de tout c'que t'auras à nous dire euh ou pas, ça fera le temps de l'entretien.

Zeta : Okay.

SJ : Donc tu es bien sur un appareil, euh tu n'es pas sur un mobile tu es bien sur un ordinateur. Euh, du coup si tu descends un peu plus bas dans la fenêtre tu verras que normalement euh y a pas de soucis, ah et oui en tout temps, si tu as besoin d'accéder à tes, à cet entretien, à ces données de recherche, tu peux nous demander sans soucis, hum, d'ici à ce qu'on fasse notre soutenance en février l'année prochaine tu es en droit de demander également à accéder à tes données. Humm elles seront en tout cas pour l'instant stockées auprès de Gaëlle Molinari, notre directrice de mémoire.

Zeta : Ok.

SJ : Donc si c'est tout bon pour toi, on peut passer à la suite, que je t'explique.

Hum du coup euh, la première partie ça va être principalement au niveau du, de, on va s'intéresser au développement du chatbot. J'aurai besoin d'avoir un peu, euh, sur la base de ton test, de quelques retours. Du coup ça va se comporter en 3 phases. La première tu vas de nouveau utiliser le chatbot, après on va juste en discuter à travers quelques questions, 3 / 4 questions. Et enfin on passera à la partie évaluation euh du chatbot. Hum donc la partie dorée. Est-ce que par hasard tu aurais déjà des questions à poser euh à ce propos ou est-ce que c'est plutôt clair pour toi ?

Zeta : C'est clair Hmm.

SJ : Ok cool. Euh Hmm pareil à tout moment, s'il y a quelque chose tu, tu comprends pas forcément ou t'as besoin de précisions t'hésites pas à me le dire.

Zeta : Ok.

SJ : Alors du coup on peut passer à la suite. Euh bah les questions démographiques pour commencer, ça c'est un p'tit peu obligatoire.

Super.

Du coup la première phase ça va être donc le teste euh du chatbot, y, il va d'abord y avoir quelques petites questions un peu de contexte et euh, après, une fois qu'on, on ira retester ensemble le chatbot, le plus important ça va surtout être que tu communiques un peu à voix haute ce que tu vas être en train de faire. Euh les exemples ça peut être « Ah bah là le chatbot il me pose une question, j'lui réponds » euh ou bien « Je n'ai aucune réponse à lui fournir euh on passe à la suite », ou ce genre d'exemple-là.

Zeta : Ok.

SJ : C'est pas vraiment au niveau du contenu. Alors ce sera, tu verras, le même scénario euh que tu as eu à distance avec le chatbot, ce sera des questions liées à, à, un peu à ta situation à distance vis-à-vis du cours ADID. Mais comme je t'ai dit avant, si tu veux répondre que des « non » ou « je ne sais pas » y a vraiment pas de soucis tu peux aussi faire ça. Euh du coup on va passer à la suite.

Donc il y a quelques petites questions pour commencer. Euh donc avant d'utiliser le chatbot, est-ce que tu en avais déjà utilisé un à ton avis ? Ça peut être euh sur internet, sur euh…

Zeta : Alors euh oui, oui oui.

SJ : …applications, jeux. Alors des fois on…

Zeta : Je les mentionne ?

SJ : Non non tu peux, alors de toute façon si tu mets oui il va te demander dans quel contexte tu les as, tu les as vu donc euh déjà c'est bien.

Zeta : Ah ok.

SJ : C'est un peu pour savoir ben du coup ouais dans quel contexte tu te souviens les avoir utilisés, c'était dans quel exemple ?

Zeta : Alors c'était pour euh un changement de forfait euh avec mon opérateur mobile.

SJ : Ok.

Zeta : En tout cas ça euh c'est l'exemple qui me vient. Après peut-être aussi pour un magasin de bricolage il me semble.

SJ : Ok. Et euhm moi j'ai une petite question du coup à ce propos. C'était donc sur internet ou sur euh une application mobile ?

Zeta : Sur internet.

SJ : Sur internet. Et quand tu as interagi avec le chatbot, c'était explicite, tu savais que c'était un chatbot, ou euh tu pensais que tu t'adressais plutôt à un humain euh?

Zeta : Alors euh… Ça m'est déjà arrivé de penser m'adresser à un chatbot, enfin à un humain pardon, et que ce soit en fait toujours euh le chatbot.

SJ : Ok d'accord, c'est juste pour savoir, une petite précision, pas de soucis. Alors avant de passer à la suite, je vais dire bonjour à AA haha.

Zeta : Bonjour AA.

SJ : Par contre on t'entends-, ah voilà

AA : Bonjour.

SJ : Ah on t'entend cette fois-ci. Euh bah alors comme tu le vois à l'écran, on est, bah on a commencé ma partie, juste que je situe, on est donc avec le participant Zeta et euh donc on va passer à la suite et on va commencer le test je pense du chatbot. J'crois qu'il y a encore j'crois une question, vas-y tu peux faire « Suivant ».

Euh voilà avant ton utilisation dans le cadre du cours, est-ce que tu savais ce que c'était un chatbot, euh bas avec tes réponses effectivement tu savais déjà ce que c'était. Super.

