HAL
open science

Approche computationnelle du regulatory focus pour des agents interactifs : un pas vers une personnalité artificielle

Caroline Faur

▶ **To cite this version:**

Caroline Faur. Approche computationnelle du regulatory focus pour des agents interactifs : un pas vers une personnalité artificielle. Intelligence artificielle [cs.AI]. Université Paris Saclay (COmUE), 2016. Français. NNT : 2016SACLS352 . tel-01401276

HAL Id: tel-01401276

https://theses.hal.science/tel-01401276v1

Submitted on 23 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NNT : 2016SACLS352

THESE DE DOCTORAT
DE
L'UNIVERSITE PARIS-SACLAY
PREPAREE A
L'UNIVERSITE PARIS-SUD

ECOLE DOCTORALE N° 580
Sciences et technologies de l'information et de la communication (STIC)

Spécialité de doctorat : Informatique

Par

Caroline Faur

Approche computationnelle du regulatory focus pour des agents interactifs : un pas vers une personnalité artificielle

Thèse présentée et soutenue à Orsay, le 21/10/2016 :

Composition du Jury :

M. Marty, Jean-Charles	Maitre de conférences – HDR, LIRIS, Université de Savoie	Rapporteur
M. Fontayne, Paul	Professeur des Universités, CeRSM, Université Paris-Ouest	Rapporteur
Mme Pelachaud, Catherine	Directrice de recherche CNRS, LTCI, Télécom ParisTech	Présidente
M. Pauchet, Alexandre	Maitre de conférences – HDR, LITIS, INSA Rouen	Examinateur
M. Martin, Jean-Claude	Professeur des Universités, LIMSI, Université Paris-Sud	Directeur de thèse
Mme Clavel, Céline	Maitre de conférences, LIMSI, Université Paris-Sud	Co-directrice de thèse

REMERCIEMENTS

> It's no use going back to yesterday, because I was a different person then.
>
> Lewis Carroll, Alice in Wonderland

On m'avait prévenu, il y a un avant et un après thèse. Si j'ai cru au départ qu'il était possible de gérer une thèse comme un "simple" projet, je sais aujourd'hui qu'on ne sort pas indemne de quatre années de doctorat, que ce soit professionnellement ou personnellement. Et je ne serais pas devenue le docteur que je suis aujourd'hui sans les personnes que je voudrais remercier maintenant.

Mes premiers remerciements vont bien évidemment à mes directeurs, Jean-Claude Martin et Céline Clavel. Jean-Claude, merci de m'avoir donné ma chance malgré mes doutes et mon parcours atypique, merci de m'avoir soutenu lorsque j'ai décidé de faire année supplémentaire. Céline, merci d'avoir été là tout au long de cette thèse, toujours disponible pour une discussion par dessus les écrans ; merci de m'avoir laissé le temps de penser quand j'en ai eu envie, merci d'avoir travaillé avec moi quand j'en ai eu besoin ; si la thèse est un marathon, tu as su placer les points de ravitaillement pour que je puisse aller au bout.

Je remercie tous les membres de mon jury, Paul Fontayne, Jean-Charles Marty, Alexandre Pauchet et Catherine Pelachaud. Merci aux rapporteurs, qui ont accepté de faire un pas de côté par rapport à leur domaine pour relire ce travail pluridisciplinaire ; merci aux examinateurs pour leurs remarques et leurs questions lors de la soutenance ; merci à tous pour le temps pris pour évaluer mon travail.

Je remercie les partenaires du projet MoCA : Wafa Johal et Florian Pécune, mes co-doctorants sur ce projet, Matthieu Courgeon, toujours disponible pour m'aider avec MARC malgré la montagne de travail qu'il abat, Dominique Duhaut, Magalie Ochs, Catherine Pelachaud et Sylvie Pesty, pour leur supervision.

Merci également à tous mes participants, ceux qui ont rempli les questionnaires web comme ceux qui ont joué à Can't Stop. Merci au collège Mondétour pour les pré-tests et au collège Alain Fournier pour mon expérimentation finale, en particulier à Mme Brouard pour son implication, M. Titaux et Mme Lafaye pour leur aide et leur disponibilité et à tous les élèves de 4ème de l'année 2015/2016.

Après quatre ans passés au LIMSI, il s'y trouve forcément beaucoup de gens que je voudrais remercier, tellement que je m'excuse d'avance auprès de ceux qui ne trouveraient pas leur nom, sachez que cette liste n'est pas exhaustive !

Merci à tous les ex-doctorants et à tous les futurs docteurs avec qui j'ai traversé ces années : Mehdi, Léonor, Tom, Florian, Yassine, Lydia, Guillaume, Renaud, Lauriane, Marwen, Kevin, Samy, Vincent, Munshi, Driss. Seuls ceux qui l'ont vécu peuvent comprendre :)

Merci à tous les membres de l'équipe CPU et à tous les habitants du bureau 205, en particulier Sylvain pour avoir été ma cible préférée et avoir supporté plus ou moins stoïquement lancers de pommes et blagues idiotes.

Merci à Laurence, Carole, Sophie, Bénédicte, Anne, Stéphanie pour leur disponibilité et leur aide sur l'administratif.

Merci à la team Alain Bonhomme qui m'a toujours laissé les mots composés. Merci à tous ceux avec qui j'ai fait des pauses, bu des cafés ou des bières, raconté des bêtises.

D'un point de vue plus personnel, je voudrais remercier Julie et Caroline. Sans votre travail, le mien n'aurait jamais pu aboutir. Je remercie mes amis, qui m'ont toujours encouragée tout au long de ces quatre ans de plaintes plus ou moins continues avec une pensée particulière pour Camille (comme toujours) et Antoine (malgré tout).

Enfin, je remercie mes parents et plus généralement ma famille, qui m'a soutenue, accompagnée, supportée de toutes les façons et dans tous les moments. Rien de ce que je pourrais écrire ne saurait dire à quel point vous comptez, dans cette thèse et dans ma vie.

RESUMÉ

L'essor de l'informatique affective ouvre aujourd'hui la porte à la création de dispositifs artificiels dotés d'une forme d'intelligence sociale et émotionnelle. L'étude de l'interaction homme-machine dans ce contexte offre de nombreuses pistes de recherche. Parmi celles-ci se trouve la question de la personnalité : comment modéliser certaines caractéristiques de la personnalité pour une entité artificielle ? Comment ces caractéristiques influencent-elles le déroulement de l'interaction entre l'homme et la machine ?

Cette question globale soulève plusieurs questions de recherche : Comment définir la personnalité ? Sur quels modèles et théories issus de la psychologie peut-on s'appuyer pour développer une personnalité artificielle ? Quelle méthodologie adopter pour aborder l'implémentation d'un concept psychologique complexe ? Qu'apporte le développement informatique d'une personnalité à l'interaction homme-machine ? Au domaine de la psychologie de la personnalité ? Comment évaluer expérimentalement ces apports ? Pour aborder ces questions, nos travaux se positionnent dans une optique pluridisciplinaire, au croisement de l'informatique et de la psychologie.

Au regard de sa pertinence pour une approche computationnelle, nous avons modélisé la régulation du Soi comme une composante de la personnalité. Ce concept est approché à partir de la théorie du regulatory focus. Sur cette base théorique, un cadre de travail conceptuel et un modèle computationnel sont proposés. Un questionnaire mesurant le regulatory focus a également été développé et validé.

Ces propositions théoriques sont mises en œuvre dans deux implémentations data-driven (dimensionnelle vs socio-cognitive) dotant des agents de regulatory focus en utilisant de l'apprentissage automatique. Deux études utilisateurs (interaction unique avec agent artificiel vs sessions répétées avec agent animé), présentées dans un cadre ludique, ont permis d'étudier la perception du regulatory focus chez un agent et son impact sur l'interaction.

Nos résultats encouragent l'utilisation du regulatory focus en informatique affective et ouvrent des perspectives sur les liens théoriques et méthodologiques entre informatique et psychologie.

Mots clés : interaction homme-machine ; informatique affective ; personnalité ; regulatory focus ; agents artificiels ; agents animés

ABSTRACT

The development of affective computing is leading to the design of artificial devices endowed with a form of social and emotional intelligence. The study of human-computer interaction in this context offers many research tracks. Among them is the question of personality : how to model some features of personality for a artificial device? How these features will influence the course of interaction with users?

This goal rises several research questions: how to define personality? On which models and theories from psychology should we rely to define some artificial personality? Which methodology will help to address the implementation of such a complex psychological concept? What could artificial personality bring to the field of human-computer interaction? And to the psychology of personality? How to experimentally evaluate these contributions? To address these issues, this thesis takes a multidisciplinary approach, at the crossing of computing science and psychology.

Given its relevance to a computational approach, we modeled self-regulation as a component of personality. This concept is approached from the regulatory focus theory. On this theoretical basis, a conceptual framework and a computational model are proposed.

Our theoretical proposals led to two data-driven implementations (dimensional vs. socio-cognitive) which endowed our artificial agents with regulatory focus by using machine-learning. A French questionnaire measuring regulatory focus was designed and validated. Two user studies (brief interaction with artificial agents vs. repeated sessions with animated agents), where the regulatory focus of agents is conveyed via game strategies, enabled the study of regulatory focus perception and its impact on the interaction.

Our results support the use of regulatory focus in affective computing and open perspectives on the theoretical and methodological links between computer science and psychology.

Keywords: human-computer interaction; affective computing; personality; regulatory focus; artificial agents; animated agents

TABLE DES MATIÈRES

i INTRODUCTION, CONTEXTE SCIENTIFIQUE ET PROBLÉMATIQUE 1

1 INTRODUCTION 3
- 1.1 La personnalité comme objet d'étude pluridisciplinaire 4
- 1.2 Contexte théorique et applicatif 6
- 1.3 Objectifs scientifiques 8
- 1.4 Organisation du manuscrit 11

2 ETAT DE L'ART 13
- 2.1 Implémenter une personnalité artificielle 14
- 2.2 Personnalité et interaction homme-machine 23
- 2.3 Positionnement et problématique 27

ii CONTRIBUTIONS THÉORIQUES ET MÉTHODOLOGIQUES 31

3 MODÉLISER LA RÉGULATION DU SOI 33
- 3.1 Personnalité et régulation du Soi 34
- 3.2 Perseed, ou représenter la régulation du Soi en tant que composante d'une personnalité artificielle 41
- 3.3 Résumé du chapitre 48

4 MESURE DU REGULATORY FOCUS : LE QUESTIONNAIRE RFQ-PF 49
- 4.1 Motivations 50
- 4.2 Développement de l'échelle 52
- 4.3 Première étude de validation 54
- 4.4 Seconde étude de validation 61
- 4.5 Résumé du chapitre 70

iii IMPLÉMENTATIONS ET EXPÉRIMENTATIONS 71

5 IMPLÉMENTATION DIMENSIONNELLE DU REGULATORY FOCUS ET INTERACTION À COURT TERME 73
- 5.1 Objectifs 74
- 5.2 Méthodologie 75
- 5.3 Hypothèses générales 82
- 5.4 Protocole expérimental 83
- 5.5 Résultats 85
- 5.6 Discussion 89
- 5.7 Résumé du chapitre 93

6 IMPLÉMENTATION SOCIO-COGNITIVE DU REGULATORY FOCUS ET INTERACTIONS RÉPÉTÉES 95
- 6.1 Objectifs 96
- 6.2 Méthodologie 99
- 6.3 Hypothèses générales 109
- 6.4 Protocole expérimental 110

6.5 Résultats 117
6.6 Discussion 136
6.7 Résumé du chapitre 142

IV CONCLUSION 143

7 CONCLUSION ET PERSPECTIVES 145
7.1 Résumé du travail de thèse et perspectives à court-terme 146
7.2 Réflexions générales et perspectives à long-terme 150

V ANNEXES 155

A ÉLÉMENTS RELATIFS À LA CONSTRUCTION ET À LA VALIDATION DU QUESTIONNAIRE RFQ-PF 157
A.1 Proverbes de la pré-sélection et représentativité perçue 157
A.2 RFQ-PF ; Version finale ; 18 items 159
A.3 Traduction du RFQ utilisée ; Angel and Steiner (2010) 160

B RÈGLES DU JEU CAN'T STOP 161
B.1 Matériel de jeu 161
B.2 But du jeu 161
B.3 Déroulement de la partie 162

C STATISTIQUES DESCRIPTIVES POUR LA SECONDE EXPÉRIMENTATION AVEC INTERACTIONS RÉPÉTÉES 165
C.1 Perception du changement 165
C.2 Regulatory focus 166
C.3 Crédibilité 168
C.4 Perception de l'interaction 170

D FORMULAIRES DE CONSENTEMENT ET AUTORISATIONS 173
D.1 Première expérimentation (adultes) 173
D.2 Seconde expérimentation (adolescents) 174

E PUBLICATIONS 177

RÉFÉRENCES 179

TABLE DES FIGURES

Figure 1 Robots Pepper (SoftBank Robotics) installés en France dans des lieux ouverts au public 7

Figure 2 Evolution du nombre de papiers par année contenant le mot "personality" (titre, abstract ou mots-clés) dans le domaine "Computer Science" (recherche sur ScienceDirect.com en juin 2016). 15

Figure 3 Activation des unités cognitivo-affectives par des caractéristiques situationnelles pour produire des comportements dans le modèle CAPS. D'après Mischel and Shoda (1995) 21

Figure 4 Sous-systèmes psychologiques à la base de la personnalité. De Mayer (2005) 28

Figure 5 Imbrication des concepts et théories utilisés dans le cadre de notre modélisation 35

Figure 6 Variables psychologiques induisant une stratégie promotion-focus (haut) et variables psychologiques affectées en conséquence (bas) 38

Figure 7 Variables psychologiques induisant une stratégie prévention-focus (haut) et variables psychologiques affectées en conséquence (bas) 39

Figure 8 Exemple d'organisation d'un réseau de self-images et d'attributs (sous forme de paires clé-valeur) pour un agent de type "copain de jeu" 44

Figure 9 Illustration de l'injection d'attributs dans une architecture cognitive via une règle promotion-focus 46

Figure 10 Etudes utilisant des jeux pour évaluer des agents ou des robots affectifs 76

Figure 11 Plateau de jeu Can't Stop réel (à gauche) et virtuel (à droite) 77

Figure 12 Ecrans montrant les adversaires filmés par webcams (en haut) et image du plateau de jeu Can't Stop tel que vu par les participants chacun sur leur écran (en bas) 78

Figure 13 Une branche de l'arbre utilisé pour la décision stop-ou-encore montrant les noeuds de prédiction (rectangles) et les noeuds de décision (ovales) 80

Figure 14 Un utilisateur jouant à Can't Stop avec un agent artificiel durant le tour de l'agent. En bas à gauche, le plateau de jeu tel que vu par l'utilisateur et en haut à droite, l'enregistrement du visage de l'utilisateur par une webcam 84
Figure 15 Score promotion pour les différents agents. *Les valeurs avec la même lettre ne sont pas significativement différentes à $p < 0{,}05$; Barre d'erreur = SD* 87
Figure 16 Score prévention pour les différents agents. *Les valeurs avec la même lettre ne sont pas significativement différentes à $p < 0{,}05$; Barre d'erreur = SD* 88
Figure 17 Interaction entre le focus chronique des participants et la personnalité de l'agent pour le score d'appréciation. *Les valeurs avec la même lettre ne sont pas significativement différentes à $p < 0{,}05$* 89
Figure 18 Agent virtuel MARC utilisé lors de la deuxième étude 97
Figure 19 Comportements non-verbaux de l'agent virtuel durant une phase de jeu (1/2) 106
Figure 19 Comportements non-verbaux de l'agent virtuel durant une phase de jeu (2/2) 107
Figure 20 Architecture de l'application Can't Stop (en vert) en interaction avec la plateforme MARC 108
Figure 21 Conditions (variables indépendantes) et mesures (variables dépendantes) de notre protocole à interactions répétées 111
Figure 22 Participants jouant à Can't Stop contre un agent durant une session expérimentale 114
Figure 23 Graphique et statistiques de l'effet d'interaction entre la présence d'un agent virtuel et le contexte d'induction pour le score promotion 121
Figure 24 Graphique et statistiques de l'effet principal du type d'agent sur l'inclination perçue 122
Figure 25 Graphique et statistiques de l'effet principal du type d'agent sur l'appréciation perçue 123
Figure 26 Graphique et statistiques de l'effet d'interaction entre le type d'agent et la session pour le score de distraction 124
Figure 27 Graphique et statistiques de l'effet de la session sur le score Stimulation pour le Gr-Low 126
Figure 28 Graphique et statistiques de l'effet d'interaction sur le score Stimulation pour le Gr-Low 127

Figure 29 Graphique et statistiques de l'effet de la session sur le score Stimulation pour le Gr-High 128
Figure 30 Graphique et statistiques de l'effet de la session sur le score Aide pour le Gr-High 129
Figure 31 Graphique et statistiques de l'effet de la session pour le score Aide pour le Gr-High 130
Figure 32 Graphique et statistiques de l'effet d'interaction sur le score Intimité pour le Gr-High 131
Figure 33 Graphique et statistiques de l'effet d'interaction sur le score Fiabilité pour le Gr-Mod 132
Figure 34 Graphique et statistiques de l'effet de la session sur le score Fiabilité pour le Gr-Low 133
Figure 35 Graphique et statistiques de l'effet d'interaction sur le score Fiabilité pour le Gr-Low 134
Figure 36 Graphique et statistiques de l'effet de la session sur le score Fiabilité pour le Gr-High 134
Figure 37 Graphique et statistiques de l'effet d'interaction sur le score Fiabilité pour le Gr-High 135
Figure 38 Graphique et statistiques de l'effet de la session sur le score Self validation pour le Gr-Mod 136
Figure 39 Graphique et statistiques de l'effet de la session sur le score Self validation pour le Gr-High 137
Figure 40 Matériel du jeu Can't stop 161
Figure 41 Illustration des 4 étapes composant une phase de jeu 163
Figure 42 Diagramme d'activité représentant une partie de Can't Stop 164