Ok. Alors du coup, on va, hum, tu vas pouvoir cliquer sur le lien d'accès au chatbot, qui est à l'écran. Hum et tu vas euh du coup appuyer sur « Commencer » et après tu vas choisir le troisième scénario : « La période à distance ». Et je te rappelle qu'aujourd'hui tu es participant Zeta s'il te demande d'entrer… Ah oui là je vais te demander de supprimer la conversation, pour ça tu cliques, exactement, sur le petit menu des trois… voilà… « Supprimer la conversation ». Comme ça on est sûrs que on repart sur une bonne base, voilà.

Alors tu peux commencer. Et là je te rappelle de communiquer un peu le euh le, toutes tes pensées…

Zeta : Ouais, ça marche.

SJ : … ce que tu vois à l'écran, ce que t'as envie de dire, plus que le contenu. Voilà donc c'est ok ? il va te demander. Et du coup la dernière utilisation ce sera « période à distance », troisième utilisation, tout en bas, voilà. Et du coup, c'est parti.

Zeta : Alors. Je m'appelle Zeta.

SJ : Exactement.

Zeta : « Dis-moi Zeta comment ça va durant cette période à distance ? » Euh… ça va bien, je te remercie.

« Comment vis-tu le travail en groupe ? » Alors euh… Durant la période à distance actuelle, euh j'ai des bons groupes de travail, en tout cas ça fonctionne bien. Donc j'vais simplement marquer euh je euh le vis bien.

SJ : Ok, très bien.

Zeta : « Et qu'est-ce que tu penses avoir appris durant cette période à distance ? » Alors, j'ai euh conti-, disons j'ai mis en pratique certain contenu euh du cours ADID notamment, sur euh la la dimension collaborative. On avait vu pas mal d'exemples euh de travail collaboratif et dans le cadre du projet, qui est cette année un projet authentique, euh on, on est amenés à, Hmm voilà à sélectionner, justifier euh certaines modalités de travail en pratique. Donc j'vais marquer euh… Euh…

SJ : En résumé parce que….

Zeta : Voilà euh euh les modalités, des modalités de travail collaboratif.

SJ : Très bien.

Zeta : « Cette période de travail à distance euh justement comment tu la vis ? » Euh bah je la vis bien. Même si on doit tout faire en 4 semaines au lieu de 5.

« Pour cette prochaine question, si tu n'as pas envie de répondre, il n'y a pas de soucis je comprendrais. Il s'agit d'évaluer ton humeur. » Je vais te répondre.

Alors euh « plaisir, espoir, fierté, colère, honte, désespoir, ennui, autre ». Bon. Euh je vais marquer « plaisir » qui me paraît être la plus, la plus positive.

Euh donc euh ok je valide euh je skip.

SJ : Exact.

Zeta : « Passons à notre jauge de motivation ». Euh. Bon. J'vais être sincère, on va quand même mettre 3 sur 5. C'est quand même la fin de l'année…

SJ : Après je, j'te répète si tu veux, si tu veux pas répondre ou si tu veux mettre 5 alors que t'es à 1 ou le contraire y a vraiment pas de problème.

Zeta : Ah oui oui c'est vraiment pas central ouais.

SJ : Non non. Et là il s'agit vraiment d'un test, les données que tu mets aujourd'hui dans le chatbot seront pas, pas récoltées en tout cas.

Zeta : Ouais, d'accord.

SJ : C'est vraiment plus l'interaction que je, qu'on évalue.

Zeta : Donc j'vais répondre à sa dernière demande, si j'ai des questions concernant les contenus. Euh. Pas de questions particulières.

« Je ne peux pas répondre à cela ». Ah oui donc là je suis renvoyé au forum des échanges si jamais j'avais euh un truc qu'il a pas compris. « On va s'arrêter là … », ok. « Je suis plus un boomer. Dernier truc si tu pourras me dire comment s'est passé notre échange, choisi parmi les étoiles » Ben j'vais mettre euh 3.

SJ : Ok. Alors du coup tu peux euh ben soit lui dire au revoir, ou bien tout simplement retourner du coup sur l'onglet de Qualtrics, ton deuxième onglet.

Zeta : Ok.

SJ : On finit le test. Donc tu peux passer à la suite.

Du coup, quelle est ton impression globale du chatbot euh en général, pas forcément le, la, le test que tu viens de faire. Ton impression…

Zeta : Euh... J'vais mettre euh… « Ni bonne ni mauvaise ».

SJ : Ok, très bien. Hum… Maintenant on a, y va y avoir des questions qui vont apparaître à la suite à l'écran on va simplement en, en discuter ensemble. Hum ce sera des réponses orales t'auras rien besoin de, d'écrire. Ce sera juste simplement j'aurai besoin d'avoir ton, ton ressenti.

Zeta : Ok.

SJ : Si tu veux passer à la suite.

Du coup première question. Tu euh, tu pourrais un tout p'tit peu nous parler du test que tu viens de fai-, de réaliser, qu'est-ce que t'en as pensé ? Vraiment là le test que tu viens de faire euh, euh, sur l'onglet précédent.