LISTE DES TABLEAUX

Tableau 1 Les 5 dimensions du Five Factor Model et les facettes associées 17
Tableau 2 Attitudes et comportements spécifiques aux focus promotion et prévention utilisés pour la catégorisation des proverbes pré-sélectionnés pour le RFQ-PF 53
Tableau 3 Règles empiriques d'interprétation des indicateurs d'adéquation du modèle produit par l'analyse factorielle exploratoire 56
Tableau 4 Moyennes (M) et écarts-types (SD) par items du RFQ-PF ; Facteurs (Fx) avec saturation des items et pourcentage de variance expliquée 57
Tableau 5 Corrélations de Spearman entre les scores du RFQ-PF et les scores du RFQ et du PANAS 59
Tableau 6 Corrélations de Spearman entre les sous-scores Promotion du RFQ-PF et les scores du RFQ 60
Tableau 7 Règles empiriques d'interprétation des indicateurs d'adéquation des modèles produits par les analyses factorielles confirmatoires 64
Tableau 8 Indicateurs d'adéquation pour les modèles en 2 et 3 facteurs concernant le questionnaire RFQ-PF à 18 items 65
Tableau 9 Corrélations de Spearman entre les scores du RFQ-PF et les scores du BFI-Fr, BIS/BAS, LOT-R, RFSS, RSES et SPSRQ 66
Tableau 10 Statistiques de la validation croisée à 10 échantillons pour trois classifieurs sur les différents modèles de jeu 81
Tableau 11 Statistiques descriptives pour les différents scores collectés *Pro Sc. = score promotion ; Pre Sc. = score prévention ; Lik. = appréciation ; Perc. Int. = intelligence perçue ; CQV = Coefficient de variation des quartiles* 86
Tableau 12 Présentation des choix de modélisation d'une partie des checks du modèle émotionnel durant le tour de l'humain 103
Tableau 13 Présentation des choix de modélisation d'une partie des checks du modèle émotionnel durant le tour de l'agent 104
Tableau 14 Moyennes des 5 scores du questionnaire McGill pour les trois clusters constitués 125

Tableau 15 Proverbes prévention pré-sélectionnés sur leur pertinence pour l'échantillon tout-venant et avis des experts 157
Tableau 16 Proverbes promotion pré-sélectionnés sur leur pertinence pour l'échantillon tout-venant et avis des experts 158
Tableau 17 Fréquences concernant la perception du changement 165
Tableau 18 Statistiques descriptives pour le score promotion 166
Tableau 19 Statistiques descriptives pour le score prévention 167
Tableau 20 Statistiques descriptives pour l'appréciation 168
Tableau 21 Statistiques descriptives pour l'intelligence perçue 169
Tableau 22 Statistiques descriptives pour le Flow - Intérêt 170
Tableau 23 Statistiques descriptives pour le Flow - Distraction 171
Tableau 24 Statistiques descriptives pour la coprésence 172

Première partie

INTRODUCTION, CONTEXTE SCIENTIFIQUE ET PROBLÉMATIQUE

1 INTRODUCTION

Sommaire

1.1 La personnalité comme objet d'étude pluridisciplinaire **4**

1.2 Contexte théorique et applicatif **6**

1.2.1 Articuler des modèles psychologiques et informatiques 6

1.2.2 Le temps des compagnons artificiels 6

1.3 Objectifs scientifiques **8**

1.3.1 Modéliser un processus de la personnalité 8

1.3.2 Proposer une méthodologie pluridisciplinaire 10

1.4 Organisation du manuscrit **11**

Cette thèse s'intéresse à la question de la personnalité et son application dans le domaine informatique, et plus précisément de son apport dans le cadre de l'interaction homme-machine. Nous proposons un cadre conceptuel ainsi qu'une spécification fonctionnelle et une méthodologie pour doter un agent artificiel d'une personnalité. Cette proposition est mise en œuvre dans une interaction ludique avec un agent virtuel. La perception de la personnalité d'un agent virtuel et son impact sur l'interaction ont été évalués par différentes expérimentations. Au travers du concept de personnalité artificielle, nous essaierons également d'illustrer et de discuter les liens théoriques et méthodologiques entre informatique et psychologie dans le cadre d'une approche pluridisciplinaire.

1.1 LA PERSONNALITÉ COMME OBJET D'ÉTUDE PLURIDISCIPLINAIRE

L'application du concept de personnalité en informatique relève du champ de l'informatique affective, en plein essor depuis une trentaine d'années. Le terme d'informatique affective *(affective computing)* dans le sens où nous l'utilisons a été défini par Picard (1995) : ce champ de recherche recouvre l'étude et le développement de systèmes et de dispositifs capables de reconnaitre, d'interpréter, de traiter et de simuler les affects humains (Tao and Tan, 2005). Avoir une machine affective lors d'une interaction homme-machine présente plusieurs intérêts comme améliorer l'interaction et réduire la frustration de l'utilisateur, en facilitant la communication affective (Picard, 1999) ou répondre à des besoins émotionnels naturellement présents chez l'humain (Picard and Klein, 2002).

Lorsqu'il s'agit d'interaction homme-machine, il est possible d'aborder le concept de personnalité aussi bien du côté "homme" que du côté "machine". Dans le premier cas, il s'agit d'utiliser un modèle de la personnalité de l'utilisateur (par détection ou par utilisation d'informations pré-existantes) pour adapter le dispositif technologique avant et/ou au cours de l'interaction, afin d'optimiser le déroulement de celle-ci. Dans le deuxième cas, il s'agit de doter une entité artificielle d'une personnalité propre, qui peut d'ailleurs être dépendante (ou non) de l'utilisateur et du contexte d'interaction. C'est à ce pan de la recherche que se rattachent nos travaux.

Si certains chercheurs en informatique font le choix de définir leur propre modèle de personnalité (Gmytrasiewicz and Lisetti, 2002), la grande majorité des travaux en informatique affective prennent appui sur des modèles psychologiques. C'est pourquoi il nous parait important de présenter ces modèles afin de pouvoir en discuter les intérêts et les limites dans leur application à l'informatique.

La personnalité est un domaine de recherche important en psychologie et qui appelle à étudier l'être humain dans sa globalité. La personnalité peut être vue de façon générale comme un ensemble de comportements individualisés, stables et organisés (Mischel et al., 2004), ou pour aller plus loin, comme la structuration cohérente d'affects, de comportements, de cognitions et de buts à travers le temps et l'espace (Revelle and Scherer, 2009).

Dans le domaine scientifique de la psychologie, nous distinguerons deux courants concernant l'approche de la personnalité et pouvant être utilisés pour expliquer et prédire des comportements humains : les approches dimensionnelles, basées sur les traits de personnalité, et les approches que nous désignerons comme socio-cognitives, intéressées par les structures et les processus de la personnalité.

Les *approches dimensionnelles* répondent à un but de description de la personnalité : elles essayent de distinguer des facteurs décrivant des patterns de comportements, d'affects et de valeurs. Il est alors possible d'évaluer, via des outils psychométriques, le degré d'adhérence d'un individu sur chacun de ces facteurs et de situer l'individu en question au sein d'une population de référence. S'il existe différentes théories proposant des traits de personnalité, le *modèle en cinq facteurs* (the Five Factors Model ou FFM, également appelé modèle du Big Five) de Costa and McCrae (1992a) qui définit, comme son nom l'indique, cinq traits de personnalité, est le plus largement utilisé.

Dans la littérature informatique, les modèles permettant de doter des agents ou des robots de personnalité sont majoritairement basés sur des approches dimensionnelles et plus particulièrement, sur le modèle du Big Five (Vinciarelli and Mohammadi, 2014a). Cependant, cette approche présente certaines limites. En effet, les liens fonctionnels entre un trait de personnalité et les comportements qui lui sont associés ne sont pas ou peu définis par la littérature. Il est alors du ressort de l'informaticien de créer le processus et de l'adapter en fonction du trait à exprimer. De plus, ces approches gomment les différences intra-individuelles qui pourraient pourtant être intéressantes à reproduire dans un contexte d'interaction homme-machine. C'est pourquoi nous proposerons d'orienter nos travaux par une approche socio-cognitive.

Les *approches socio-cognitives* répondent quant à elles à un but de compréhension des structures et des mécanismes psychiques soustendant les comportements donnant lieu à des différences interindividuelles et dont la stabilité justifie l'application du terme de personnalité. Ces approches mettent également un accent particulier sur l'importance du contexte dans l'expression des comportements illustratifs de la personnalité d'un sujet (Mischel and Shoda, 1995).

Bien que ces approches socio-cognitives soient moins utilisées en informatique affective, nous soulignerons leurs intérêts (et leurs limites) pour la modélisation d'une personnalité artificielle.

1.2 CONTEXTE THÉORIQUE ET APPLICATIF

1.2.1 *Articuler des modèles psychologiques et informatiques*

D'un point de vue théorique, les liens entre informatique et psychologie sont multiples. Ces deux sciences s'influencent mutuellement. Concernant le lien allant de la pensée informatique vers la psychologie, nous pouvons citer le courant computationnaliste de la cognition, particulièrement développé par Fodor (1983) et sa vision modulaire de l'esprit. Le test de Turing (Turing, 1950), connu comme essayant de répondre à la question « Une machine peut-elle penser ? », peut être vu comme nous interrogeant autant, sinon plus, sur notre propre humanité (Hayes and Ford, 1995). De même, la psychologie influence la pensée informatique et l'informatique affective en est un exemple flagrant.

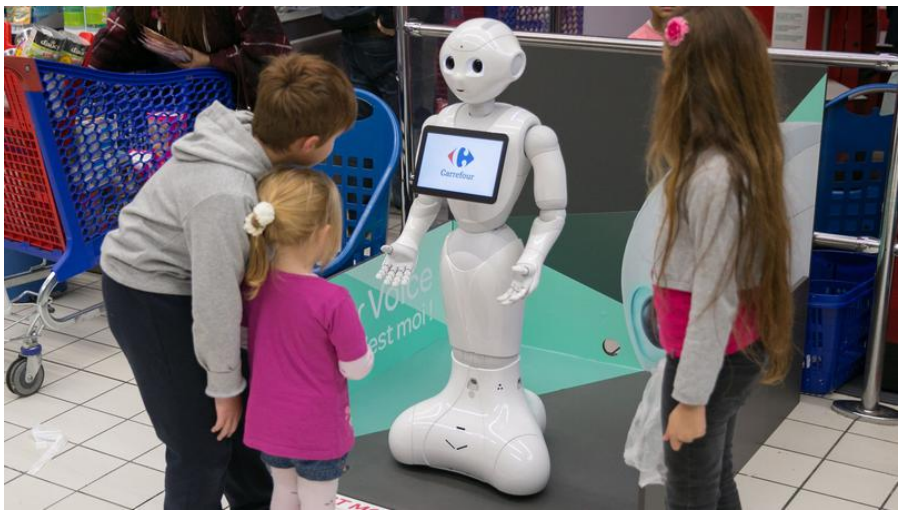
Malgré ces liens, les deux domaines ont souvent été vus comme opposés, dans une vision dualiste, assignant l'informatique au domaine du raisonnement et la psychologie au domaine du sentiment. Depuis, l'importance de l'émotion dans les processus de décision a été mise en avant dans la psychologie humaine et l'informatique affective amène ce concept au sein du domaine informatique. Cependant, l'articulation de l'informatique et de la psychologie reste délicate à mettre en œuvre.

C'est donc avec une démarche résolument pluridisciplinaire que nous essaierons de contribuer par nos travaux aussi bien à la recherche en informatique qu'à la recherche en psychologie.

1.2.2 *Le temps des compagnons artificiels*

D'un point de vue applicatif, l'étude de la personnalité artificielle et de son impact sur l'interaction homme-machine représente un enjeu particulièrement intéressant dans le domaine des compagnons artificiels. Un compagnon artificiel peut être défini comme un dispositif technologique qui fournit un service à son utilisateur sur le long-terme et qui suscite le développement d'une relation avec celui-ci (Benyon and Mival, 2010). Ces dernières années, plusieurs projets de recherche européens se sont d'ailleurs intéressés aux compagnons artificiels, comme les projets Companions (persistent multi-modal interfaces to the Internet ; 2006-2010), SERA (Social Engagement with Robots and Agents ; 2009-2010) et LIREC (LIving with Robots and intEractive Companions ; 2008-2012). Au niveau sociétal, la question

(a) Pepper dans une gare SNCF (image Ouest-France)



(b) Pepper dans un supermarché (image Le Figaro)

FIGURE 1 – Robots Pepper (SoftBank Robotics) installés en France dans des lieux ouverts au public

ne semble d'ailleurs plus être de savoir *si* nous allons cohabiter avec des compagnons artificiels mais *quand*. Le succès d'une application comme Siri ou l'installation de robots Pepper dans des gares SNCF et des supermarchés (cf. Figure 1) pourraient en être les prémisses.

Dans ce contexte, l'étude de la dynamique de la relation entre l'utilisateur et lesdits compagnons est donc une nouvelle problématique scientifique. C'est justement une des questions posées par le projet ANR MoCA [1] (Mon petit monde de Compagnons Artificiels) sur lequel ces travaux ont été financés. Le projet MoCA propose d'étudier

1. Partenaires du projet : LTCI (Laboratoire Traitement et Communication de l'Information), LIG (Laboratoire d'Informatique de Grenoble), Lab-STICC (Laboratoire des Sciences et Techniques de l'Information, de la Communication et de la Connaissance) et LIMSI

un collectif de compagnons, avec divers équipements (robots, agents virtuels) remplissant divers rôles (jeu, aide aux devoirs, sécurité...) et dont les utilisateurs finaux pourraient être des enfants. Ce collectif de compagnons artificiels accompagnerait les enfants dans des activités quotidiennes et serait totalement intégré dans leur cadre de vie.

Dans un tel dispositif, nous faisons l'hypothèse que la personnalité pourrait permettre de créer un "fil rouge", une continuité permettant d'une part d'offrir une identification des différents compagnons malgré la diversité des formes et des rôles, et d'autre part, de permettre à l'utilisateur de comprendre, prédire et justifier les comportements adoptés par les compagnons. Ce dernier point pourrait permettre d'augmenter la crédibilité d'un compagnon artificiel aux yeux de l'utilisateur, c'est-à-dire la capacité du compagnon à être perçu comme vraisemblable et convaincant (Burgoon et al., 2000). L'augmentation de la crédibilité est importante car, au-delà de la forme et de la fonction, la notion centrale dans la définition de ce qu'est un compagnon artificiel est bien sa potentielle capacité à établir une relation à long-terme avec l'utilisateur (Sviatlana et al., 2012).

La littérature en psychologie montre qu'il existe des formes de compatibilité en termes de personnalité dans l'établissement de différentes relations entre humains, par exemple pour les coachs ou les enseignants (Sue-Chan et al., 2012; Karwowski, 2011). Si notre hypothèse se vérifie et que la personnalité d'une entité artificielle a bien un impact sur l'interaction homme-machine, il devient alors crucial de maîtriser la personnalité transmise par le compagnon artificiel. En effet, si celle-ci est adaptée, l'interaction s'en trouverait alors facilitée, offrant alors un contexte propice au développement d'une relation entre l'humain et les dispositifs technologiques qui l'accompagnent.

1.3 OBJECTIFS SCIENTIFIQUES

Les travaux présentés dans ce manuscrit ont été dirigés par deux objectifs scientifiques. Notre premier objectif est thématique et concerne la personnalité, abordée depuis les versants informatique et psychologique. Notre deuxième objectif scientifique est plus transversal et questionne l'aspect pluridisciplinaire de notre projet au travers de la démarche intellectuelle et de la méthodologie que nous mettons en place pour articuler informatique et psychologie.

1.3.1 *Modéliser un processus de la personnalité*

Les recherches concernant la personnalité recouvrent un large champ d'étude car le concept de personnalité peut être abordé à différents niveaux. D'après Revelle (1995), trois niveaux peuvent être distingués :

- « All people are the same » (tous les gens se ressemblent) : du point de vue de l'espèce, certains comportements sont typiques de l'être humain
- « Some people are the same » (certaines personnes sont les mêmes) : on se réfère ici à l'étude des différences et des similarités inter-individuelles
- « No person is the same » (aucune personne n'est la même) : le type de comportement d'une personne lui est unique

Nous ne nous intéresserons pas ici au premier niveau (celui de l'espèce). Bien que les robots et agents artificiels ne fassent pas partie de l'espèce humaine *per se*, nous pouvons observer que ces dispositifs, par leur autonomie grandissante et leur apparence parfois humanoïde, suscitent assez naturellement chez l'utilisateur une tendance à l'anthropomorphisme. L'anthropomorphisme se définit comme la tendance à attribuer des caractéristiques humaines (comme des buts, des motivations, des intentions, des croyances, etc) à des choses non-humaines (Epley et al., 2007). Nous pouvons voir l'anthropomorphisme comme un mécanisme d'adaptation permettant de gérer ces nouvelles technologies (Waytz et al., 2010). Pour un utilisateur, appliquer ses connaissances sur les humains pour interagir avec une machine complexe permet de donner un cadre explicatif et prédictif aux comportements de ladite machine.

Les deux autres niveaux sont particulièrement intéressants lorsqu'il s'agit de modéliser une personnalité artificielle. Dans un premier temps, essayer de comprendre pourquoi "certaines personnes sont les mêmes" nous orientera dans le choix des structures et des processus constitutifs de notre modèle. Dans un deuxième temps, se demander pourquoi "aucune personne n'est la même" peut aider à créer des dispositifs plus proches d'un individu que d'un simple gadget technologique.

En gardant ces vues à l'esprit, notre premier objectif est de proposer une modélisation computationnelle permettant de doter une entité artificielle d'une personnalité. Pour cela, nous commencerons par étudier les modèles psychologiques de la personnalité. En se basant sur cette littérature issue de la psychologie, nous proposerons un modèle ainsi qu'une méthodologie permettant d'implémenter un processus aboutissant à des comportements distinctifs d'une personnalité. Il conviendra alors de tester notre implémentation, en termes de perception par un utilisateur aussi bien qu'en termes d'apport de cette personnalité pour l'interaction.

1.3.2 *Proposer une méthodologie pluridisciplinaire*

Comme décrit dans les paragraphes précédents, nous pensons que le domaine de la psychologie peut apporter des concepts et des modes de réflexion permettant d'enrichir les modèles informatiques et par là, l'interaction homme-machine. Cependant, les modèles psychologiques utilisés répondent à des objectifs de recherche différents de ceux des chercheurs en informatique affective. Ces modèles visent à comprendre l'humain et à proposer des concepts permettant d'organiser une pensée expliquant leur fonctionnement. Ce sont donc des modèles généraux proposant une certaine vue d'un objet de recherche et non des spécifications fonctionnelles destinées à l'implémentation d'humain artificiel.

D'autre part, l'informatique affective peut également enrichir le domaine de la psychologie de différentes manières. Se rapprocher de l'humain dans la modélisation n'est pas forcément un but souhaité (ou souhaitable ?) dans le cadre d'une implémentation destinée à l'interaction homme-machine. Mais en s'appuyant sur des modèles psychologiques et en essayant de reproduire le comportement humain, les modèles informatiques et leur utilisation peuvent apporter des informations utiles à la psychologie. En effet, l'informatique permet d'observer des variables psychologiques "internes" comme la manipulation de croyances, l'utilisation des connaissances ou les processus de prise de décision qui sont par nature inaccessibles autrement que par l'introspection.

Néanmoins, il est nécessaire de garder à l'esprit que l'expérimentation informatique n'est pas une "validation" en soi d'un modèle psychologique. Si l'implémentation d'un processus n'amène pas aux comportements prédits par la théorie, cela permet de pointer d'éventuels problèmes au sein de ladite théorie (ou dans l'interprétation qui en a été faite par le concepteur du programme informatique). Mais en cas de succès, il faut être conscient qu'il peut exister des modèles concurrents en psychologie qui pourraient également amener aux mêmes comportements observables. De même, le succès d'une étude perceptive (e. g. un agent virtuel est perçu comme possédant la caractéristique implémentée) n'est pas une "validation" en soi d'un choix d'implémentation. En effet, une implémentation différente (potentiellement même dénuée de base théorique) pourrait peut-être entraîner la même perception chez l'utilisateur. L'intérêt de l'expérimentation ici ne réside donc pas dans la "validation" en tant que telle, mais dans la réflexion bidirectionnelle qu'elle nourrit (aussi bien par rapport au concept implémenté que par rapport aux choix computationnels qui ont été faits).

Notre deuxième objectif est donc de proposer une méthodologie permettant d'articuler modèle psychologique et modèle informatique

tout en interrogeant les concepts utilisés. De fait, il est intéressant d'observer les transformations que nous apportons à ces concepts en les faisant passer de la psychologie vers l'informatique, dans un sens (modélisation computationnelle) comme dans l'autre (interprétation des résultats d'expérimentations). Dans notre travail, le concept de personnalité sera notre objet d'étude, du point de vue de son apport pour l'interaction homme-machine, mais également dans la réflexion à laquelle notre modélisation pourra mener du point de vue de la psychologie.

1.4 ORGANISATION DU MANUSCRIT

Dans ce manuscrit, nous présenterons au *chapitre* 2 un état de l'art sur l'utilisation du concept de personnalité en informatique et nous présenterons les modèles psychologiques sur lesquels s'appuient ces travaux. À la lumière des travaux actuels et de leurs limites, nous exposerons notre problématique de recherche (chapitre 2 - section 2.3). En réponse à cette problématique, nous développerons le travail de recherche effectué en deux parties.

La première partie se concentrera plus particulièrement sur les aspects théoriques et méthodologiques de notre travail.

Au *chapitre* 3, nous proposons de nous intéresser à un processus particulier de la personnalité, la régulation du Soi et ce au travers de la théorie du regulatory focus (chapitre 3 - section 3.1). Sur cette base théorique, nous décrirons notre première contribution : Perseed, à la fois cadre de travail conceptuel et modèle computationnel pour la représentation de la régulation du Soi comme composante d'une personnalité artificielle (chapitre 3 - section 3.2). Nous mettrons également en avant la façon dont cette première modélisation modifie et affine notre problématique scientifique (chapitre 3 - sous-section 3.2.3).

Dans le *chapitre 4*, nous introduisons le questionnaire RFQ-PF (Regulatory Focus Questionnaire - Proverbs Form) qui constitue notre deuxième contribution. Ce questionnaire, que nous avons conçu et validé au regard des normes psychométriques en vigueur, est destiné à mesurer le regulatory focus aussi bien en auto-attribution chez l'humain qu'en hétéro-attribution chez un agent virtuel.

La deuxième partie de cette thèse présentera nos contributions en termes d'implémentation du regulatory focus chez des agents artificiels et d'expérimentations en interaction homme-machine.

Le *chapitre 5* propose une première implémentation dimensionnelle du regulatory focus pour un agent artificiel jouant à un jeu de plateau. Cette implémentation est basée sur une méthodologie data-driven (chapitre 5 - section 5.2). Les agents ainsi produits ont été mis

en interaction avec des participants adultes afin d'évaluer la personnalité perçue des agents et l'apport de l'ajout du regulatory focus pour la crédibilité de l'agent et de l'interaction (chapitre 5 - section 5.4 et section 5.5).

À partir des limites et des enseignements de cette expérimentation, le *chapitre 6* présente une seconde implémentation socio-cognitive du regulatory focus dans le même cadre applicatif. L'implémentation des stratégies de jeu reste basée sur une méthodologie data-driven, mais le dispositif est enrichi par l'ajout d'un agent virtuel capable de réaliser des expressions faciales. Ces expressions sont calculées par un modèle émotionnel dont l'implémentation theory-driven se nourrit également des apports du data-driven (chapitre 6 - section 6.2).

Cette seconde implémentation fait également l'objet d'une étude en interaction homme-machine. Des collégiens ont interagi avec nos agents durant plusieurs sessions. Nous nous intéressons toujours à la personnalité perçue et à la crédibilité des agents et de l'interaction, mais cette expérimentation est également l'occasion d'explorer la question de l'évolution de la relation entre un utilisateur et un agent virtuel au cours de sessions répétées (chapitre 6 - section 6.4 et section 6.5).

Enfin, le *chapitre 7* conclut ce manuscrit en deux parties. La section 7.1 résume les contributions principales du travail présenté ici et offre des perspectives à court-terme pour poursuivre ces travaux de recherche. La section 7.2 propose des réflexions plus générales sur nos objectifs scientifiques et suggère des pistes de développement à long-terme en considération des questions soulevées par nos travaux.

2

ETAT DE L'ART

Sommaire

2.1 Implémenter une personnalité artificielle **14**
- 2.1.1 Les approches dimensionnelles 15
- 2.1.2 Les approches sociocognitives 20

2.2 Personnalité et interaction homme-machine **23**
- 2.2.1 Percevoir la personnalité 24
- 2.2.2 Influencer l'interaction 25
- 2.2.3 Limites des travaux en interaction 26

2.3 Positionnement et problématique **27**

Dans ce chapitre, nous présenterons les travaux existants concernant l'ajout de personnalité dans des dispositifs artificiels (section 2.1). Ces travaux informatiques seront mis en regard avec les modèles psychologiques utilisés. Comme présenté au chapitre 1, nous distinguerons les travaux présentés suivant l'approche de la personnalité utilisée : d'une part, les travaux s'appuyant sur des approches dimensionnelles (sous-section 2.1.1) et d'autre part, ceux utilisant des approches socio-cognitives (sous-section 2.1.2). Nous présenterons ensuite les travaux s'intéressant à l'impact de la personnalité sur l'interaction homme-machine (section 2.2). À la lumière des travaux existants et de leurs limites, nous préciserons le positionnement de nos travaux au sein de notre champ de recherche (section 2.3).

2.1 IMPLÉMENTER UNE PERSONNALITÉ ARTIFICIELLE

Parmi les objets d'étude de l'informatique affective, si l'émotion est sans doute le concept le plus étudié, la personnalité est peut-être le plus général. Dans sa taxonomie des affects, Lisetti (2002) positionne la personnalité comme le nœud le plus haut dans la hiérarchie. Classiquement, l'émotion est considérée comme un affect à court-terme en réaction à une situation particulière alors que la personnalité est une disposition à long-terme qui n'est pas dépendante de la course immédiate des évènements (Moffat, 1997; Scherer, 2000). Revelle and Scherer (2009) utilisent l'analogie suivante pour concevoir les liens entre ces deux concepts : la personnalité est à l'émotion ce que le climat est à la météo [1]. Au-delà de l'émotion, la personnalité est un système global qui implique de nombreux sous-systèmes comme l'attention, la mémoire, la motivation ou la prise de décision (Mayer, 2005).

Comme le montrent Vinciarelli and Mohammadi (2014b) dans leur revue de littérature, la communauté informatique a montré un intérêt croissant pour la question de la personnalité au cours des années 2000. Une recherche sur ScienceDirect sur le mot "personality" dans le domaine "Computer Science" montre effectivement une évolution nette du nombre d'articles sur ce thème au cours des 15 dernières années (cf. Figure 2).

Parmi les nombreuses théories existant en psychologie de la personnalité, les approches dimensionnelles sont majoritairement utilisées en informatique, que ce soit pour la génération de comportements artificiels ou pour la détection de la personnalité de l'utilisateur. Et parmi ces approches dimensionnelles, celle dite du "Big Five" domine nettement : 76 des 81 articles présentés dans la revue de Vinciarelli and Mohammadi (2014b) utilisent cette approche.

1. « A helpful analogy is to consider that personality is to emotion as climate is to weather. » (Revelle and Scherer, 2009, p. 304)

FIGURE 2 – Evolution du nombre de papiers par année contenant le mot "personality" (titre, abstract ou mots-clés) dans le domaine "Computer Science" (recherche sur ScienceDirect.com en juin 2016).

2.1.1 *Les approches dimensionnelles*

2.1.1.1 *En psychologie*

Les approches dimensionnelles de la personnalité définissent des traits de personnalité. Les traits de personnalité sont des dimensions qui permettent de décrire une personne. En psychologie, il existe différentes taxonomies concernant les traits de personnalité, qui se distinguent par leur base théorique et le nombre nécessaire et suffisant de dimensions permettant de décrire la personnalité.

Comme dit plus haut, le modèle de personnalité majoritairement utilisé dans le domaine des différences individuelles (en psychologie comme en informatique) est le Five Factor Model (FFM) (McCrae and Costa, 1987). Ce modèle est également appelé modèle OCEAN ou modèle des "Big Five". Les cinq dimensions définies sont :

- l'Ouverture à l'expérience (O) : décrit la tendance à la curiosité envers soi-même et le monde extérieur,
- le caractère Consciencieux (C) : décrit le degré d'organisation et de contrôle de ses impulsions,
- l'Extraversion (E) : décrit la tendance à l'activité, à la sociabilité et à l'optimisme,
- l'Agréabilité (A) : décrit le degré de sympathie et de bonne disposition envers autrui
- le Névrosisme (N) : décrit la tendance à ressentir des émotions négatives et être instable émotionnellement.

Ce modèle est basé sur l'hypothèse lexicale. L'hypothèse lexicale peut être décrite par deux postulats : 1) les différences de personnalité saillantes et pertinentes sont traduites dans le langage naturel ; 2) les caractéristiques les plus importantes sont plus susceptibles d'être exprimées par un seul mot (Allport, 1937; Goldberg, 1982). Par conséquent, une analyse factorielle appliquée au vocabulaire ayant trait à

la personnalité doit pouvoir dégager des traits importants de la personnalité. Allport et Goldberg ont chacun proposé une taxonomie du vocabulaire anglais se rapportant à la personnalité. Avec la taxonomie d'Allport, Cattell (1943) crée un modèle de la personnalité à seize facteurs. En se basant sur une analyse des travaux de Cattell et de la taxonomie de Goldberg, McCrae and Costa (1987) créent le FFM et définissent donc 5 traits de personnalité. Chaque trait correspond à un "domaine" de la personnalité découpé en six facettes, qui définissent les différents aspects que recouvre un trait (cf. Tableau 1). Chacune de ces facettes correspond à une sous-échelle dans le questionnaire NEO-PI-R, proposé par Costa and McCrae (1992b) pour mesurer la personnalité suivant le FFM. Les six facettes choisies ont été définies en partie de manière empirique (Costa and McCrae, 1995).

Tableau 1 – Les 5 dimensions du Five Factor Model et les facettes associées

Trait	**Facette**	**Description**
Ouverture à l'expérience	Ouverture aux rêveries	*Imagination vive et active, création d'un monde intérieur riche*
	Ouverture à l'esthétique	*Appréciation de l'art et de la beauté*
	Ouverture aux sentiments	*Sensibilité à ses propres émotions*
	Ouverture aux actions	*Intérêt pour les activités différentes et nouvelles*
	Ouverture aux idées	*Curiosité intellectuelle*
	Ouverture aux valeurs	*Capacité à remettre en question les conventions ainsi que ses propres valeurs*
Caractère consciencieux	Compétence	*Croyance en sa propre efficacité*
	Ordre	*Organisation personnelle et soin*
	Sens du devoir	*Importance accordée aux principes éthiques*
	Recherche de réussite	*Besoin d'accomplissement dans le travail*
	Autodiscipline	*Capacité à entreprendre et à terminer des tâches malgré les distractions*
	Délibération	*Tendance à planifier ses comportements*
Extraversion	Chaleur	*Attitude amicale envers les autres*
	Grégarité	*Préférence pour la compagnie des autres*
	Assertivité	*Tendance à se montrer dominant et énergique*
	Activité	*Rythme rapide et vigueur*
	Recherche de sensations	*Fort besoin de stimulation*
	Emotions positives	*Tendance à éprouver beaucoup d'émotions positives*
Agréabilité	Confiance	*Croyance en l'honnêteté et aux bonnes intentions d'autrui*
	Droiture	*Sincérité et franchise*
	Altruisme	*Préoccupation active du bien-être d'autrui*
	Compliance	*Tendance à se soumettre lors de conflits interpersonnels*
	Modestie	*Tendance à l'humilité et à l'effacement*
	Sensibilité	*Préoccupation et sympathie pour autrui*
Névrosisme	Anxiété	*Tendance à être timide, craintif ou inquiet*
	Colère-Hostilité	*Tendance à ressentir de la colère, de la frustration et l'amertume*
	Dépression	*Tendance à se sentir coupable, triste, impuissant ou seul*
	Timidité sociale	*Tendance à éprouver de la honte et de l'embarras*
	Impulsivité	*Incapacité à maîtriser ses désirs et ses émotions*
	Vulnérabilité	*Vulnérabilité au stress*

2.1.1.2 *Utilisation en informatique*

Les Big Five sont très souvent utilisés pour modéliser informatiquement des personnalités. Les traits sont alors utilisés comme des variables numériques qui viennent pondérer différents mécanismes internes des agents virtuels concernés.

Par exemple, Egges et al. (2004) s'intéressent à la personnalité pour un agent conversationnel animé dans un scénario minimaliste : l'agent, Julie, porte un carton qu'elle doit amener dans un appartement en montant les escaliers. Un homme passe et lui propose de l'aide. Comment Julie va-t-elle réagir face à cet événement en fonction de ses émotions et de sa personnalité ? Ces chercheurs utilisent un modèle OCC (Ortony et al., 1990) pour modéliser les émotions. Le modèle OCC est un modèle de type appraisal qui se base sur les buts, les valeurs et les attitudes pour déterminer l'émotion ressentie à un moment t. Un vecteur à cinq dimensions correspondant aux Big Five est utilisé d'une part, comme base pour sélectionner les buts, les valeurs et les attitudes initiales de l'agent et d'autre part, comme influençant l'intensité des émotions ressenties par l'agent.

Dans son modèle affectif ALMA, Gebhard (2005) définit l'humeur par défaut de personnages virtuels en se basant sur des valeurs numériques attribuées aux cinq dimensions. Les travaux de Mehrabian (1996), qui relient les "Big Five" au modèle émotionnel Pleasure-Arousal-Dominance (PAD) sont utilisés pour permettre le calcul de cette humeur "de base". L'humeur de base d'un personnage non joueur (PNJ) est également définie par Signoretti et al. (2010) sur la base des dimensions de la personnalité. De plus, la personnalité conditionne le déclin des états émotionnels, ainsi que des préférences dans le processus de raisonnement.

Toujours dans le contexte d'un PNJ, les traits de personnalité sont intégrés par Lim et al. (2012) à l'architecture d'un agent au sein d'un jeu de rôle éducatif. Les traits de personnalité définissent ainsi l'importance de différentes motivations possibles et servent également à déterminer des facteurs de déclin pour ces mêmes motivations, au sein d'un modèle combinant l'architecture cognitive FatiMA et le modèle orienté-besoin PSI.

Il est courant d'utiliser un sous-ensemble de traits particulièrement pertinents par rapport à la problématique du modèle. Extraversion, agréabilité et névrosisme forment un sous-ensemble populaire car ces traits sont particulièrement liés aux émotions et aux comportements sociaux.

Ce sous-ensemble de traits est utilisé par André et al. (2000) pour modéliser la personnalité d'agents virtuels servant d'interfaces homme-

machine affectives dans différentes applications (Puppet, the Inhabited Marketplace, Presence). Les traits déterminent l'importance des buts de haut niveau lors de l'évaluation émotionnelle et, dans les applications pertinentes, les traits influent également sur les buts conversationnels et le choix des actes de langage utilisés par des agents conversationnels animés. Kshirsagar and Magnenat-Thalmann (2002) s'intéressent à l'influence de ce sous-ensemble sur l'humeur, qui influence à son tour les émotions exprimées par un agent virtuel.

Van den Bosch et al. (2012) choisissent l'extraversion et l'agréabilité pour façonner le style et le contenu d'une conversation tenue par un PNJ qui joue le rôle du client dans un serious game destiné à des agents immobiliers. Par exemple, un agent avec un haut score d'extraversion aura une plus grande probabilité d'exprimer des souhaits et des opinions concernant sa recherche immobilière, alors qu'un agent avec un bas score d'extraversion aura une plus grande probabilité de demander des informations factuelles. Les traits de personnalité peuvent donc être utilisés pour influencer des processus de haut niveau mais également pour diriger le comportement non-verbal d'un agent virtuel. Neff et al. (2011) se focalisent sur le trait de névrosisme pour étudier son influence sur des gestes non-signifiants pendant le discours (comme se gratter le cou ou se frotter le front).

2.1.1.3 *Limites des approches dimensionnelles*

Les travaux présentés montrent que les traits de personnalité sont utilisés pour pondérer les processus dirigeant le comportement des agents virtuels implémentés. Ces processus sont de différents types : émotionnel, prise de décision, comportements verbaux et non-verbaux. S'il est vrai que l'ensemble de ces processus est impacté par la personnalité, utiliser une approche dimensionnelle comme FFM pose certains problèmes. Les approches de type trait ont été conçues pour donner une description de la personnalité, pour fournir des catégories. Le modèle FFM se concentre donc sur l'expression des comportements associés à la personnalité (quel comportement est caractéristique d'une personnalité ?) plutôt que sur les causes (pourquoi ce comportement est-il caractéristique de cette personnalité ?) ou sur la manière (comment cette personnalité entraîne-t-elle ce comportement ?). Par conséquent, les informaticiens doivent choisir par eux-mêmes comment relier traits et comportements pour reproduire les corrélations comportementales observées en psychologie.

Cette limite a d'ailleurs été soulignée par Bevacqua et al. (2012) qui choisissent d'utiliser plutôt le modèle en trois traits (Extraversion, Névrosisme et Psychoticisme) d'Eysenck (1970). Les dimensions d'extraversion et de névrosisme sont utilisées pour moduler la fréquence et le type des backchannels utilisés par des agents virtuels qui simulent des comportements d'écoute. D'après Bevacqua et al. (2010),

ce modèle a été choisi car il donne des bases psychobiologiques pour expliquer le lien entre traits et comportements. Par exemple, concernant l'extraversion, Eysenck suggère que les personnes extraverties auraient un niveau d'activation corticale naturellement plus bas. En conséquence de quoi, ces personnes supporteraient, voire rechercheraient, des situations présentant des conditions d'activation plus importantes que les personnes introverties. Ces explications causales permettent à la fois d'orienter les choix de modélisation et de justifier les patterns comportementaux des agents ainsi produits.