Zeta : Alors euh… Vis-à-vis du test, en fait c'est une expérience assez ludique on va dire d'utiliser le chatbot, donc ça c'est la première chose, c'est, c'est, y a un côté rigolo. Mais ça, ça n'a pas vraiment

euh, ‘fin ça ne m’apporte rien de spécial à mon niveau, en fait c’est, je le fais plus euh bah parce que ça, ça vous aide dans votre travail. Mais euh, je, je, je réponds aux questions euh, pour répondre aux questions, voilà c’est un peu ça mon, mon ressenti général. Mais après euh, oui c’est, en tout cas le chatbot de manière générale c’est un outil intéressant et sympathique, et pis là ben j’ai l’impression que ce qui est central en termes de contenu, s’il faut parler des contenus, c’est euh de retracer euh, en fait de, de, c’est un questionnaire amélioré sur euh ce qu’on a pu, sur une évaluation d’un cours en fait après le cours, voilà.

SJ : Ok, très bien. Ben je te remercie. Tu peux passer à la suite.

Prochaine question : si je te dis qu’un chatbot sera utilisé dans la suite de ton master MALTT dans d’autres cours, qu’est-ce que tu aimerais qu’il soit capable de faire ? Est-ce que tu penses à des fonctionnalités en particulier ?

Zeta : Oui euh, je pense à, à aider, enfin à permettre le partage de réflexions, et à réagir de manière contextualisée selon le type de réflexion euh faite par l’utilisateur. Euh… Notamment ben si on doit, si on n’a pas bien saisi un contenu de cours, que p’t’être le chatbot puisse reprendre des éléments, nous rediriger vers des, des liens qui euh…

SJ : Hmm Hmm.

Zeta : Voilà, qui euh… Si, si c’est par reconnaissance de, de mots-clés hein, j’sais pas exactement comment ça fonctionne mais, qui, que ça puisse nous suggérer des ressources pour qu’on puisse euh reprendre des notions, aller plus loin sur certains apprentissages par exemple.

SJ : Ok.

Zeta : Euh bon après euh… La dimension émotionnelle euh j’me vois pas trop euh… partager mes émotions avec un chatbot.

SJ : Hmm Hmm.

Zeta : Euh j’en ai ni l’envie ni l’besoin, c’est, c’est mieux j’pense à ce niveau-là de parler directement avec le professeur quand c’est nécessaire.

SJ : Ok.

Zeta : Donc oui ce serait plus, voilà, pour euh, pour euh, améliorer des apprentissages on va dire.

SJ : Ok, très bien. C’est noté. Tu peux passer du coup à la suite.

Zeta : « Selon vous… »

SJ : Euh voilà, « qu’est-ce qu’il pourrait t’apporter d’utile dans tes apprentissages » euh bah du coup tu, tu viens de le répondre dans cette amélioration dans les apprentissages.

Zeta : Hmm Hmm.

SJ : Effectivement, dans cette, euh, ces suggestions d’aller plus loin, plus dans… Je note en tout cas ça, comme ça ça me permet de… À moins que t’aies autre chose à ajouter dans ce domaine-là ?

Zeta : En fait je, je parle de ça en me disant que c’est quelque chose qui serait réalisable, bien sûr dans l’idéal il faudrait que ce soit euh un robot qui répond directement à des questions euh, euh, euh, de, de, des questions conceptuelles par exemple mais euh j’imagine que, que que, bah c’est quelque chose d’impossible à réaliser donc ça sert à rien à mon avis que euh, qu’on, que j’ajoute d’autres éléments, parce que voilà c’est… dans ce qu’il me semble réalisable.

SJ : Rien n’est impossible à réaliser mais je pense que… haha. Non tout à fait c’est une bonne remarque et, et c’est vrai que j’ai noté cet aspect-là.

Du coup tu peux passer -

Zeta : Disons alors, si on euh, si on se fait euh, si, si on, on est un peu dans l'idéal, l'idéal serait que ce soit un tueur informatisé en fait…

SJ : Hmm Hmm.

Zeta : ...On pourrait l'appeler comme ça.

SJ : Hmm Hmm, ouais, tout à fait. Qu'il ait un rôle de tuteur, tout à fait. Ok.

Hum, et du coup précisément est-ce que t'aurais un besoin spécifique pour toi, vis-à-vis d'un cours, vis-à-vis d'un, d'une tâche actuellement ? T'aurais une idée euh, qui te viendrais…?

Zeta : Au, au-delà de ADID ?

SJ : Oui exactement, alors en général, qui pourrait être utilisé via un chatbot ?

Zeta : Bon il me semble que, à plusieurs reprises, que ce soit moi ou d'autres étudiants de ma volée, on avait des doutes sur euh des consignes de, de tâche à réaliser à distance. Et donc on était invités à les… À les formuler soit sur l'EdutechWiki, soit euh les forums Moodle. Est-ce que le chatbot il pourr-, ne pourrait pas être un outil qui centralise euh, euh, ces aspects ? C'est-à-dire que si on a un doute on s'adresse au chatbot qui vient nous clarifier les consignes de travail à distance.

SJ : Hmm Hmm.

Zeta : Donc clarifier des modalités d'évaluation, ce genre de chose.

SJ : Ok, très bien. Oui c'est, c'est un bon exemple en effectivement. Euh tu peux passer à la suite, s'il te plait. Ok alors là on arrive à la partie d'évaluation, et du coup je vais laisser la parole à AA qui va tout t'expliquer la suite.