Une autre limite, commune à FFM et au modèle d'Eysenck, se trouve dans le fait que les modèles de trait masquent des différences intra-individuelles. Par exemple, certaines personnes présentent des comportements extravertis avec leurs amis, mais montrent des caractéristiques d'introverti dans leur environnement professionnel. Les traits ne nous donnent pas ou peu d'informations quant à ces variations comportementales car ils cherchent à décrire le comportement "moyen" des individus et ne prennent pas en compte les caractéristiques situationnelles. Si cet aspect n'est pas forcément problématique quand le but est de décrire globalement un individu, ne pas prendre les caractéristiques situationnelles dans le cadre d'une interaction homme-machine peut être gênant. En effet, un dispositif artificiel destiné à interagir avec un utilisateur doit montrer une certaine forme d'adaptabilité par rapport à la situation. Ajouter une personnalité artificielle doit apporter une cohérence globale mais ne doit pas rigidifier le comportement de la machine.

2.1.2 *Les approches sociocognitives*

2.1.2.1 *En psychologie*

L'approche sociocognitive de la personnalité souligne l'importance de la situation lors de l'expression de comportements liés à la personnalité (Bandura, 1986) et tente de comprendre les processus cognitifs et sociaux qui conduisent à la personnalité. Pour ce faire, l'accent est donc mis sur l'interaction entre la personne et le contexte social (le paradigme "personne $\times$ situation") et sur les différences intra-individuelles. Sachant que les gens montrent une variabilité dans leurs comportements en fonction des situations, la cohérence perçue de la personnalité est attribuée à la stabilité des signatures comportementales (Mischel et al., 2004). Comprendre ces signatures, c'est chercher à éliciter les mécanismes qui sous-tendent les comportements liés à la personnalité.

Mischel, contributeur majeur de cette approche, a conçu avec Shoda le modèle CAPS (Cognitive Affective Processing System) (Mischel and Shoda, 1995). CAPS est un cadre métathéorique de la personnalité. Il définit le système de la personnalité comme étant caractérisé

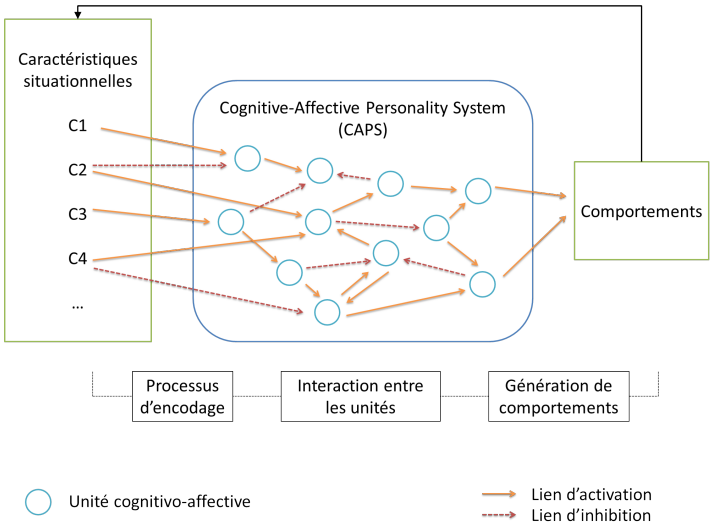


FIGURE 3 – Activation des unités cognitivo-affectives par des caractéristiques situationnelles pour produire des comportements dans le modèle CAPS. D'après Mischel and Shoda (1995)

par un réseau stable d'unités cognitivo-affectives, reliées entre elles par des liens d'activation et d'inhibition. Les comportements sont le résultat de la propagation de l'activation engendrée par les caractéristiques situationnelles au sein de ce réseau (cf. Figure 3).

Cinq types d'unités cognitivo-affectives sont définis : stratégies d'encodage de l'information, attentes/croyances, affects, buts/valeurs et compétences/stratégies de self-régulation. Ces catégories sont larges. En effet, les auteurs ont volontairement laissé les unités non spécifiées.

KAPA (Knowledge-and-Appraisal Personality Architecture) (Cervone, 2006) est un exemple de spécification de CAPS. KAPA définit six types de variable sociocognitive. Il distingue les variables de structure ("knowledge", les connaissances) des variables de processus ("appraisal", l'évaluation). Cette distinction connaissances/évaluation est croisée avec une distinction concernant la direction d'adéquation, qui sépare les croyances, les normes d'évaluation et les objectifs/buts.

Dans le même esprit, Morf (2006) utilise les concepts de CAPS pour construire sa proposition de représenter la personnalité comme un pattern unique de mécanismes de régulation du Soi qui sont utilisés dans le but de construire et de maintenir une certaine image de soi. Dans ce cadre de travail, la personnalité est une représentation dynamique, qui émerge des interactions de différents concepts qui organisent un système représentant le Soi. Ce système est constitué de connaissances sur le Soi, de mécanismes régissant les relations intra- et interpersonnelles et de représentations de certaines relations sociales signifiantes. Ce système du Soi est lui-même inscrit dans un

contexte social. Les situations sociales rencontrées entrent en interaction avec le système du Soi et toute à la fois, modèlent ce système et font émerger les comportements qui seront considérés comme représentatifs de l'individu.

2.1.2.2 *Utilisation en informatique*

Les approches sociocognitives sont assez peu utilisées en informatique, mais certains travaux s'y rattachent néanmoins. Même si tous ne revendiquent pas forcément l'appelation "socio-cognitive", ces travaux s'intéressent tous à des mécanismes particuliers qui, en interaction avec la situation, déterminent des comportements spécifiques pouvant être perçus comme distinctif d'une personnalité.

En se basant sur les théories sociocognitives, Moffat (1997) a conçu Will, un agent artificiel autonome capable de jouer au dilemne du prisonnier itératif. Will possède une architecture modulaire qui comprend entre autres un module de mémoire et un module émotionnel mêlant appraisal et tendances à l'action. La personnalité est mise en place par différents paramètres au sein de l'architecture comme les éléments mémorisés considérés comme pertinents dans une situation, la façon dont l'évaluation émotionnelle est calculée ou les tendances à l'action disponibles. Ces paramètres sont définis de façon empirique par rapport à la théorie.

Sandercock et al. (2006) propose un framework cognitivo-affectif dans lequel la personnalité se développe en fonction de l'environnement. Les agents déterminent leurs comportements à l'aide d'une architecture Belief-Desire-Intention (BDI). Au sein de cette architecture, la personnalité intervient au niveau des préférences relatives aux stratégies de coping (quels plans vont être utilisés pour la gestion des émotions) et de différentes pondérations au sein des processus d'évaluation (l'importance des différentes variables permettant d'évaluer si l'exécution d'un plan est un succès ou non).

Certains combinent également les deux approches comme dans le modèle BASIC (Romano et al., 2005). Ce modèle combine les traits du Big Five avec des facteurs sociocognitifs pour influencer un modèle émotionnel basé sur OCC. Tous deux sont utilisés en tant que variables de pondération dans des équations permettant de déterminer l'état émotionnel des agents. Poznanski and Thagard (2005) combinent également Big Five et sociocognitif pour guider le comportement d'agents virtuels autonomes. Leur modèle SPOT (Simulating Personality Over Time) est un réseau de neurones, mais il combine des nœuds représentant les « Big Five » avec des nœuds qui représentent des facteurs situationnels.

Read et al. (2010) utilisent également un réseau de neurones avec deux systèmes de buts : un système d'évitement et un système d'approche. Ils opérationnalisent ainsi le modèle BIS/BAS (Behavioral Inhibition System/Behavioral Approach System) de Gray (1990). Cette approche de la personnalité utilise le travail d'Eysenck (1970) et postule deux dimensions sur des bases également psychobiologiques. Le modèle BIS/BAS s'intéresse aux motivations (biologiques et psychologiques) qui poussent un individu à adopter des comportements d'approche ou d'évitement en se basant principalement sur le principe de renforcement. Le système BIS est sensible au renforcement négatif (punitions) alors que le système BAS est sensible au renforcement positif (récompenses). Dans la modélisation informatique, les deux systèmes de buts (approche/évitement) sont liés à plusieurs couches dont une couche représentant des caractéristiques situationnelles. Après apprentissage, les simulations montrent que le réseau est capable de produire des signatures comportementales stables (i. e. des associations durables entre des situations et des comportements).

2.1.2.3 *Limites des approches socio-cognitives*

Si les approches dimensionnelles s'adaptent bien à une implémentation informatique en se transformant en variables numériques, les approches sociocognitives sont plus difficiles à utiliser. En effet, nous passons des dimensions chiffrables à de nombreuses structures et de processus en interaction. Si nous voyons ici apparaître l'utilisation de techniques d'apprentissage automatique comme les réseaux de neurones, ceux-ci présentent tout de même un désavantage majeur : contrairement aux implémentations symboliques, ils représentent une boîte noire et manquent donc de pouvoir explicatif quand il s'agit de comprendre les processus liés à la personnalité (Rothmund et al., 2012).

De plus, la plupart des modèles présentés ci-dessus sont utilisés dans des études perceptives (c'est-à-dire une situation où un utilisateur observe et juge une interaction entre un humain et un agent (ou parfois une interaction entre agents). Or les études perceptives en troisième personne n'interrogent pas la même chose que les études en deuxième personne, où l'utilisateur est partie prenante de l'interaction avec le dispositif technologique étudié (De Bruin et al., 2012).

2.2 PERSONNALITÉ ET INTERACTION HOMME-MACHINE

Comme nous l'avons dit au chapitre 1, les progrès constants dans le domaine de la robotique et des agents artificiels ont fait de ces technologies des acteurs de notre vie quotidienne. Des aspirateurs autonomes aux assistants personnels, nous cohabitons avec des technologies qui font désormais partie de notre environnement.

Nass et al. (1994) définissent le paradigme "Computers Are Social Actors" (CASA). Le paradigme CASA définit les ordinateurs comme des acteurs sociaux : bien qu'ils soient conscients qu'un ordinateur n'est pas un être humain, les utilisateurs tendent à adopter des attitudes sociales avec des machines. Il est donc logique de penser qu'un robot ou un agent artificiel élicitent des heuristiques sociales chez son utilisateur, d'autant plus si l'entité en question a un aspect humanoïde. Nous présenterons ici deux volets concernant le sujet de la personnalité dans l'interaction homme-machine : 1/ la perception de la personnalité d'une entité artificielle par son utilisateur et 2/ l'influence de la personnalité de l'entité artificielle sur l'interaction elle-même.

2.2.1 *Percevoir la personnalité*

Les gens sont capables d'attribuer une personnalité à des formes géométriques en mouvements et de les qualifier avec des termes comme "timide" ou "agressif" (Scholl and Tremoulet, 2000). De même, les participants à l'étude de Benfield et al. (2007) n'ont soulevé aucune objection lorsqu'il leur a été demandé de remplir un questionnaire mesurant les dimensions du Big Five pour leur voiture. La consigne était la suivante : "Imaginez que votre véhicule a une personnalité. Répondez aux questions suivantes en vous basant sur la personnalité du véhicule."[2] et les participants ont ainsi évalué si leur véhicule "fait généralement confiance aux autres" ou s'il "est créatif, plein d'idées originales".

Que l'utilisateur perçoive une personnalité n'est donc pas vraiment au cœur du sujet, la véritable question est : les utilisateurs perçoivent-ils la personnalité telle que les chercheurs ont essayé de l'implémenter ? La majorité des travaux en informatique affective sur la personnalité choisissent FFM comme cadre de travail. Les auteurs utilisent donc le plus souvent tout ou partie des échelles associées, bien que celles-ci soient conçues pour les humains. À l'aide de cette méthode, les résultats observés sont généralement positifs : les utilisateurs perçoivent les traits implémentés comme attendu (e. g. Van den Bosch et al. (2012); Neff et al. (2011)).

Cependant, il convient de rester prudent car l'aspect méthodologique est important. Comme le soulève Liu et al. (2015), la façon dont est reportée la personnalité joue sur ce qu'exprime l'utilisateur. Ces auteurs comparent deux méthodes d'évaluation de la personnalité d'un agent virtuel : ils utilisent d'une part, des échelles et de l'autre, des questions ouvertes (du type "Quelle personnalité transmet ce personnage ?"). Liu et al. (2015) s'appuient sur les travaux de Neff et al.

2. « Imagine that your vehicle had a personality. Now rate the following items based on the vehicle's personality. » (Benfield et al., 2007, p. 250)

(2011), qui utilisent un agent virtuel dont les mouvements ont permis de créer une perception d'un névrosisme haut ou bas. L'agent utilisé par Liu et al. (2015) fait des gestes saccadés pour transmettre un niveau de névrosisme haut ou des gestes à la trajectoire plus lisse pour transmettre un niveau de névrosisme bas. Les participants regardaient une vidéo de 15 secondes de l'un ou l'autre des agents, puis ils répondaient à la question ouverte ("Quelle personnalité transmet ce personnage ?") et enfin, ils remplissaient un questionnaire du Big Five en hétéro-attribution.

Liu et al. (2015) montrent alors que les réponses à la question ouverte n'évoquent généralement pas la dimension Névrosisme implémentée. Globalement, l'agent névrotique est décrit comme désagréable et l'agent non-névrotique comme extraverti. Par contre, lors de l'analyse des questionnaires en hétéro-attribution, la perception des agents diffère cette fois sur les dimensions d'Agréabilité et de Névrosisme. Ils montrent ainsi que les deux méthodes apportent des informations différentes sur la perception de la personnalité de l'agent par l'utilisateur.

Par ailleurs, si la perception de la personnalité, consciente et reportée par l'utilisateur, est une information intéressante, l'influence de cette personnalité (que celle-ci ait été "correctement" perçue ou non) sur l'interaction l'est tout autant. Nass et al. (1995) montrent que si la personnalité d'un ordinateur peut être perçue par un utilisateur en utilisant des indices de surface (comme le style linguistique), cette personnalité peut également influencer la perception de l'interaction comme dans un cadre humain-humain.

2.2.2 *Influencer l'interaction*

Concernant l'influence de la personnalité d'un dispositif sur l'interaction, les travaux en informatique affective se sont principalement intéressés à la question de la compatibilité. L'idée sous-jacente est d'étudier si certaines personnalités sont plus "adaptées" pour certains utilisateurs et/ou pour certaines situations. La notion de "bonne" compatibilité peut être évaluée en termes de qualité de la relation (durée dans le temps, satisfaction) (Robins et al., 2000) aussi bien que par rapport à l'efficacité dans un contexte où la relation a un but spécifique (par exemple, la mesure de l'apprentissage dans le cadre d'un enseignement (Karwowski, 2011)).

Concernant la préférence des utilisateurs en termes de personnalité artificielle, la compatibilité est surtout étudiée à travers le prisme du modèle des Big Five et les auteurs se sont particulièrement concentrés sur le trait d'Extraversion. Cela peut s'expliquer par le fait que l'Extraversion est un trait qui semble pouvoir être transmis plus facilement que les autres par des comportements verbaux et non-verbaux.

Dans ce cadre particulier, certains résultats plaident pour la similitude (e. g.une personne introvertie préfèrera interagir avec un dispositif introverti) (Nass and Lee, 2001; Tapus et al., 2008) quand d'autres parlent plutôt de complémentarité (e. g.une personne introvertie préfèrera interagir avec un dispositif extraverti) (Isbister and Nass, 2000; Lee et al., 2006).

Au final, il semble donc difficile de trancher entre "qui se ressemble s'assemble" et "les contraires s'attirent", d'autant que ces effets sont modérés par le rôle social joué par le dispositif en question et les stéréotypes associés à ce rôle (Joosse et al., 2013; Tay et al., 2014).

Concernant les mesures utilisés lors de l'évaluation d'agents ou de robots, le concept de *crédibilité* est très souvent utilisé. La crédibilité représente la capacité de la machine à être perçue comme ayant des comportements plausibles et convaincants (Burgoon et al., 2000) ainsi qu'à être digne de confiance (Tseng and Fogg, 1999). Bien que la définition de la crédibilité soit largement débattue, celle-ci s'évalue souvent par rapport à la constance et à la cohérence des éléments constitutifs de l'agent, et ce à différents niveaux (psychologique comme physique ; intra-personnel aussi bien que social) (Isbister and Doyle, 2002; Niewiadomski et al., 2010). En suivant cette définition, la personnalité peut donc jouer un rôle dans le perception de la crédibilité de l'interaction et du dispositif artificiel.

2.2.3 *Limites des travaux en interaction*

La principale limite des travaux en interaction s'intéressant à la personnalité se situe dans leur durée. En effet, pour des raisons pratiques évidentes, les expérimentations évaluant la perception de personnalité artificielle par des utilisateurs se basent souvent sur une unique interaction généralement courte (inférieure à 30 minutes). Malheureusement, la durée d'une telle interaction ne permet pas à l'utilisateur de se construire une réelle idée de la personnalité. L'utilisateur reste plutôt sur une formation d'impression (Fiske and Neuberg, 1990), un processus sujet à de nombreux biais cognitifs. Dans la formation d'impression, les personnes commencent par catégoriser rapidement et automatiquement l'individu à juger, en fonction d'informations de surface et de stéréotypes. En fonction de l'intérêt porté à l'individu à juger et du nombre d'informations disponibles sur celui-ci, les personnes pourront au cours du temps affiner leur jugement, jusqu'à considérer cet individu comme quelqu'un d'unique (et pas seulement comme faisant partie d'une catégorie).

Cependant, cette formation d'impression peut parfois être considérée comme suffisante dans le cadre des interactions homme-machine

à cause d'une autre limite de ces travaux : l'utilisation de personnalités "extrêmes". En prenant l'exemple d'études utilisant les Big Five, les valeurs attribuées pour les traits utilisés sont souvent poussées au maximum afin d'obtenir des comportements bien différenciés. Cette méthode permet de mettre en avant les traits étudiés mais donne des comportements très stéréotypés. Chez l'humain, ce type de personnalité est extrêmement rare. Les études psychométriques du NEO-PI-R montrent qu'au plus, les personnes présentent quelques facettes fortes mais rarement un trait sur-affirmé. De plus, certaines études combinent une expression de la personnalité par les actions avec l'aspect visuel de l'agent. De ce fait, il est difficile de distinguer la part du visuel et la part du comportemental dans le jugement produit.

2.3 POSITIONNEMENT ET PROBLÉMATIQUE

À l'instar de Vinciarelli and Mohammadi (2014a), nous concevons le champ de recherche concernant la personnalité comme un domaine commun où de multiples disciplines, incluant l'informatique et la psychologie, peuvent contribuer et bénéficier mutuellement des avancées de chacun : les progrès sur les théories de la personnalité devraient aider à construire plus efficacement des machines avec une personnalité et inversement[3].

Dans cette optique, il apparaît comme important de s'interroger sur les concepts psychologiques utilisés en informatique afin de permettre un bénéfice mutuel. Bien que celles-ci soient moins utilisées, les théories socio-cognitives, intéressées par les structures et les processus de la personnalité, fournissent une base de modélisation intéressante pour les chercheurs en informatique. Et si les modèles computationnels produits sur cette base restent interprétables, alors l'informatique affective peut permettre de nourrir la réflexion des psychologues sur leurs modèles. S'il ne s'agit pas d'une validation à proprement parler, des modèles informatiques construits avec cette approche pourraient permettrent d'interroger en retour les modèles psychologiques utilisés et leur capacité à représenter des processus produisant des comportements cohérents.

Au vu de la littérature et des limites évoquées, notre problématique s'articule donc autour de deux axes : proposer un modèle de personnalité basé sur une approche socio-cognitive et réaliser une implémentation de façon à informer le concept de personnalité du point de vue de la psychologie.

3. « a common ground where multiple disciplines, including computing and psychology, can contribute and mutually benefit from each other : progress in personality theory should help to build more effective personality machines and vice versa » (Vinciarelli and Mohammadi, 2014a, p. 291)

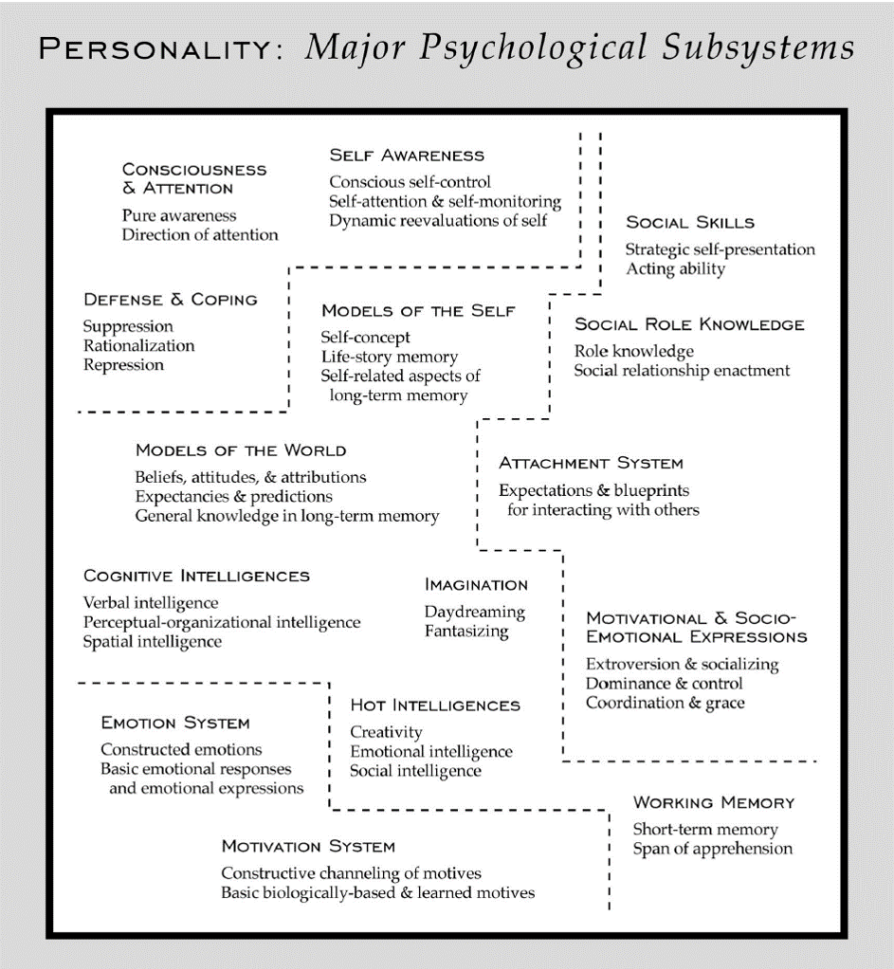


FIGURE 4 – Sous-systèmes psychologiques à la base de la personnalité. De Mayer (2005)

Comme nous l'avons vu, la conception socio-cognitive nous pousse à nous intéresser aux structures et processus impliqués dans l'expression de la personnalité. Et ceux-ci sont nombreux comme le montre la Figure 4. Ces processus ne peuvent être tous abordés ensemble en un seul modèle. Il convient donc de réduire notre objet d'étude à un processus en particulier. Nous proposons de nous intéresser à la régulation du Soi, choix défendu au chapitre 3 dans lequel nous présenterons également un modèle computationnel basé sur la théorie du regulatory focus (Higgins, 1997).

Concernant l'implémentation, nous proposons d'utiliser une approche mixte, à la fois theory-driven et data-driven. En collectant des données humaines et en les utilisant à l'aide de techniques d'apprentissage automatique, une première implémentation dénuée d'aprioris pourra être obtenue et utilisée pour réaliser des expérimentations (cf. 5). Nous pourrons également utiliser les résultats de cet apprentis-

sage automatique pour guider l'implémentation de processus symboliques (cf. chapitre 6).

Deuxième partie

CONTRIBUTIONS THÉORIQUES ET MÉTHODOLOGIQUES

3

MODÉLISER LA RÉGULATION DU SOI

Sommaire

3.1 Personnalité et régulation du Soi **34**
- 3.1.1 Soi et personnalité 34
- 3.1.2 Utiliser la théorie du regulatory focus 36
- 3.1.3 Intérêt de la théorie du regulatory focus pour l'implémentation d'une personnalité artificielle 41

3.2 Perseed, ou représenter la régulation du Soi en tant que composante d'une personnalité artificielle **41**
- 3.2.1 Perseed : Cadre conceptuel 42
- 3.2.2 Perseed : modèle computationnel 43
- 3.2.3 Discussion 45

3.3 Résumé du chapitre **48**

Comme présenté au chapitre 2, le fait de prendre une approche socio-cognitive rend la modélisation informatique délicate. En effet, en comparaison d'une approche dimensionnelle comme celle du Big Five, nous passons de cinq traits dits orthogonaux à un ensemble de multiples structures et de processus en interaction. Pour tirer parti au mieux de cette approche, nous restreignons le champ des possibles pour nous concentrer sur un concept spécifique : la régulation du Soi dont nous montrerons l'intérêt pour la modélisation informatique et l'interaction homme-machine (section 3.1). Nous présenterons ensuite la théorie du regulatory focus sur laquelle nous basons nos travaux (sous-section 3.1.2). Enfin, nous proposerons une première contribution : Perseed, à la fois cadre conceptuel pour l'utilisation de la régulation du Soi en informatique affective (sous-section 3.2.1) et modèle computationnel pour l'implémentation du regulatory focus (sous-section 3.2.2).

3.1 personnalité et régulation du soi

Utiliser une approche socio-cognitive pour modéliser une personnalité artificielle nous pousse à avoir une démarche intellectuelle de spécifications successives afin d'aller du concept de personnalité à l'utilisation d'une théorie pour baser notre modélisation d'*une* composante de la personnalité pour une entité artificielle.

En effet, il est impossible de produire en un seul modèle computationnel la représentation de l'ensemble des structures et des processus impliqués dans le phénomène psychologique désigné par le terme "personnalité". C'est pourquoi nous proposons de nous intéresser à la personnalité au travers du prisme du concept de Soi. Le concept du Soi sera lui-même réduit au concept de régulation du Soi, abordé au travers de la théorie du regulatory focus (Higgins, 1997). L'organisation des concepts et théories qui seront évoqués au cours de ce chapitre est présentée en Figure 5.

Ces différents concepts et théories sont présentés successivement dans les sous-sections suivantes, du plus général au plus spécifique.

3.1.1 *Soi et personnalité*

En l'état actuel des connaissances, il n'existe pas de définition communément admise de ce qu'est le Soi (Klein, 2012). Dans le modèle historique de James (1890), le Soi se partage entre le « Je » (le soi-agent, connaissant) et le « Moi » (le soi-objet, connu). Le « Moi » peut être considéré comme l'ensemble des représentations cognitives d'une personne sur elle-même. Le « Je » se réfère à une forme de conscience de soi, au sentiment interne d'être soi. À l'opposé, pour Skowronski (2012), le Soi n'existe pas en tant que tel. De la même façon que la

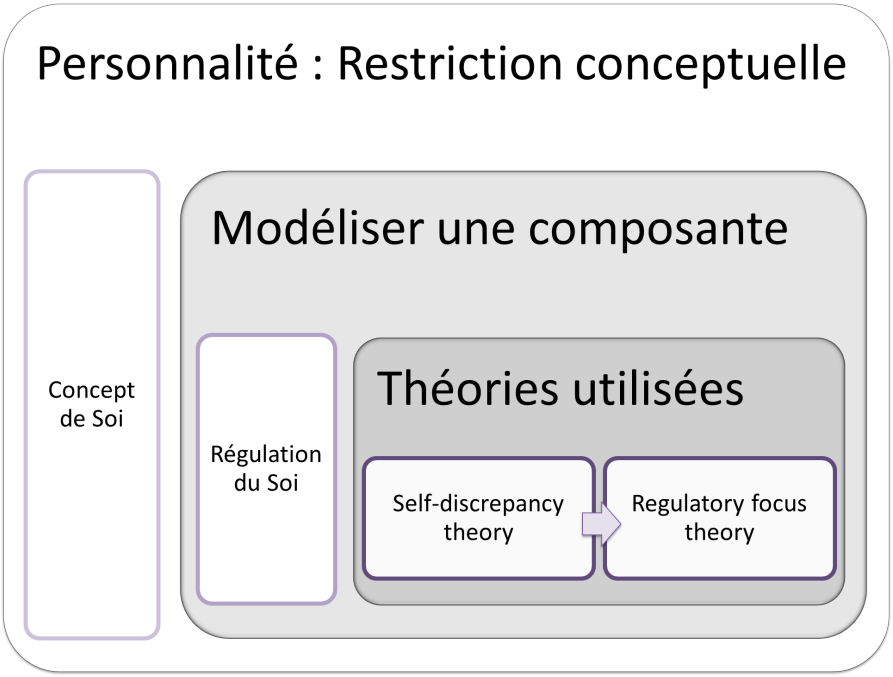


FIGURE 5 – Imbrication des concepts et théories utilisés dans le cadre de notre modélisation

mémoire humaine peut se décliner en différents sous-systèmes, le concept de Soi pourrait se décliner en différents sous-systèmes possédant certes des similitudes mais surtout des caractéristiques qui les distinguent dans leur structure et leur fonctionnement. C'est alors l'interaction de différents sous-systèmes mentaux qui crée l'impression d'être soi-même.

Malgré les questions ouvertes régnant autour de la définition même du Soi, beaucoup de chercheurs insistent sur le rôle du Soi dans la personnalité (Block, 1993; Damian and Robins, 2012; Rothbart, 2007) : pour certains, le concept de Soi est nécessaire pour comprendre pleinement les processus de la personnalité – les processus qui génèrent et régulent les pensées, les sensations et les comportements [1]. En outre, le Soi est utilisé par Cervone (2005) pour expliquer ce qui peut paraître être un paradoxe de la définition socio-cognitive de la personnalité : le fait que les personnes se définissent comme ayant une personnalité stable malgré leurs variations comportementales. Pour Cervone (2005), les aspects stables des connaissances sur le Soi (par exemple, les schémas de soi) contribueraient à maintenir une cohérence inter-situationnelle globale des comportements, alors que ses aspects dynamiques (comme les évaluations d'auto-efficacité dans une

1. « In fact, the concept of Self is necessary for a complete understanding of personality processes—the processes that generate and regulate thoughts, feelings, and behaviors. » (Damian and Robins, 2012, p. 432)

situation particulière) expliqueraient la variabilité inter-situationnelle desdits comportements.

Le concept de Soi est peu utilisé en informatique affective. Pourtant, les théories du Soi sont pertinentes pour éclairer les recherches s'intéressant à créer des dispositifs qui puissent être considérés comme des acteurs sociaux individualisés car le Soi est justement lié au sentiment d'individualité. De plus, doter des agents virtuels de mécanismes de type "conscience de Soi" pourrait permettre d'augmenter leur crédibilité (Ijaz et al., 2011).

Cependant, si le concept de Soi est intéressant en lui-même mais trop complexe pour être modélisé tel quel. Dans cette thèse, nous nous limiterons au concept de régulation du Soi, qui recouvre les mécanismes dédiés à l'alignement des attentes ou des désirs avec la réalité (Hoyle, 2010). Les processus de régulation du Soi peuvent s'interpréter dans un cadre socio-cognitif comme le méta-modèle CAPS, comme le montrent Ayduk and Gyurak (2008). De plus, les processus de régulation du Soi sont à la base de comportements pouvant être liés au modèle FFM (Fein and Klein, 2011).

3.1.2 *Utiliser la théorie du regulatory focus*

Pour étudier l'apport de la personnalité dans le domaine de l'interaction homme-machine, nous proposons donc d'étudier le concept du Soi, et plus spécifiquement la régulation du Soi. Et toujours plus spécifiquement, nous proposons d'étudier la régulation du Soi au travers de la *regulatory focus theory* (Higgins, 1997), qui se base sur un première théorie du même auteur, la *self discrepancy theory* (Higgins, 1987). Nous présenterons ici ces deux théories dans l'ordre chronologique avant de justifier de leur intérêt pour notre modélisation computationnelle d'une composante de la personnalité.

3.1.2.1 *La théorie de la self discrepancy*

La théorie de la self discrepancy (Higgins, 1987) postule trois domaines du Soi et deux points de vue. Les domaines du Soi sont le soi actuel, le soi idéal et le soi imposé.

SOI ACTUEL Le soi actuel (*actual self*) se réfère au concept de Soi d'une personne P. Il est constitué des représentations des attributs que quelqu'un (P inclus) pense que P possède à un instant t.

SOI IDÉAL Le soi idéal (*ideal self*) est constitué des représentations des attributs que quelqu'un (P inclus) aimerait que P possède. Le soi idéal est associé aux gains et aux récompenses : P pense qu'il gagnerait à se rapprocher de son soi idéal.

SOI IMPOSÉ Le soi imposé (*ought self*) est constitué des représentations des attributs que quelqu'un (P inclus) pense que P devrait posséder. Le soi imposé est lié aux pertes et aux sanctions : P pense qu'il a à y perdre en ne se conformant pas à son soi imposé.

Ces domaines du Soi se distribuent sur deux points de vue : le propre point de vue de P et les points de vue de personnes signifiantes pour P (c'est-à-dire des personnes qui comptent pour P). Les représentations associées aux points de vue de ces personnes signifiantes sont des croyances que P leur attribue. Cette approche rappelle la théorie de l'esprit, la capacité à attribuer des états mentaux à d'autres (Baron-Cohen et al., 1985). L'étude de la théorie de l'esprit s'intéresse à la construction des croyances qui sont attribuées à autrui et à la façon dont ces croyances sont utilisées pour expliquer ou prédire le comportement des autres. Ici, nous nous concentrons uniquement sur la façon dont ces croyances dirigent les propres comportements de celui qui les crée.

La self discrepancy theory déclare que les sois idéaux et imposés sont donc des "guides de soi" pour P. Ils constituent des images de soi à atteindre pour P. Les personnes utilisent donc leurs "guides de soi" (soi idéal et soi imposé) et les comparent à leur soi actuel afin d'adapter leur comportement dans le but de réduire l'écart entre ces différentes vues du Soi. Les différences comportementales peuvent être expliquées par les différences de contenu de ces sois, mais également par les différentes manières d'utiliser ces contenus. L'aspect théorique concernant ces différentes façons d'utiliser ces sois est développé dans la regulatory focus theory.

3.1.2.2 *La théorie du regulatory focus*

La théorie du regulatory focus (Higgins, 1997) fait la distinction entre deux stratégies de régulation du Soi nommées *promotion-focus* et *prévention-focus*. Ces deux stratégies se distinguent par l'utilisation des guides de soi mais également par les besoins que la personne cherche à satisfaire et par son appréhension des situations, ce qui donne lieu à des différences aussi bien comportementales qu'émotionnelles détaillées ci-après.

STRATÉGIE PROMOTION-FOCUS (FIGURE 6) L'utilisation d'une stratégie promotion-focus permet de répondre en premier lieu aux besoins d'épanouissement du sujet. Pour cela, le sujet va accorder une plus grande importance à ses sois idéaux et évaluer les situations en termes de gains (situations positives) ou de non-gains (situations négatives). Pour illustrer la notion de gain et de non-gain, imaginons une loterie où, pour une mise de 1 €, le joueur peut recevoir l'un des trois retours suivant : 2 €, 1 € ou 0 €. Dans cette situation, obtenir

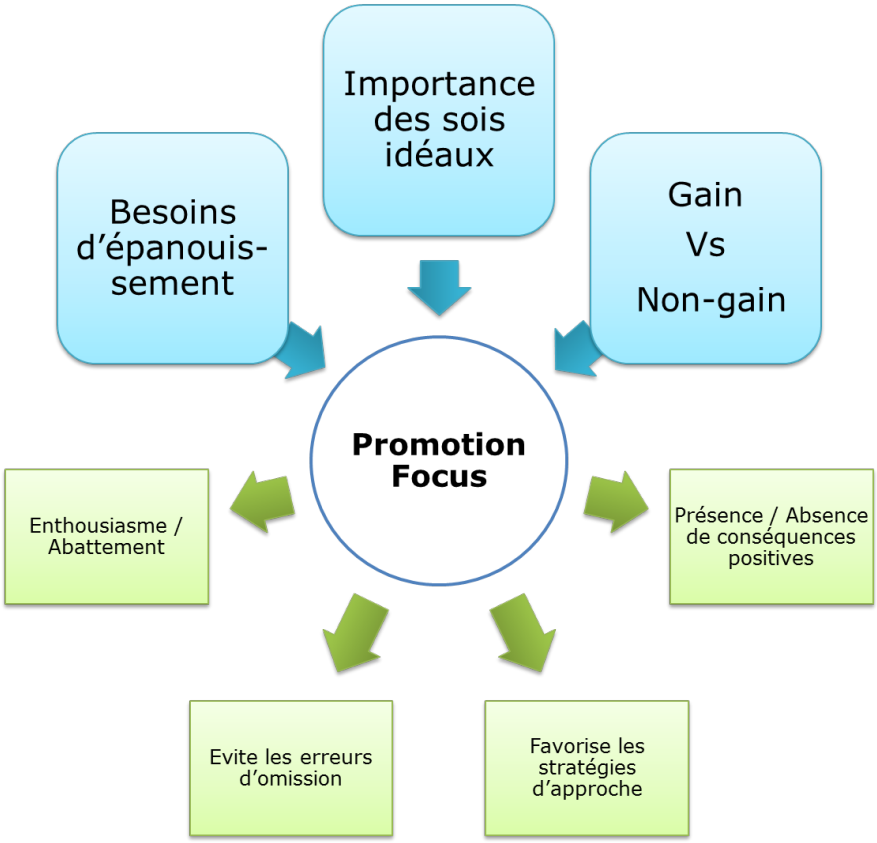


FIGURE 6 – Variables psychologiques induisant une stratégie promotion-focus (haut) et variables psychologiques affectées en conséquence (bas)

2 € est un gain alors que les deux autres possibilités, 1 ou 0 €, sont des non-gains.