Zeta : Ça marche.

AA : Euh donc c'est dans la continuité finalement de euh, de ce qu'on a fait avec SJ. Donc là on va essayer d'évaluer le chatbot. Donc toujours bien garder à l'esprit euh, l'expérience du chatbot.

J'ai divisé la, l'expérience en deux étapes. Et pis tu peux euh, cliquer sur le lien, probablement. Voilà.

Alors euh le pseudo ben c'est ton nom euh d'attribution d'aujourd'hui. Alors euh, est-ce que déjà, tu vois euh, des critères d'évaluation qui permettraient d'évaluer le, le chatbot ?

Zeta : Hmm. La première chose qui m'vient c'est la pertinence des réponses données.

AA : Hmm.

Zeta : Euh... La… La pertinence slash spécificité. Donc c'est son degré de reconnaissance des, des, des contenus. Et puis euh… Peut-être, bah y a des critères ergonomiques quand même aussi mais qui sont plus secondaires donc euh… En premier lieu c'est l'aspect contenu, en deuxième lieu y a l'aspect interface. J'ai mentionné que c'est une expérience ludique, peut-être que… Plus l'interface est aussi jolie, agréable à utiliser, plus on a l'impression de s'adresser peut-être aussi à une vraie personne, plus on va l'utiliser.

AA : Hmm. Ok.

Zeta : Je continue ?

AA : Voilà. Du moment que t'as, si t'as fini tu peux continuer.

Alors ici je te propose une liste de critères d'évaluation. Euh… Donc euh… En fait l'idée c'est de… cocher ceux qui te paraissent euh pertinents et puis expliquer pourquoi, et puis à l'inverse si tu les coches pas aussi expliquer bah pourquoi tu les as pas cochés.

Zeta : Bon alors je coche les trois premiers parce que ça recouvre ce qui a trait à l'ergonomie si j'ai bien retenu les contenus du cours haha. Euh… Bon, donc. Là est-ce que j'ai besoin de détailler ou… Ou c'est simplement de dire que c'est ce que je mets dans ergonomie donc euh…

AA : Ok.

Zeta : … En lien avec le plaisir, la facilité, et la pertinence de… Pour l'utilisateur à utiliser l'interface.

AA : D'accord.

Zeta : Bon « qualité de l'environnement en ligne » je le coche pas parce que ça me paraît un peu redondant avec euh les précédents.

AA : Hmm.

Zeta : « Expérience utilisateur » pareil. Hmm… « Attractivité »... Alors j'vais regar - j'vais juste les parcourir un peu…

Est-ce qu'on est limité dans le nombre de choix ?

AA : Non.

Zeta : Ok. Bon alors je vais quand même mettre celui-là [Expérience utilisateur]. Euh… « attractivité »... En fait c'est oui jusqu'à un certain point donc en fait je le priorise pas. « Transparence » euh je vois pas trop dans quelle mesure ça a besoin d'être transparent. Peut-être sur l'utilisation des données mais pareil euh c'est pas ce à quoi j'aurai pensé en premier lieu.

« Efficacité ». « Stimulation » non. « Originalité » non parce que c'est pas forcément nécessaire d'être original. Euh « confiance »... « Crédibilité ».... « Crédibilité » je le mets aussi, dans la même optique là d'a- d'apporter des réponses pertinentes.

« Langage naturel », « humanité », Hmm… « Humanité » oui dans la formulation euh… des phrases c'est, c'est, c'est j'pense que c'est un peu la logique que vous avez suivie et c'est appréciable justement de, d'avoir l'impression qu'on, qu'on s'adresse à quelqu'un à qui on pourrait s'adresser par exemple euh, euh en étant à l'université donc c'est un plus quand même.

« Comportement », « présence », « émotion », « motivation »... Ceux-là je, je les mets pas.

AA : Ok.

Alors là on arrive à la, l'expérience à proprement parler. Donc en, en fait sur la gauche tu verras des, des items qui ont pour objectif de, d'évaluer le chatbot. Donc euh, bah ça s'appelle des listes de cartes, dans la partie de gauche. Donc l'idée ce serait de les lire, et puis euh de dire selon toi qu'est-ce, qu'est-ce que cette phase elle évalue.

Zeta : Ouais.

AA : Euh et puis… Alors, soit tu le fais, tu le fais, tu lis et tu dis chaque critère qu'est-ce que ça évalue. Ou tu peux simultanément, après, faire ça et puis classer, puisque la deuxième étape en fait c'est d'y classer euh, soit un item que tu trouves utile pour évaluer le chatbot, soit selon toi qui est pas pertinent et pourquoi, ou alors un item que tu comprends pas le placer dans la colonne donc « incompréhensible ».

Zeta : Ok. Très bien. Je clique sur ok ?

AA : Oui.

Zeta : Les, les différents critères je les ai pas sous les yeux hein du coup ?

AA : Non.

Zeta : Parce que je suis pas sûr de me souvenir exactement des termes.

AA : Non après c'est - voilà les termes c'est pas grave. C'est vraiment l'idée qu'il y a derrière le… En fait, comme t'as fait juste avant sur les critères.

Zeta : Ouais.