En conséquence, un sujet adoptant une stratégie promotion-focus favorise l'atteinte des buts par des stratégies d'approche, c'est-à-dire qu'il va chercher à s'approcher d'une correspondance entre ses buts et la réalité. Il va également s'assurer contre les erreurs d'omission (faux négatifs). L'adoption d'une stratégie promotion-focus donne au sujet une sensibilité émotionnelle et évaluative : cette sensibilité émotionnelle s'articule sur un axe enthousiasme (valence positive) / abattement (valence négative). En termes d'évaluation d'une situation, le sujet sera sensible à la présence (ou à l'absence) de conséquences positives de ses comportements.

STRATÉGIE PRÉVENTION-FOCUS (FIGURE 7) La mise en place d'une stratégie prévention-focus permet de répondre principalement aux besoins de sécurité du sujet. Pour cela, le sujet va accorder une plus grande importance à ses sois imposés et évaluer les situations en termes de pertes (situations négatives) ou de non-pertes (situations positives). Si nous reprenons l'exemple de la loterie présentée plus

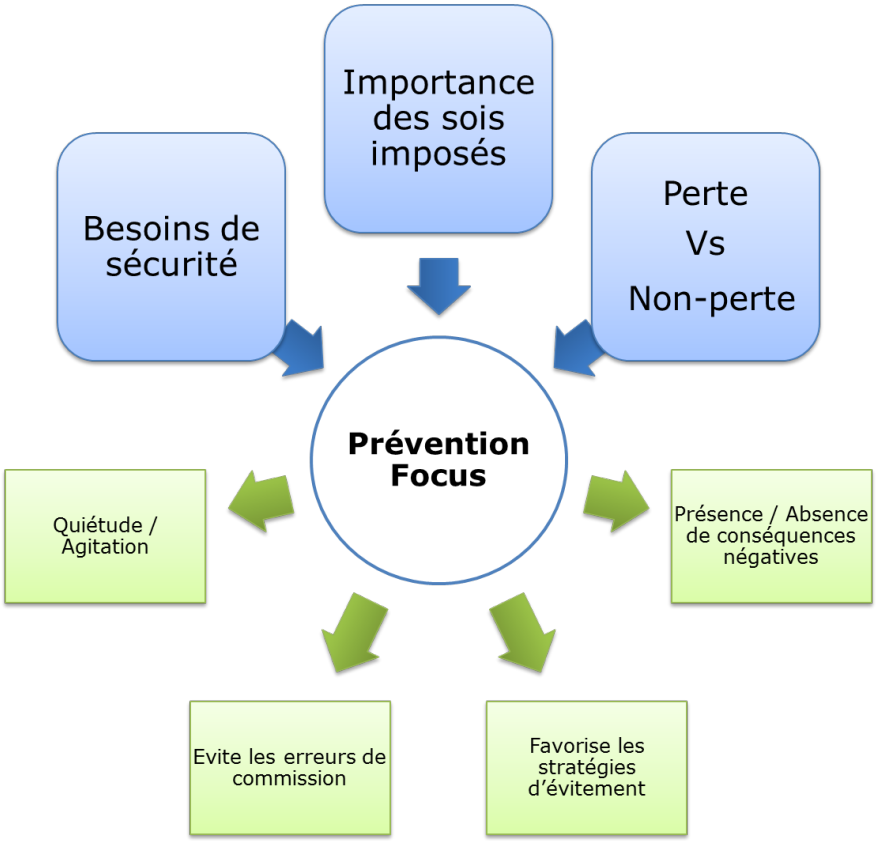


FIGURE 7 – Variables psychologiques induisant une stratégie prévention-focus (haut) et variables psychologiques affectées en conséquence (bas)

haut, obtenir 0 € est une perte alors que les deux autres possibilités, 1 ou 2 €, sont des non-pertes.

En conséquence, un sujet adoptant une stratégie prévention-focus favorise l'atteinte des buts par des stratégies d'évitement, c'est-à-dire qu'il va chercher à éviter une non-correspondance entre ses buts et la réalité. Il va s'assurer contre les erreurs de commission (faux positifs). L'adoption d'une stratégie prévention-focus donne au sujet une sensibilité émotionnelle et évaluative : cette sensibilité émotionnelle s'articule sur un axe quiétude (valence positive) / agitation (valence négative). En termes d'évaluation d'une situation, le sujet sera sensible à la présence (ou à l'absence) de conséquences négatives de ses comportements.

Ces deux modes stratégiques (promotion et prévention) sont disponibles pour un même individu. En fonction de son éducation et de ses expériences passées, un individu développe une préférence pour l'un ou l'autre de ces modes (Keller, 2008) : on parle alors de focus chronique. Plus cette préférence est marquée, plus le focus chronique est fort. Cependant, le choix d'une stratégie peut également être in-

duit par le contexte : présenter une tâche avec une consigne mettant en avant des aspects liés à la stratégie promotion-focus (comme les besoins d'accomplissement ou les gains/non-gains) induit un focus promotion (*Résolvez un labyrinthe pour qu'une souris obtienne du fromage* (Friedman and Förster, 2001) ; *En réussissant 90% des essais dans cette tâche, vous pouvez gagner 1$ supplémentaire* (Shah et al., 1998)). La même consigne mettant en avant des aspects liés à la stratégie prévention-focus (comme les besoins de sécurité ou les pertes/non-pertes) induit un focus prévention (*Résolvez un labyrinthe pour éviter qu'une souris soit mangée par une chouette* (Friedman and Förster, 2001) ; *En ne ratant pas plus de 10% des essais dans cette tâche, vous ne perdrez pas 1$* (Shah et al., 1998)).

3.1.2.3 *Regulatory focus et regulatory fit*

Lorsque un but est poursuivi d'une façon qui correspond à l'orientation du sujet, il se produit un effet de *regulatory fit* (Higgins, 2005). En d'autres termes, lorsque les moyens utilisés par une personne pour approcher un état désiré correspondent aux moyens rattachés au regulatory focus qu'elle éprouve, la personne ressent un sentiment de justesse à propos du but poursuivi et son engagement dans la poursuite de ce but s'en trouve augmentée. Expérimenter un effet de regulatory fit a une influence sur les jugements produits par la personne, sur sa prise de décision, son attitude envers un objet ou ses performances dans une tâche (Higgins, 2005). Par exemple, un utilisateur dans un focus promotion jugera une marque plus positivement si celle-ci lui a été présentée par des messages mettant en avant les gains potentiels lors de l'utilisation d'un produit de cette marque (et respectivement pour un focus prévention et des messages mettant en avant les risques évités lors de l'utilisation du produit) (Lee and Aaker, 2004). De la même façon, des participants ont montré de meilleures performances dans une tâche de prédiction lorsque le contenu de la tâche et la consigne s'inscrivent dans le même regulatory focus (contenu promotion/consigne présentant la tâche comme amusante ; contenu prévention/consigne présentant la tâche comme importante) (Bianco et al., 2003).

L'effet de regulatory fit augmente donc le sentiment de justesse ressentie dans l'utilisation d'une stratégie concordante avec le focus éprouvé (qu'il soit chronique ou induit). Ce sentiment de justesse peut alors augmenter la valeur perçue (indépendamment de sa valeur hédonique : une valence positive sera perçue comme meilleure, une valence négative sera perçue comme moins bonne) et le sentiment d'importance de la tâche. Ce transfert n'est pas le résultat d'une humeur positive induite par le sentiment de justesse mais se produit par confusion des valeurs : la "justesse" de la stratégie employée dans la tâche déteint sur la perception de la tâche ou des objets la concer-

nant, qui sont également perçus comme plus "justes" (Camacho et al., 2003).

3.1.3 *Intérêt de la théorie du regulatory focus pour l'implémentation d'une personnalité artificielle*

La théorie du regulatory focus offre une vue intéressante sur les structures et processus mis en jeu dans la régulation du Soi en allant au-delà du fait que les personnes cherchent à approcher le plaisir et à éviter la souffrance (Higgins, 1997). Par exemple, la théorie de la personnalité de Gray (1990) (présentée au chapitre 2) s'intéresse également à la personnalité et à la régulation du Soi au travers du prisme de l'approche/évitement et des systèmes motivationnels associés. Cependant, cette théorie n'évoque pas directement les différentes manières d'atteindre un état désiré, pas plus qu'elle ne distingue les différences possibles sur la nature de ces états (e. g. approcher une "récompense" est équivalent à approcher une "non-punition"). Utiliser la théorie du regulatory-focus pour guider l'implémentation d'une personnalité artificielle est donc intéressant car cette théorie fournit des informations sur les différentes croyances et les différents modes d'évaluation des situations qui donnent lieu à des différences inter- et intra-individuelles.

De plus, l'effet de regulatory fit a été exploré dans différentes domaines comme les environnements de travail (Park et al., 2015), le sport (Debanne et al., 2014) ou la communication (Ludolph and Schulz, 2015) mais pas dans le domaine de l'informatique affective. Cependant, si un dispositif comme un compagnon artificiel peut créer un état de regulatory fit avec un utilisateur, cela pourrait augmenter l'engagement de cet utilisateur dans la tâche réalisée et envers le dispositif lui-même. Le sentiment de justesse créé par la situation de fit pourrait potentiellement être transféré au dispositif, augmentant par là son agréabilité et/ou sa crédibilité aux yeux de l'utilisateur.

3.2 PERSEED, OU REPRÉSENTER LA RÉGULATION DU SOI EN TANT QUE COMPOSANTE D'UNE PERSONNALITÉ ARTIFICIELLE

Il n'existe pas de modèle en informatique affective permettant la conception et l'implémentation d'une personnalité artificielle basée sur la régulation du Soi. En prenant appui sur la psychologie de la personnalité et la théorie du regulatory focus, nous proposons Perseed [2], un cadre conceptuel et computationnel permettant de repré-

2. Concaténation de *personality*, la personnalité, et de *seed*, la graine, au sens où le modèle va "planter" des éléments de personnalité au sein d'une architecture cognitive qui vont "éclore" pour se manifester dans les comportements de l'entité dirigée.

senter la régulation du Soi (comme composante de la personnalité) chez une entité artificielle.

3.2.1 *Perseed : Cadre conceptuel*

Les différents éléments de notre cadre de travail sont inspirés par la théorie de la self discrepancy (présentée en sous-sous-section 3.1.2.1) et par le dynamic self-regulatory processing framework (présenté au chapitre 2 - sous-section 2.1.2). Le coeur de Perseed est composé de deux parties : 1/ un réseau de sois et d'attributs (i. e. des structures et des processus concernant le Soi) et 2/ des règles d'injection (i. e. des règles sélectionnant les attributs à utiliser).

3.2.1.1 *Images de Soi*

Le premier élément du cadre conceptuel que nous proposons est une collection d'images de Soi (ou self-images), inspirée par la self-discrepancy theory (Higgins, 1987). Comme dans la théorie, ces self-images sont liées à différents points de vue : le point de vue du dispositif lui-même et ceux des personnes signifiantes. Par exemple, dans le cas de compagnons artificiels pour les enfants, l'utilisateur principal, c'est-à-dire l'enfant, est signifiant. Nous pouvons également ajouter d'autres acteurs jouant un rôle important dans l'environnement de l'enfant, comme les parents de l'enfant ou un enseignant. Pour chaque point de vue, un soi idéal et un soi imposé sont représentés. Ces sois sont reliés à des attributs, décrits ci-après.

3.2.1.2 *Attributs*

Les trois types d'attributs présentés ci-après reprennent les concepts proposés dans le dynamic self-regulatory processing framework (Morf, 2006) en les opérationnalisant dans un cadre informatique.

CONNAISSANCES SUR LE SOI Les connaissances sur le Soi sont des attributs qui sont en rapport avec la représentation interne que l'entité artificielle a d'elle-même. Il peut s'agir par exemple de buts (e. g. être drôle) et de croyances (e. g. être efficace dans une certaine tâche) concernant le Soi.

STRATÉGIES INTRAPERSONNELLES Les stratégies intrapersonnelles sont des processus permettant à l'entité de maintenir une certaine image d'elle-même. Ces attributs influencent les mécanismes internes de l'entité artificielle, qui sont principalement invisibles pour l'utilisateur. Il peut s'agir par exemple de modes d'évaluation particuliers (e. g. pondération importante sur la contrôlabilité de la situation), de filtres perceptifs (e. g. prise en compte des informations à valence positive uniquement), de règles d'attribution causale (e. g. les effets po-

sitifs ont toujours une cause externe) ou de sensibilité aux émotions (e. g. les émotions à valence négative sont ressenties plus fortement).

STRATÉGIES INTERPERSONNELLES Les stratégies interpersonnelles sont des processus permettant à l'entité de maintenir une certaine image d'elle-même dans ses relations avec les autres. Ces attributs influencent des mécanismes qui joue particulièrement sur l'interaction en tant que telle. Il peut s'agir par exemple de schémas de comportements spécifiques (e. g. heuristiques personnelles), de préférences d'action (e. g. pondération sur certaines actions) ou de comportements non-verbaux (e. g. ajout de gestes non signifiants).

3.2.1.3 *Règles d'injection*

Nous choisissons d'appeler règles d'injection les règles déterminant comment sont utilisés les différents attributs. L'injection est donc un processus qui sélectionne les attributs et les transpose au sein du dispositif artificiel, en altérant alors les fonctionnalités comme la prise de décision, la sélection d'action, l'autoévaluation ou les mécanismes émotionnels. L'injection suit des règles définies, qui correspondent à des façons d'utiliser les contenus des sois (ici, les attributs) s'inspirant de la psychologie. Les stratégies promotion-focus et prevention-focus sont des exemples de règles pouvant être utilisées. Par exemple, un agent virtuel promotion-focus injectera des attributs liés à ses sois idéaux, alors qu'un agent prevention-focus injectera des attributs liés à ses sois imposés. Dans la même idée, un agent égoïste injectera principalement des attributs liés à ses propres sois, tandis qu'un agent absolument altruiste n'injectera que des attributs liés aux points de vue des autres.

3.2.2 *Perseed : modèle computationnel*

Concernant le modèle computationnel, nous nous concentrerons sur la mise en place d'un modèle permettant d'instancier un regulatory focus chez un agent virtuel. Notre modèle fonctionne en conjonction avec une architecture offrant des fonctionnalités cognitives. Nous ne faisons pas de présupposés quant à l'architecture cognitive spécifique qui pourrait être utilisée. Le modèle modifie seulement les fonctionnalités offertes par cette architecture quelles qu'elles soient. Le module de personnalité peut influer sur la façon dont l'information est perçue ou sur la manière dont les actions sont effectuées. Il peut également modifier une fonctionnalité cognitive proprement dite (e. g. la représentation des connaissances, la sélection d'action ou les mécanismes émotionnels). Nous laisserons donc le terme "architecture cognitive" non spécifié dans la proposition ci-dessous, mais, à des fins de compréhension, nous préciserons les fonctions attendues de cette architecture dans le cadre d'un exemple.

Afin de simplifier la structure, les self-images idéal et imposé seront respectivement considérées comme des self-images promotion (Pro) et prévention (Pre) mais sont toujours représentées suivant différents points de vue. Les attributs sont représentés par un ensemble de paires clés-valeurs. Les clés représentent les attributs définis plus haut en utilisant les concepts propres à l'architecture visée. Les clés sont donc dépendantes de l'architecture utilisée, une interface ad hoc doit être programmée pour chaque architecture spécifique. Un attribut sera donc différemment représenté en fonction des spécificités de l'architecture. Par exemple, dans une architecture de type BDI (Rao and Georgeff, 1995), un attribut de type but (connaissance sur le Soi) peut être représenté par une clé "désir". Dans une architecture CLARION (Sun, 2006), un attribut de type but peut être représenté par une clé "but explicite" ou "motivation". Les valeurs rattachées aux clés correspondent aux éléments concrets à injecter dans l'architecture. Pour reprendre l'exemple précédent, dans une architecture de type BDI, la valeur attachée à une clé "désir" peut être un désir associé à des croyances et des plans correspondants à sa réalisation. Les valeurs peuvent être génériques (pour affecter un comportement général) mais sont généralement dépendantes de l'application développée.

La Figure 8 illustre cette organisation pour un agent de type "copain de jeu". Dans cet exemple, nous postulons une architecture cognitive ayant les fonctionnalités suivantes : 1/ réalisation d'actions ; 2/ perception de l'état émotionnel de l'utilisateur ; 3/ sélection d'actions orientée buts et 4/ un modèle émotionnel pour l'agent. L'agent

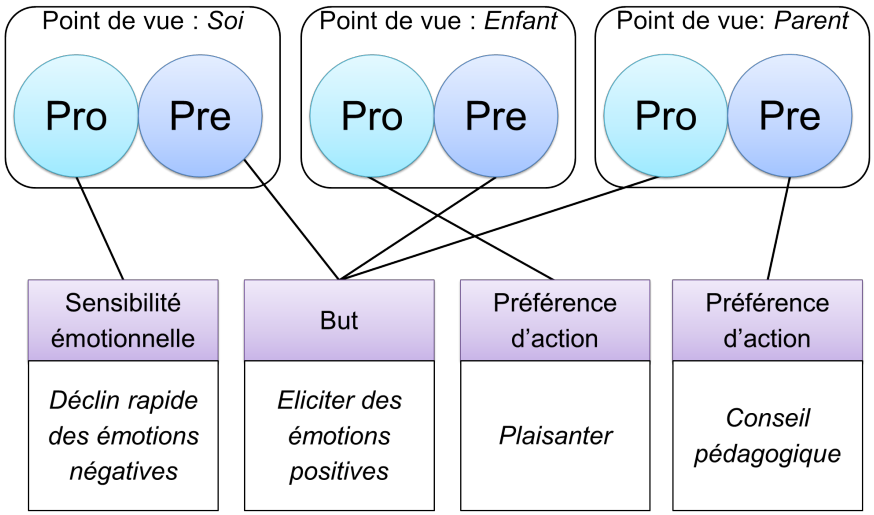


FIGURE 8 – Exemple d'organisation d'un réseau de self-images et d'attributs (sous forme de paires clé-valeur) pour un agent de type "copain de jeu"

a trois points de vue différents : le sien propre, le point de vue de l'enfant et le point de vue d'un parent de l'enfant. Par rapport à lui-même, l'agent voudrait avoir une faible sensibilité aux émotions négatives (aspect Pro) et pense qu'il devrait susciter des émotions positives chez son utilisateur (aspect Pre). Ce dernier but est nécessaire du point de vue de l'enfant (Pre) mais seulement désirable pour le parent (Pro). L'enfant aimerait que l'agent fasse des plaisanteries (Pro) alors que pour le parent, l'agent doit profiter du jeu pour donner des conseils d'ordre pédagogique (Pre).

L'injection peut avoir lieu une seule fois, à l'initialisation (représentation du focus chronique) ou être rejouée en fonction de retours fournis par l'architecture cognitive. Dans ce dernier cas, on peut imaginer qu'en fonction du feedback de l'architecture, le module puisse détecter un contexte promotion/prévention et adapter l'injection en fonction. De cette façon, la régulation est rendue plus dynamique (représentation du focus induit). Les attributs injectés dans cette architecture ajoutent de nouvelles données ou de nouvelles fonctionnalités à l'architecture ou bien remplacent des données ou des fonctionnalités existantes. Si le langage de programmation de l'architecture permet de faire du hot-swapping logiciel, les attributs représentant des processus pourraient se substituer aux méthodes existantes uniquement à l'exécution (e. g. exécution d'un nouveau planificateur dans une architecture BDI). Si les règles de sélection des valeurs à injecter sont propres à Perseed, la réalisation concrète de l'injection est également dépendante de l'architecture (par exemple, dans une architecture CLARION, injecter une "motivation" revient à insérer un élément dans le sous-système motivationnel).

Comme illustré en Figure 9, un agent présentant l'organisation présentée plus haut et utilisant une stratégie promotion-focus utilisera donc les attributs liés aux images Pro. La fonction de déclin des émotions négatives du modèle émotionnel supposé sera remplacée par une nouvelle fonction fournie par le module de personnalité. Un but de haut-niveau concernant le fait d'éliciter des émotions positives chez l'utilisateur sera ajouté à la liste des buts. L'action "plaisanter" sera pondérée de façon à être plus facilement sélectionnée si l'occasion se présente.

3.2.3 *Discussion*

3.2.3.1 *Intérêts*

Perseed offre plusieurs intérêts pour explorer la façon dont les différences dans les structures et les processus de régulation du Soi sont perçues en termes de profils de personnalité chez des dispositifs artificiels.

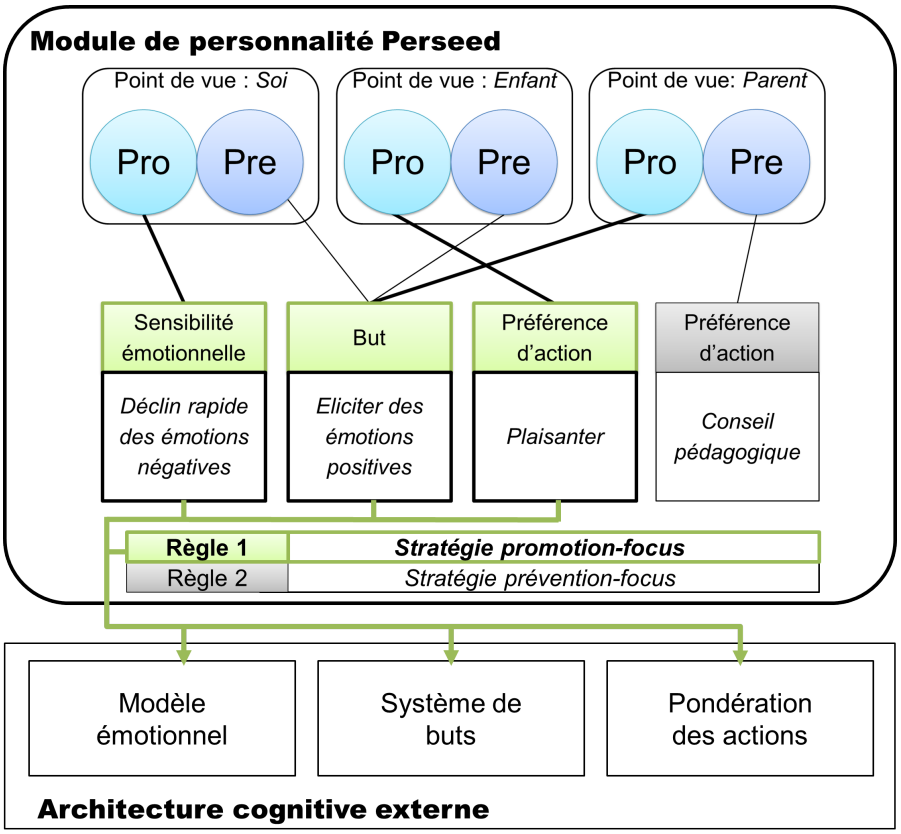


FIGURE 9 – Illustration de l'injection d'attributs dans une architecture cognitive via une règle promotion-focus

Les différences individuelles sont modélisées à trois niveaux : 1/ par le contenu des attributs, 2/ par les liens entre les sois et les attributs, et 3/ par les règles d'injection des attributs. Il est donc possible de mettre en place des différences avec un niveau de granularité assez fin.

Notre modèle est extensible car de nouveaux attributs et de nouvelles règles d'injection peuvent être ajoutés sans modifier le cœur du module ou d'éventuelles configurations préexistantes. De la même façon, il est possible d'étendre le modèle afin que celui-ci prenne en compte différents contextes sociaux. Par exemple, un compagnon qui tiendrait deux rôles sociaux distincts (comme compagnon de jeu et tuteur) pourrait avoir deux collections de self-images, une pour chaque rôle social. Les attributs reliés aux self-images seraient soit partagés (représentation de la stabilité de la personnalité), soit spécifique à un rôle (représentation de la variabilité intra-individuelle).

De plus, Perseed est flexible puisque ce modèle est conçu comme un plug-in externe plutôt que comme une brique intégrée à une architecture cognitive spécifique.

3.2.3.2 *Limites*

En l'état, Perseed présente également un certain nombre de limites. Par exemple, d'un point de vue théorique, la théorie de la self-discrepancy met en avant le processus de comparaison du Soi actuel avec les autres images du Soi comme guidant la régulation du Soi. Cependant, notre modèle n'apporte pas pour le moment de spécifications quant au maintien et la mise à jour d'une image interne représentant un Soi actuel.

D'un point de vue computationnel, la gestion d'éventuels conflits intervenant lors de la sélection d'attributs ayant des valeurs antagonistes ne fait pas partie intégrante du modèle. Enfin, si Perseed offre un cadre de travail orientant la mise en place de la régulation du Soi comme processus permettant l'émergence d'une personnalité, il est à noter que les éléments concrets du modèle (valeurs des attributs) restent logiquement sous-spécifiés.

3.2.3.3 *Nouvelles problématiques*

Le modèle Perseed, qui représente notre première contribution, amène de nouvelles problématiques quant à notre recherche. Comme nous le disions plus haut, les valeurs des attributs ne peuvent être spécifiées au niveau général. Il est donc nécessaire de définir une application pour nous permettre de tester une implémentation du regulatory focus. Cependant, la théorie du regulatory focus n'a jamais été utilisée dans le cadre de l'informatique affective à notre connaissance. Il convient donc dans un premier temps de tester si ce concept peut être transmis par un agent virtuel, à l'instar des dimensions du Big Five. Cette implémentation devra également permettre de répondre à notre problématique d'informer le concept de regulatory focus du point de vue de la psychologie comme présenté en section 2.3.

De plus, puisque nous envisageons de réaliser des tests de perception de la personnalité avec des utilisateurs en interaction avec des agents, il est nécessaire de définir une mesure du regulatory focus adapté à ce besoin. Cet outil doit permettre de mesurer des scores promotion et prévention aussi bien pour mesurer le focus chronique d'un humain que pour le report de la perception d'un utilisateur suite à son interaction avec un agent. Notre seconde contribution, présentée dans le chapitre suivant, sera donc de proposer un questionnaire (validé psychométriquement) permettant de mesurer le regulatory focus en auto-attribution chez l'humain comme en hétéro-attribution pour l'agent.

3.3 RÉSUMÉ DU CHAPITRE

Nous nous concentrons sur la modélisation d'un processus psychologique spécifique : LA RÉGULATION DU SOI. Les structures et les processus concernant le Soi sont souvent mis en avant dans les théories sociocognitives de la personnalité (Cervone, 2005). Le Soi est également un concept central pour l'individualité (Damian and Robins, 2012). Ainsi, nous pensons que les utilisateurs pourraient percevoir plus facilement une personnalité chez un agent artificiel possédant des mécanismes relatifs au Soi.

Pour sous-tendre notre modélisation, nous proposons de nous appuyer sur LA THÉORIE DU REGULATORY FOCUS de Higgins (1997) qui distingue deux types de stratégies de régulation : promotion-focus et prévention-focus. La stratégie promotion-focus met l'accent sur les besoins d'accomplissement et l'analyse des situations en termes de gains/non-gains. La stratégie prévention-focus met en avant les besoins de sécurité et l'analyse des situations en termes de pertes/non-pertes. L'orientation d'un individu vers l'une ou l'autre de ces stratégies peut être dispositionnelle ou induite par la situation.

Lorsqu'il y a adéquation entre le focus (chronique ou situationnel) d'un individu et l'environnement d'une tâche que cet individu réalise, il se produit un effet de REGULATORY FIT. Expérimenter une situation de regulatory fit conduit un individu à éprouver un sentiment de "justesse" dans la réalisation de sa tâche (Camacho et al., 2003) et contribue à augmenter l'engagement de l'individu dans la poursuite de cette tâche (Bianco et al., 2003).

Sur cette base, nous avons créé le MODÈLE PERSEED pour la modélisation computationnel du regulatory focus comme composante de la personnalité. Notre modèle propose un cadre conceptuel et un modèle computationnel. Le cadre conceptuel Perseed propose un ensemble de concepts pour guider les concepteurs dans la caractérisation des éléments utiles lors de la modélisation d'un processus de régulation du Soi. Le modèle computationnel Perseed propose une spécification fonctionnelle pour la mise en place du regulatory focus comme processus de régulation du Soi, sous la forme d'un plug-in destiné à être connecté à une architecture cognitive.

4

MESURE DU REGULATORY FOCUS : LE QUESTIONNAIRE RFQ-PF

Sommaire

4.1 Motivations **50**
- 4.1.1 Limites des questionnaires existants 50
- 4.1.2 Utiliser des proverbes pour mesurer le regulatory focus 51

4.2 Développement de l'échelle **52**
- 4.2.1 Pré-sélection d'items 52
- 4.2.2 Evaluation de la représentativité des items 53

4.3 Première étude de validation **54**
- 4.3.1 Méthode 54
- 4.3.2 Structure factorielle 56
- 4.3.3 Liens avec le RFQ et le PANAS 58

4.4 Seconde étude de validation **61**
- 4.4.1 Méthode 61
- 4.4.2 Structure factorielle 64
- 4.4.3 Validités convergente et discriminante 65
- 4.4.4 Discussion 68

4.5 Résumé du chapitre **70**

Dans ce chapitre, nous présenterons le développement et la validation de notre propre questionnaire interrogeant le regulatory focus, le *Regulatory Focus Questionnaire - Proverbs Form* (RFQ-PF). Dans une première partie, nous commencerons par présenter les motivations sous-tendant notre choix de créer une nouvelle échelle de mesure du regulatory focus (section 4.1). Nous décrirons ensuite notre processus de sélection des items (section 4.2) et une première étude dans laquelle nous examinons la structure factorielle du RFQ-PF (section 4.3). Grâce à cette étude, nous avons créé une version finale du RFQ-PF à 18 items. Enfin, cette version finale a fait l'objet d'une seconde étude pour étudier ses propriétés psychométriques, principalement en termes de validité convergente (section 4.4).

4.1 MOTIVATIONS

Au chapitre 3, nous avons présenté la théorie du regulatory focus à laquelle nous proposons de nous intéresser pour la modélisation d'une composante de la personnalité. Afin d'interroger l'apport de l'implémentation du regulatory focus dans une interaction homme-machine, nous prévoyons de réaliser des expérimentations mettant en interaction un agent virtuel et un utilisateur humain. Pour pouvoir explorer la perception de l'agent par l'utilisateur et l'effet de regulatory fit, il nous faut obtenir une mesure du regulatory focus.

4.1.1 *Limites des questionnaires existants*

Plusieurs questionnaires existent pour mesurer le focus chronique chez l'humain. Deux échelles sont majoritairement utilisées dans la littérature : le *Regulatory Focus Questionnaire* (RFQ ; Higgins et al. (2001)) et le *General Regulatory Focus Measure* (GRFM ; Lockwood et al. (2002)).

Le RFQ propose deux sous-échelles permettant d'évaluer le focus chronique d'un individu par rapport à son histoire personnelle, en termes de succès (subjectifs) passés dans la poursuite de buts promotion et prévention. Le RFQ est plutôt centré sur la conceptualisation des idéaux et des obligations (utilisation des guides de Soi, cf. sous-sous-section 3.1.2.1). Environ la moitié des items font volontairement référence aux interactions parentales. Le RFQ comprend 11 items, 6 pour l'échelle Promotion et 5 pour l'échelle Prévention. D'après Higgins et al. (2001) qui rapporte les résultats (non-publiés) de Harlow et al. (1997), le RFQ a une bonne cohérence interne (échelle promotion : $\alpha = 0,73$; échelle prévention : $\alpha = 0,80$) et présente également une validité convergente et discriminante correcte.

Lockwood et al. (2002) proposent un autre questionnaire : le GRFM, basé sur les mêmes construits théoriques que le RFQ et s'intéressant aux buts qu'un individu poursuit. Le GRFM a été conçu de façon

à évaluer l'orientation promotion et prévention d'étudiants, certains items font référence aux études académiques. Les termes utilisés par Higgins (1997) pour décrire les deux focus sont repris très directement, comme le montre les items suivants : "I see myself as someone who is primarily striving to reach my "ideal self" - to fulfill my hopes, wishes, and aspirations" [1] (promotion), "I see myself as someone who is primarily striving to become the self I "ought" to be - to fulfill my duties, responsibilities, and obligations" [2] (prévention). Le GRFM est constitué de 18 items, 9 pour la sous-échelle promotion et 9 pour la sous-échelle prévention. D'après Lockwood et al. (2002), la cohérence interne est bonne (échelle promotion : $\alpha = 0,81$; échelle prévention $\alpha = 0,75$) et les deux sous-échelles sont légèrement corrélées ($r = 0,17$). Contrairement au RFQ, le GRFM montre des corrélations avec l'affectivité Summerville and Roese (2008).

Cependant, le RFQ et le GRFM ne semblent pas adaptés pour que nous les utilisions en hétéro-attribution, pour caractériser la personnalité perçue d'une entité virtuelle. Le RFQ est clairement tourné sur le passé, et en particulier sur le rapport avec les parents. S'il peut déjà être compliqué pour un adulte de répondre à une question comme "Avez-vous souvent mis vos parents sur les nerfs ?" pour lui-même, cela peut rapidement apparaitre comme très décalé de devoir y répondre pour un agent virtuel. Le GRFM est lui aussi tourné vers les buts profonds qu'une personne (et plus spécifiquement un étudiant) cherche à atteindre, ce qui peut représenter un problème pour l'application à une machine. Et la tournure des items désignent peut-être trop précisément les concepts mesurés. Et par ailleurs, le GRFM est particulièrement tourné sur l'approche des gains (promotion) et l'évitement des pertes (prévention) et délaisse donc certaines subtilités de la théorie du regulatory focus (approche des non-pertes pour le focus prévention et évitement de non-gains pour le focus promotion). De plus, il n'existe pas de de version française validée de ces questionnaires.

4.1.2 *Utiliser des proverbes pour mesurer le regulatory focus*

Un questionnaire néerlandais original, le RFQ-Proverb, a été développé et validé par Van Stekelenburg (2006), questionnaire dans lequel les participants indiquent leur accord avec des proverbes et des dictons. Ce questionnaire comprend 16 items (8 pour chaque focus) et utilise une échelle de Likert à 7 points. Le RFQ-Proverb a été utilisé dans différentes études sur le regulatory focus (Hamstra et al., 2011; Rietzschel, 2011). Selon Hamstra et al. (2011), les proverbes sont an-

1. Je me vois comme quelqu'un qui fait tout son possible pour atteindre mon "moi idéal" - pour satisfaire mes espérances, mes désirs et mes ambitions

2. Je me vois comme quelqu'un qui fait tout son possible pour devenir celui que je "devrais" être - pour remplir mes devoirs, mes responsabilités et mes obligations

crés culturellement et permettent d'évaluer les préférences en termes de stratégies de régulation de façon discrète et subtile. De plus, les proverbes permettent de saisir beaucoup de choses touchant aux buts poursuivis ou aux stratégies mises en place pour les atteindre en une seule et simple phrase.

Les proverbes n'ayant pas vraiment de sens absolu, les répondants à un tel questionnaire saisissent probablement de façon moins évidente le construit mesuré et peuvent plus facilement utiliser leur intuition pour répondre. De plus, la simple modification de la consigne peut suffire à modifier le contexte d'utilisation du questionnaire. Par exemple, Rietzschel (2011) a utilisé le RFQ-Proverb pour évaluer un regulatory focus collectif plutôt qu'individuel.

Ces raisons nous laissent penser que la création d'un tel questionnaire pourrait être particulièrement adapté pour nos expérimentations. En effet, le côté "flou" des proverbes jouera en notre faveur lors de l'utilisation du questionnaire en hétéro-attribution. Il nous parait moins problématique de demander si une machine vous évoque le proverbe "qui ne tente rien n'a rien" plutôt que si elle semble concentrée sur la satisfaction de ses ambitions et de ses désirs.

Du fait de l'ancrage culturel des proverbes et dictons, il ne nous parait pas pertinent d'utiliser une simple traduction du questionnaire néerlandais RFQ-Proverb. C'est pourquoi nous proposons de développer une échelle équivalente en français, mesurant le regulatory focus en utilisant des proverbes : le Regulatory Focus Questionnaire - Proverb Form (RFQ-PF).

4.2 développement de l'échelle

4.2.1 *Pré-sélection d'items*

Pour déterminer les items qui composent le questionnaire, nous avons commencé par créer un large groupe de proverbes évoquant une sensibilité promotion ou prévention. Ce groupe a été constitué à partir d'un site web proposant un inventaire de proverbes en langue française[3]. Les 1778 proverbes recensés par ce site[4] ont été passés au crible pour ne retenir que ceux présentant une validité apparente pour la mesure du regulatory focus. 1689 proverbes ont été éliminés soit parce qu'ils ne présentaient pas de rapport sémantique avec les focus, soit parce qu'ils étaient de pures reformulations de proverbes déjà sélectionnés. Un ensemble de 89 proverbes a ainsi été créé : 48 évoquant le focus prévention et 41 le focus promotion.