AA : Voilà, juste énoncer euh…

Zeta : Bon donc celui-là, « je trouve le chatbot facile à utiliser », c'est euh… Ah ok, faut peut-être pas le décaler tout de suite.

AA : Tu peux, faut juste le mettre dans le, dans le bleu.

Zeta : Directement, ok. Alors celui-là, « je trouve le chabot facile à utiliser », je mettrais utilisabilité, 'fin je rapporte au critère utilisabilité, et donc utile pour l'évaluation.

« Le chabot me permet d'accomplir les tâches plus facilement » euh… Oui ben par rapport à ce que j'ai mentionné en tout cas ça me paraît utile, et là euh le critère c'est euh… Ben utilité.

« Il y a trop d'incohérence dans ce chabot ». Donc euh… Bon je, je suis un peu mitigé là. Donc là c'est peut-être au niveau de la pertinence justement des réponses. En tout cas je sais que je l'ai fait boguer à la première - lors de la première utilisation. Hmm mais bon c'est parce que j'avais pas coché un des onglets, j'avais euh plutôt rédigé quelque chose donc il avait pas aimé. Après il me semble que là à deux reprises, euh, il m'a dit quelque chose là, qu'il arrivait pas à traiter la réponse donc peut-être qu'on peut, euh, euh ouais on peut le mettre aussi « utile » en termes d'itération en fait ouais. Euh.

AA : Pour évaluer…

Zeta : Donc le critère c'est la pertinence, et utile pour évaluer le chabot parce que… 'Fin… En fait ouais, la, la formulation je, j'ai un peu du mal avec la formulation dans le sens où euh, c'est, c'est un état de fait. Ce serait plutôt euh, 'fin ouais, vous voyez ce que je veux dire. C'est que les incohérences actuelles sont à soulever dans le cadre de - pour une itération pour améliorer le chabot donc euh, voilà, je le mets ici.

AA : Hmm Hmm.

Zeta : Euh. « Plaisant à utiliser »... Je, j'vais parcourir un peu les autres aussi pour me faire une idée de l'ensemble.

AA : Ok.

Zeta : [*Lecture des autres items*]

Ok. Alors euh, au vu de la formulation, je vais plutôt classer « Il y a trop d'incohérences dans ce chat - chabot » dans « incompréhensible » parce que je vois pas vraiment euh, quelle est la visée de, de la phrase.

« Le chabot est plaisant à utiliser » donc là c'est euh, le, l'attractivité. Euh Hmm… Et… Je trouve ça important dans un, un deuxième niveau, on va dire.

« Les différentes fonctionnalités de ce chatbot sont bien intégrées »... Ouais, utile pour l'évaluation donc euh je vais le mettre là.

«J'appréhende l'utilisation du chatbot » euh non pertinent pour l'évaluation. J'pense pas que c'est vraiment une source d'appréhension comme outil.

« Le chatbot fait tout ce que j'attends de lui » euh bah utile pour l'évaluation mais, mais à ce stade je suis pas d'accord avec la formulation alors je sais pas si vous pouvez me préciser… Est-ce que ça veut dire qu'on est d'accord avec la formulation quand on classe dans « utile » ou est-ce que c'est simplement que c'est un critère utile.

AA : Euh alors...

Zeta : C'est, par exemple la finalité, ce serait la finalité de l'outil.

AA : En fait ça serait utilisé dans le questionnaire, et puis on demanderait à la personne de se positionner sur une échelle.

Zeta : Ah oui. Ouais ouais euh. L'échelle de… J'crois qu'on l'avait utilisée. Entre 1 et 7 là.

AA : Voilà par exemple. Et puis après si la formulation bah elle est pas compréhensible, c'est mieux de le mettre dans « non compréhensible ».

Zeta : Mouais. Non non j'vais le mettre dans « non pertinent » parce que on lui demande pas de faire tout ce que, ce qu'on attend haha, ça doit être ciblé.

« J'utilise le chabot parce qu'on me le demande ». Ben, oui je suis d'accord mais euh, mais en termes d'évaluation je trouve pas pertinent. Euh par rapport - faudrait que je mentionne chaque fois le critère d'évaluation c'est ça ?

AA : Euh, oui idéalement qu'est-ce que ça, qu'est-ce que ça évalue en fait, pis du coup en fait euh la réponse « normalement » devrait orienter le choix dans l'un des trois euh... Carrés.

Zeta : D'accord, donc là ce serait peut-être la polyvalence, euh 'fin disons pour celui-là là, le « il fait tout ce que j'attends de lui ». La polyvalence ça me paraît pas prioritaire.

J'utilise ensuite le chabot parce qu'on me le demande, là je… C'est vrai. Je le mets dans « Incompréhensible ».

Euh « l'utilisation du chabot me rend plus efficace », donc, euh… On va dire ça c'est donc l'utilité, et c'est utile pour l'évaluation.

AA : Hmm Hmm.

Zeta : « J'ai l'intention d'utiliser le chabot ces prochains mois », euh… Ce serait peut-être aussi sous utilité ? Et donc aussi utile pour l'évaluation.

« Mon interaction avec le chabot est claire et compréhensible »... Euh là c'est donc la, l'inter - la lisibilité de l'interface. Y a peut-être un autre terme mais ce qui me vient c'est là, la lisibilité.

« Avec le chabot c'est plus facile d'apprendre », donc, oui. Et là c'est à nouveau l'utilité, d'après moi.

« Je suis satisfait du chabot ». Je le relirai aussi à l'utilité. Utilité / plus-value en fait, plus-value de l'outil par rapport à, à ce qu'on a ou pas actuellement.

« J'utilise le chabot parce que les autres apprenants l'utilisent », euh, ça c'est euh, non c'est pas pertinent. Donc là c'est, j'sais pas euh, effet de mimétisme ou quelque chose comme ça, j'vois pas trop euh pourquoi ce serait pertinent.

« Je trouve le chabot utile pour mon apprentissage », utilité, utile.

« Je recommanderai l'utilisation de chabot à un ami ». Euh, pareil.

« Je me suis senti très en confiance en utilisant le chabot » ça c'est peut-être le, les, au niveau des appréhensions. Pff pour moi euh c'est pas pertinent.

« Travailler avec le chabot est amusant », oui donc là c'est euh, un peu bah, le, lud-, ludicité, ou euh, utilisab-, ouais, aspect ludique que je trouve quand même euh, ouais, intéressant.

« Le chabot est utile ».

« Je voudrais utiliser ce chabot fréquemment » non, ce, c'est pas forcément pertinent (inaudible!!), mais c'est ouais, une question vraiment d'utilité, lorsqu'on a besoin, donc j'mets pas cette phrase dans… Je la mets dans euh « non pertinent ».

« Le chabot permet de réaliser des tâches en peu d'étapes » donc là c'est euh, oui c'est utile pour l'évaluation parce que ça a trait en fait à la, Hmm, la pertinence et à la réactivité. En fait le moins de temps on a besoin de passer dessus pour trouver réponse à notre question le mieux c'est.

« J'aime utiliser le chabot » donc euh oui, ben là c'est un peu euh, multiniveau hein. Y a l'utilité mais aussi l'aspect ludique et l'aspect esthétique.

AA : Hmm Hmm.

Zeta : « Le chabot me fait gagner du temps ». Donc euh… C'est euh, pareil à l'avant dernière phrase. Donc euh… Ouais on, oui on, on pourrait formuler le critère gain de temps, économie de temps aussi.

AA : Hmm Hmm.

Zeta : Même si euh... Ouais. Oui. Si. Oui oui. C'est, c'est pertinent.

« Le chatbot rend l'apprentissage plus intéressant»... Non c'est pas son rôle, je pense. Ça c'est le rôle du, du professeur. Lui il a un rôle plus de tuteur...

AA : Hmm Hmm.

Zeta : … D'accompagnant de l'é, de l'apprenant.

« Le chabot correspond à ce que j'ai besoin » donc, utilité. Utile. Enfin, pertinent.

« Ce chabot est facile à utiliser » donc euh, oui utilisabilité, je le mets dans pertinent.

Et « Utiliser le chabot me permet d'accomplir les tâches plus rapidement » bon, c'est pareil, économie de temps.

AA : Hmm hmm.

Zeta : Donc on peut le mettre là.

AA : Ok. Est-ce que tu vois par hasard des, des dimensions qu'on aurait pas évaluées ?

Zeta : Euh… Bon, la dimension « pertinence » par rapport à la formulation des phrases…

AA : Hmm Hmm.

Zeta : Je sais pas si là les, les expressions elles recouvrent ça ou pas. Mais on-, oui, j'pense que oui. J'pense que oui. Euh… L'aspect ludique on en a parlé. L'aspect euh… Euh non j'crois que c'est, j'crois que c'est bon comme ça.

AA : Ok. Alors tu peux cliquer sur « j'ai terminé » pour passer à la dernière phase.

Donc c'est exactement le même exercice, seulement que, au lieu des phrases on a des paires de mots.

Zeta : D'accord. Alors «fâcheux - réjouissant »... Euh, je parcours un p'tit peu hein le, les étiquettes…

[*Lecture des items*]

Ok. Donc « fâcheux - réjouissant » je le mets pas… En fait c'est, c'est une question de priorisation hein. Donc c'est pas que je trouve non-pertinent, mais c'est pas prioritaire. C'est pas le, les émotions pre-, pri-, 'fin premières qu'on aimerait euh faire ressentir je pense quand on développe ce genre d'outil.

« Innovant - traditionnel » donc, oui, tout ce qui va recouvrir l'utilité je vais le mettre dans pertinent et… Utilité / plus-value donc innovant oui.

« Imprécis - précis » je le mets dans utile.

« Inefficace - efficace » pareil.

«Surchargé - ordonné » euh… Ah là on est plus sur euh… Sur le, le, l'interface euh, le… 'Fin l'aspect de l'interface. Mais oui c'est clair que quand même on va le mettre là euh dans, dans utile parce que… Si c'est surchargé, je pense que les utilisateurs ils, ils délaissent l'outil au bout d'un moment. Donc il y a ce risque-là.

« Incompréhensible - compréhensible » ben pareil, ça bon c'est important quand même.

« Ludique - académique » on le met aussi. Parce que justement ben, celui-là il est intéressant parce que, c'est une manière peut-être plus ludique de, de renforcer des apprentissages ou de, de comprendre certaines choses, que dans le cadre d'un cours, par exemple. 'Fin c'est, c'est un, un outil qui peut avoir un aspect ludique.

«Soporifique - stimulant » pff... Celui-là… Non. Celui-là c'est… Non. Je le mets dans non-pertinent parce que euh, en fait la, la stimulation c'est, c'est, c'est le, à l'ini-, à l'initiative de l'utilisateur de, de solliciter le, le chatbot, et pas l'inverse.

« Laborieux - aisé » oui, utilisabilité.

« Impatient - calme » Euh ça non.

« Mal préparé - bien préparé » ben oui euh, bien préparé par rapport à la pertinence des réponses.

« Naturel - robotisé » par rapport à la formulisa-, formulation des phrases, donc aussi je, je trouve utile.

« Pas sérieux - sérieux »... Alors oui, faut que ça ludi-, ludique mais sérieux. Donc on peut le mettre.

AA : Hmm Hmm.

Zeta : « Satisfait - insatisfait » oui.

Euh… « Conservateur - innovateur » je le mets aussi parce qu'il y a innovateur.

« Rassurant - déroutant » ça ça touche plus à nouveau les émotions donc j'crois pas que c'est prioritaire.

« Non logique - logique » oui, par rapport à la pertinence.

« Stressé - serein » plus euh, émotions, donc non.

« Actif - passif »... Non à nouveau, vu que ça, 'fin les questions doivent être posées par le, l'utilisateur, et on demande juste au chatbot de répondre de manière ciblée, mais pas d'être proactif en fait dans… 'Fin pas, peut-être pas à ce stade. Oui, bien sûr idéalement oui, mais euh, je le mets dans non-pertinent parce que si on est pragmatique, c'est, c'est pas ça qu'on lui demande.

« Environnement réel - environnement fictif » Hmm… Pff non, je le mets dans non-pertinent parce que… Ou dans incompréhensible » parce que, 'fin… Parce que euh, 'fin je vois pas trop le, la dimension environnement dans ce genre d'outil. Parce que c'est, c'est très euh, situé j'veux dire, comme outil. Donc en fait c'est, c'est, c'est des boîtes de dialogue, donc l'environ-, 'fin ouais, je… J'sais pas. Spontanément pour moi l'environnement, c'est par exemple l'environnement d'un jeu vidéo, et j'vois pas le lien là du coup trop avec euh, avec cet outil.

« De mauvais qualité - de bonne qualité » Ben… Oui, de bonne qualité, ça voudrait dire que euh, il est innovant et adapté, donc je le mets là.

« Déçu - stimulé » je le mets dans non-pertinent pour les mêmes raisons que précédemment.

« Non fiable - fiable » je le mets dans utile.

« Pas intéressant - intéressant »... Oui, ben… Plus il est pertinent plus il sera intéressant donc euh, on, on peut considérer que c'est utile.

« Pas sympathique - sympathique »... Oui, aller, pour l'aspect ludique.

« Désagréable - agréable » pareil.

« Lent - rapide » oui. Utilité. 'Fin utilisabilité.

« Conventionnel - original »... Pff… Oui aller. Je le mets parce que euh, original pas forcément en termes d'interface par rapport à d'autres chatbot mais par contre par rapport au contenu traité euh, innovant, original, ça me paraît aller dans, un peu ensemble.

« Ennuyeux - passionnant» c'est un peu too much, donc non-pertinent.

« Dépourvu d'imagination - créatif » je le mets dans, dans non-pertinent. Parce que vu-, pareil comme je disais tout à l'heure quoi, c'est, c'est plutôt le, l'utilisateur qui doit être proactif donc on ne demande pas à l'outil d'être créatif.

« Non bénéfique - bénéfique » oui pour l'utilité.

Et « Pas pratique - pratique » oui pour l'utilisabilité. Voilà.

AA : Ok. Est-ce que tu vois d'autres paires de mots par hasard ?

Zeta : Oh pardon.

AA : Y a pas de soucis.

Zeta : Euh… Ben… N-Non, j'crois que ça en fait déjà pas mal.

AA : Ça marche.

Zeta : Sincèrement j'en aurai pas rajouté.

AA : Alors tu peux retourner sur le Qualtrics.

Voilà. Et cliquer sur « suivant ».

Alors. Euh là je pose la question en, j'ai fait exprès de couper les phrases et les paires de mots, parce que j'aimerai savoir un peu ton ressenti par rapport à qu'est-ce qui évalue le mieux, euh, le chatbot.

Zeta : Hmm Hmm.

AA : Donc ça peut être euh autant l'un que l'autre ou euh est-ce qu'il y a quelque chose plus que euh l'autre ou euh voilà.

Zeta : Euh Hmm… Bon moi j'préfère l'utilisation des phrases je pense. Parce que vous pouvez mieux personnaliser, donc euh… ça évolu-, évalue bien. Alors que les paires de mots… Euh… Pff. C'est intéressant oui, pour des, des graphiques euh donc je le mets, je le mets plus ou moins bien.

AA : Hmm Hmm.

Donc là, pareil que tout à l'heure, les… Les questions s'affichent à l'écran et puis on discute, pas besoin de taper la réponse.

Zeta : Hmm Hmm.

AA : Est-ce que… Bah, outre le fait des phrases et des paires de mots, est-ce qu'il y aurait d'autres moyens, d'autres méthodes pour évaluer le chabot ?

Zeta : Bah, bon, je pense que des, des entretiens euh, avec, avec questions semi-ouvertes sont important parce que euh, il y a quand même le ressenti utilisateur à la suite de l'utilisation donc c'est, je pense qu'on est quand même limité avec ce genre de questionnaire, et que si l'outil devient vraiment un outil d'accompagnement à la réflexion, compréhension de contenu, résolution de problèmes finalement de, de, de l'étudiant, de l'apprenant, euh alors il y a quand même euh une dimension ressenti qui est intéressante à partager. Donc si je, voilà si on peut faire rentrer là-dedans les entretiens semi-ouverts, ça me paraît complémentaire. Ouais.

AA : Ok.

Zeta : Mais individuels hein, y a pas de… De focus group ou autre je pense à envisager.

AA : Ok.

Et pis, du coup, alors là, là on parlerait que du questionnaire d'évaluation. Est-ce que y aurait, bah, quelque chose qu'on pourrait faire pour améliorer ? C'était un peu déstructuré hein mon questionnaire mais, 'fin voilà y aurait ces quelques items euh rangés d'une certaine manière avec des échelles, et est-ce que euh, voilà, est-ce qu'il y aurait des pistes d'améliorations ? Alors je, j'imagine bah, pour reprendre ce que t'as dit avant, ben d'intégrer une partie euh, plus euh entretiens individuels.

Zeta : Hmm Hmm.

AA : Et pis est-ce qu'il y a autre chose ?

Zeta : Ben j'sais pas, j'ai été un peu déstabilisé par la, les, les euh, les boîtes, 'fin les, les intitulés de boîtes qui disaient « utile pour l'évaluation » ou « non pertinent pour l'évaluation ». Euh Hmm… Mais bon après, enfin après quand tu m'as détaillé un peu le truc je, ça allait mais… J'sais pas est-ce qu'il faudrait pas les formul-, les, les, les intituler « d'accord - pas d'accord » ?

AA : Hmm Hmm.

Zeta : Dans un premier temps, peut-être faire un pré-tri.

AA : Hmm Hmm.

Zeta : Puis « utile » ou « inutile » ou « pertinent - non-pertinent pour l'évaluation ».

AA : Ok.

Zeta : Mm… Je, je, je, je sais pas. 'Fin faudrait euh, réfléchir si euh, le faire en deux temps ça rallonge un peu est-ce c'est, est-ce que ça a une plus-value. Mais vu que j'étais un peu déstabilisé j'sais pas c'est ce qui, c'est la réflexion qui me vient en tout cas.

AA : Hmm Hmm. Très bien. Voilà haha.

Zeta : Ok.

SJ : Cool. Ben merci beaucoup. L'expérience euh, touche donc à sa fin. Euh… Est-ce que tu aimerais qu'on revienne sur certains points euh de l'expérience, des choses que tu voudrais partager en plus euh, une impression euh, tout ce qui te passe par la tête. Histoire de clôturer un peu ce, cet entretien.

Zeta : Alors, bon, par rapport à l'entretien j'ai rien de particulier à ajouter. Euh par rapport aux, à l'objet de l'entretien je pense que, 'fin, je pense que c'est quand même un outil qui a des belles perspectives donc euh… J'ai, j'ai un intérêt réel en tant qu'utilisateur potentiel on va dire euh, à suivre ça. C'est, ouais, c'est, j'aime bien le sujet et puis euh, j'pense que tout l'enjeu c'est de voir jusqu'où on peut aller euh vis-à-vis de la, l'humanisation on va dire, du chatbot. Et… Ouais. Mais c'est un, c'est un domaine ouais je, je, je connais pas, bah, je sais pas trop les possibilités mais je, ça m'intéresserait ouais d'avoir un peu un, comme un, un robot sous la main finalement euh, qu'on peut solliciter en cas de besoin. Notamment pour ce que j'évoquais de-, tout à l'heure, de poser des questions sur des forums et tout ça, c'est bien parce que tout le monde y a accès, mais y a des, parfois des délais de réponse euh, peut-être qu'on réfléchit plus à la formulation des phrases etc. donc ça nous prend plus de temps à rédiger, alors que si on s'a-, s'adresse à un chabot on va peut-être être un peu plus spontané.

Donc euh ouais, je, j'vois ça comme vraiment un, un outil complémentaire qui a des, des perspectives, du potentiel plutôt.

SJ : Ok. Ok super. Effectivement c'est vrai que ce qui ressort un peu de ton entretien c'est le, le, l'aspect euh ouais de l'utilité, de la pertinence et l'aspect ludique ou innovateur. Donc c'est, c'est bien, dans tes réponses c'est bien ce qui, ce qui en ressort.

Zeta : Hmm Hmm.

SJ : Donc euh, donc non. Super. Ben en tout cas euh, merci beaucoup euh de nous avoir aidés euh pour cet entretien.

Zeta : Pas de soucis.

SJ : On va couper l'enregistrement.