3. www.linternaute.com/proverbe
4. En septembre 2013

4.2.2 *Evaluation de la représentativité des items*

Un questionnaire en ligne a ensuite été réalisé pour deux raisons : s'assurer que les proverbes choisis soient catégorisés correctement (comme promotion ou prévention) par une majorité de personnes et sélectionner un nombre réduit d'items parmi cet ensemble. Ce questionnaire a été diffusé à une population tout venant via une liste de diffusion. Chaque proverbe a été présenté avec 16 descriptions (8 par focus ; cf. Tableau 2) d'attitudes et de comportements en rapport avec les deux focus tels que décrits par Higgins. Il était demandé au répondant de cocher les attitudes et les comportements évoqués pour lui par le proverbe affiché. Deux cases à cocher ("Ce proverbe n'est pas pertinent" et "Je ne comprends pas ce proverbe") ont été ajoutées pour s'assurer de la clarté et de la compréhensibilité des items. Les 89 proverbes pré-sélectionnés étaient présentés dans un ordre aléatoire.

Tableau 2 – Attitudes et comportements spécifiques aux focus promotion et prévention utilisés pour la catégorisation des proverbes pré-sélectionnés pour le RFQ-PF

Type	Promotion	Prévention
Buts (x2)	Chercher à faire ce que l'on veut	Chercher à faire ce que l'on doit
	Avoir de l'ambition	Avoir le sens des responsabilités
Besoins	Prendre des risques	Veiller à sa sûreté
Evaluation	Evaluer les situations en termes de gains / non-gains (Réussir, c'est gagner quelque chose ; Echouer, c'est ne rien gagner)	Evaluer les situations en termes de pertes / non-pertes (Réussir, c'est ne rien perdre ; Echouer, c'est perdre quelque chose)
Conséquences	Chercher à obtenir des conséquences positives	Chercher à éviter des conséquences négatives
Style	Etre ouvert au changement	Assurer une stabilité
Actions (x2)	S'assurer d'agir quand il le faut	S'assurer de ne pas agir à tort
	Eviter de ne pas agir alors qu'il le faudrait	Eviter d'agir à tort

78 personnes (21 hommes et 57 femmes), âgés de 20 à 78 ans (M âge = 44,9 ; SD age = 17,3), ont participé à cette étude de catégorisation. 9 d'entre elles ont été exclues des analyses car leurs réponses ont été jugées incohérentes.

Pour chaque proverbe, la pertinence a été calculée comme le pourcentage des répondants trouvant le proverbe en accord avec un focus particulier. Pour un répondant, un proverbe est considéré comme en accord avec le focus promotion (resp. le focus prévention) dans deux cas : 1/ seules des catégories promotion (resp. prévention) ont été cochées pour ce proverbe ou 2/ il y a au moins deux catégories promotion (resp. prévention) de plus de cochées par rapport au nombre de catégories prévention (resp. promotion) cochées.

Un ensemble de 40 proverbes (20 pour chaque focus) a été déterminé comme contenant les proverbes les plus pertinents selon notre échantillon : la plage de pertinence pour les items promotion va de 100% à 61% ; la plage de pertinence pour les items prévention va de 86% à 58%.

Ces 40 items ont été soumis à la validation de deux experts (un homme et une femme, agrégés en littérature) pour qu'ils jugent de la concordance entre les proverbes et les focus définis. Ces experts ont évalué la pertinence de chaque proverbe par rapport au focus associé sur une échelle de Likert à 4 points ("Non pertinent", "Plutôt non pertinent", "Plutôt pertinent", "Pertinent"). Les experts pouvaient également donner des commentaires sur les items qui ne leur paraissaient pas suffisamment clairs. Sur la base de l'évaluation et des commentaires, seuls les proverbes faisant consensus entre les experts ont été retenus. Les 40 proverbes évalués ainsi leur pertinence pour le tout-venant et pour les experts sont rapportés en Annexe A.

En suivant ce processus, 20 proverbes (10 items promotion ; 10 items prévention) ont finalement été sélectionnés pour constituer la première version de l'échelle RFQ-PF.

4.3 PREMIÈRE ÉTUDE DE VALIDATION

Une première étude de validation sur deux échantillons a été conduite pour évaluer les propriétés psychométriques du RFQ-PF, en particulier sa fiabilité et sa structure factorielle. Une première exploration des validités convergente et discriminante a également été mise en place.

4.3.1 *Méthode*

4.3.1.1 *Matériel*

Les questionnaires utilisés dans cette étude étaient les suivants :

RFQ-PF La version préliminaire à 20 items du RFQ-PF a été utilisée avec la consigne suivante : "Pour chaque proverbe, indiquez dans

quelle mesure celui-ci peut s'appliquer à votre façon de voir la vie en général". Les réponses se faisaient sur une échelle de Likert à 7 points allant de 1 ("Pas du tout") à 7 ("Tout à fait").

RFQ Afin de commencer à établir la validité convergente du RFQ-PF, une version française du RFQ (Higgins et al. (2001) ; version française : Angel and Steiner (2010)) a été choisie. Celle-ci est disponible en Annexe A. Le questionnaire comporte 11 items permettant de calculer un score promotion et un score prévention. Les réponses se font sur une échelle de Likert à 5 points. Les deux échelles sont fiables (échelle promotion : $\alpha = 0,79$; échelle prévention : $\alpha = 0,81$).

PANAS La version dispositionnelle du PANAS (Positive and Negative Affect Schedule Watson et al. (1988) ; version française : Gaudreau et al. (2006)) a été choisie pour examiner la validité discriminante du RFQ-PF. Théoriquement, les focus promotion et prévention sont définis comme indépendants de la valence des affects. Les personnes ressentent toutes des émotions positives et négatives, mais ces émotions sont de différents types en fonction du focus chronique de la personne (par exemple, pour les émotions positives, un focus promotion entraine plutôt de la joie alors qu'un focus prévention entraine plutôt du calme).

Ce questionnaire de 20 items permet de mesurer la disposition d'une personne à ressentir certains types d'affects avec deux échelles, une pour les affects positifs et une pour les affects négatifs. Les items correspondent à des adjectifs représentant des affects (e. g. intéressé, angoissé, excité...) et le répondant doit indiquer à quelle fréquence il ressent ce sentiment ou cette émotion. Les réponses se font sur une échelle de Likert à 5 points allant de 1 ("Jamais") à 5 ("Très souvent"). Les deux échelles sont fiables (échelle affect positif : $\alpha = 0,88$; échelle affect négatif : $\alpha = 0,87$).

4.3.1.2 *Participants*

343 personnes (83 hommes et 260 femmes), âgées de 18 à 78 ans (M âge = 34,2 ; SD âge = 15,2) ont participé à cette étude. Les participants étaient répartis en deux échantillons qui ont répondu à plusieurs questionnaires. L'échantillon A était composé de 184 personnes (47 hommes et 137 femmes), âgées de 18 à 78 ans (M âge = 34,4 ; SD âge = 15,0) et l'échantillon B de 159 personnes (36 hommes et 123 femmes), âgées de 18 à 71 ans (M âge = 33,9 ; SD âge = 15,5). Tous les participants ont été contactés via une liste de diffusion.

4.3.1.3 *Procédure*

Les questionnaires étaient remplis en ligne par les participants. Sur la page de présentation, il était donné des informations sur les ob-

Tableau 3 – Règles empiriques d'interprétation des indicateurs d'adéquation du modèle produit par l'analyse factorielle exploratoire

	Bonne adéquation	Adéquation acceptable
RMSEA	< 0,03	compris entre 0,03 et 0,07
TLI	> 0,95	compris entre 0,95 et 0,80

D'après Hooper et al. (2008)

jectifs de l'étude ainsi que sur la procédure et la confidentialité des données. Il était demandé aux participants d'être aussi honnête que possible et de répondre à l'étude individuellement et dans un environnement calme.

Après avoir donné ses informations démographiques, chaque participant répondait à deux questionnaires : 1/ la version préliminaire du RFQ-PF et 2/ le questionnaire RFQ pour l'échantillon A ou le PANAS pour l'échantillon B.

Pour l'analyse des données, toutes les statistiques ont été réalisées avec R version 3.2. Les analyses factorielles, les alphas de Cronbach et les corrélations ont été réalisés avec le package "psych", version 1.5.8.

4.3.2 *Structure factorielle*

Pour étudier la structure factorielle du RFQ-PF, les données des échantillons A et B ont été réunies. En suivant les recommandations de Schmitt (2011) et Courtney and Gordon (2013), une analyse factorielle exploratoire (Exploratory Factor Analysis ou EFA) avec rotation oblimin [5] a été réalisée sur les 20 items de la version préliminaire du RFQ-PF, avec la factorisation en axe principal comme méthode d'extraction [6].

À cette étape, deux items de l'échelle Promotion ont été supprimés car ils montraient des saturations croisées. La même analyse a été reconduite sur les 18 items restants. Le nombre de facteurs à extraire a été déterminé via l'analyse parallèle de Horn (avec une méthode d'extraction de factorisation en axe principal également). Cette analyse suggère une structure en 3 facteurs (eigenvalues = 2,82, 2,45, 0,56). Pour évaluer l'adéquation du modèle, nous avons observé deux indicateurs : 1/ le RMSEA (Root Mean Square Error of Approximation),

5. La rotation oblimin autorise la corrélation entre les facteurs

6. Cette méthode est recommandée quand le postulat de la normalité de la distribution multivariée des items n'est pas respecté. Ce postulat n'est pas respecté dans le cas de données issues d'échelles de Likert, même si les coefficients de dissymétrie (*skewness*) et les excès d'aplatissement (*kurtosis* - 3) de nos données sont toujours compris entre -1 et 1 (plage pour les coefficients de dissymétrie : de -0,7 à 0,1 ; plage pour les excès d'aplatissement : de -0,9 à 0,44)

un indicateur absolu et son intervalle de confiance à 90% (IC 90%) et 2/ le Tucker-Lewis Index (TLI ou NNFI), un indicateur relatif. Les seuils déterminant l'interprétation de ces indicateurs sont présentés dans le Tableau 3.

La solution à 3 facteurs montre une adéquation acceptable (RMSEA = 0,06 avec IC 90% = [0,048 ; 0,069] ; TLI = 0,87). Les statistiques descriptives pour les différents items et les saturations sont reportées dans le Tableau 4.

Le premier facteur (F1) réunissant les items conçus comme prévention est donc désigné comme le facteur "prévention focus". Le

Tableau 4 – Moyennes (M) et écarts-types (SD) par items du RFQ-PF ; Facteurs (Fx) avec saturation des items et pourcentage de variance expliquée

Items	M	SD	F1	F2	F3
(Pre-02) Il ne faut pas quitter le certain pour l'incertain	3,81	1,67	0,54		
(Pre-03) Qui veut voyager loin ménage sa monture	4,80	1,49	0,63		
(Pre-04) Il vaut mieux tenir que courir	4,59	1,49	0,60		
(Pre-06) Dans le doute abstiens-toi	4,36	1,65	0,47		
(Pre-08) Un tiens vaut mieux que deux tu l'auras	4,78	1,45	0,48		
(Pre-09) On sait ce qu'on quitte mais on ne sait pas ce qu'on trouve	5,12	1,60	0,42		
(Pre-11) Il faut garder une poire pour la soif	4,56	1,45	0,44		
(Pre-13) Mieux vaut prévenir que guérir	5,52	1,30	0,50		
(Pre-16) Prudence est mère de sûreté	4,47	1,28	0,62		
(Pre-17) Il ne faut pas mettre tous ses œufs dans le même panier	5,40	1,24	0,43		
(Pro-01) Qui ne tente rien n'a rien	5,46	1,36		0,59	
(Pro-10) À cœur vaillant, rien d'impossible	5,07	1,58		0,64	
(Pro-14) La fortune sourit aux audacieux	4,73	1,49		0,50	
(Pro-18) Quand on veut, on peut	5,10	1,58		0,66	
(Pro-05) Il faut prendre la balle au bond	4,90	1,40			0,48
(Pro-07) Trop de prudence n'atteint pas son but	4,46	1,52			0,67
(Pro-12) Il faut faire tourner le moulin lorsque le vent souffle	5,25	1,31			0,44
(Pro-15) Trop de précautions nuit	4,35	1,45			0,63
Pourcentage de variance expliquée			16%	10%	10%

second facteur (F2) est interprété comme un facteur "promotion focus - approche" car il regroupe 4 items qui soulignent des comportements d'approche vers un gain. Le troisième facteur (F3) est interprété comme un facteur "promotion focus - omission" car il regroupe 4 items qui mettent en garde contre les erreurs d'omission. Les deux facteurs "promotion" montrent une corrélation substantielle de 0,46. Aucun des facteurs "promotion" ne présente de corrélation avec le facteur prévention.

Pris dans leur ensemble, les trois facteurs expliquent 36% de la variance. Le facteur "prévention focus" explique 16% de la variance. Les facteurs "promotion focus - approche" et "promotion focus - omission" expliquent chacun 10% de la variance. Aucun des items conçus comme prévention ne sature sur les facteurs promotion et aucun des items conçus comme promotion ne sature sur le facteur prévention.

Nous pensons que les scores des deux facteurs "promotion" peuvent être moyennés et que l'ensemble des items promotion peuvent être considérés comme une seule échelle. Cette préconisation se base sur le fait que, d'une part, l'approche vers un gain et l'attention portée aux erreurs d'omission sont deux composantes du concept du focus promotion et que d'autre part, les deux facteurs montrent une corrélation honnête. Ce choix sera discuté plus avant par l'étude des corrélations avec le RFQ. Les deux échelles étaient fiables : pour l'échelle promotion, l'alpha est de 0,78 (intervalle de confiance à 95% = [0,73 ; 0,83]) et pour l'échelle prévention, l'alpha est de 0,77 (intervalle de confiance à 95% = [0,72 ; 0,82]).

Dans l'ensemble, ces résultats soulignent la fiabilité du RFQ-PF. Les résultats de l'analyse factorielle exploratoire s'inscrivent dans la suite des résultats de Higgins et al. (2001) qui rapporte deux facteurs avec une légère corrélation ($r = 0,21$, $p < 0,001$) qui expliquent 50% de la variance (respectivement 29% et 21%). De plus, nos résultats sont en cohérence avec Hamstra et al. (2011) qui utilisent également un questionnaire basé sur des proverbes.

4.3.3 *Liens avec le RFQ et le PANAS*

Les validités convergentes et discriminantes ont été approchées par l'analyse des corrélations entre les scores du RFQ-PF et ceux du RFQ (validité convergente) et du PANAS (validité discriminante).

D'après Higgins et al. (2001), les participants peuvent être classés comme ayant un focus chronique promotion ou prévention en observant la différence entre leur score sur l'échelle promotion et leur score sur l'échelle prévention du RFQ (i. e. strictement supérieur à 0 équivaut à promotion ; inférieur ou égal à 0 équivaut à prévention).

Tableau 5 – Corrélations de Spearman entre les scores du RFQ-PF et les scores du RFQ et du PANAS

	RFQ-Pro	RFQ-Pre	RFQ-Inc	PA	NA
PF-Pro	0,32	-0,27	0,38	0,47	*ns*
PF-Pre	-0,18	0,32	-0,32	*ns*	0,21
PF-Inc	0,34	-0,43	0,50	0,43	-0,24

Tous les résultats reportés sont significatifs à $p <= 0,05$ *; ns = non significatif*

En conséquence, trois scores ont été calculés pour le RFQ et pour le RFQ-PF : un score Promotion (resp. RFQ-Pro et PF-Pro), un score Prévention (resp. RFQ-Pre et PF-Pre) et ce que nous appelons l'inclination (soit score Promotion - score Prévention ; resp. RFQ-Inc et PF-Inc). Pour le PANAS, deux scores ont été calculés : un score pour les affects positifs (PA) et un score pour les affects négatifs (NA). La relation entre le RFQ-PF et ces deux autres questionnaires est étudiée au travers de corrélations de Spearman.

Les corrélations entre les différents scores obtenus sont présentées dans le Tableau 5.

4.3.3.1 *Liens RFQ et RFQ-PF*

Concernant le RFQ, nous avons trouvé des corrélations positives moyennes entre les scores Promotion du RFQ et du RFQ-PF ($\rho = 0,32$; $p < 0,01$) ainsi qu'entre les scores Prévention ($\rho = 0,32$; $p < 0,01$). Une forte corrélation a été trouvée entre les scores d'inclination du RFQ et du RFQ-PF ($\rho = 0,50$; $p < 0,01$). D'après ces résultats, ils semblent que les échelles Prévention et Promotion du RFQ et du RFQ-PF ne mesurent pas exactement le même concept. Mais le RFQ-PF semble néanmois être un bon indicateur pour la classification du focus chronique.

Pour étudier plus en détails les différences entre le RFQ et le RFQ-PF, des sous-scores PF-Pro-Approche et PF-Pro-Omission ont été calculés pour le RFQ-PF. Ces sous-scores correspondent aux facteurs "Promotion-Approche" et "Promotion-Omission" mis en évidence par l'analyse factorielle. Les corrélations entre ces sous-scores et les scores du RFQ-PF sont présentés dans le Tableau 6.

La corrélation entre le score PF-Pro-Approche et le score RFQ-Pro ($\rho = 0,39$; $p < 0,01$) est plus importante que celle entre le score PF-Pro-Omission et le score RFQ-Pro ($\rho = 0,19$; $p < 0,05$). Le RFQ est centré sur les succès passés d'un individu lors de la mise en œuvre de stratégies promotion ou prévention pour atteindre ses buts (Higgins

Tableau 6 – Corrélations de Spearman entre les sous-scores Promotion du RFQ-PF et les scores du RFQ

	RFQ-Pro	RFQ-Pre	RFQ-Inc
PF-Pro-Approche	0,39	*ns*	0,31
PF-Pro-Omission	0,19	-0,33	0,36

Tous les résultats reportés sont significatifs à $p <= 0,05$ *; ns = non significatif*

et al., 2001). Il semble donc logique que le score d'approche du RFQ-PF soit plus fortement corrélé avec le score Promotion du RFQ.

Ces résultats mettent également en avant un autre point intéressant : la corrélation moyenne entre le score d'inclination du RFQ et le sous-score PF-Pro-Omission ($\rho = 0,36$; $p < 0.01$). D'après Higgins (1997), les individus promotion-focus s'assurent contre les erreurs d'omission (faux négatif), alors que les individus prévention-focus seront plus sensibles à éviter les erreurs de comission (faux positif). Nous pouvons donc penser que plus la différence entre le score Promotion et le score Prévention est importante, plus les individus notent les mises en garde contre les erreurs d'omission comme signifiantes dans leur vision de la vie.

La corrélation négative entre le score RFQ-Pre et le sous-score PF-Pro-Omission ($\rho = -0,33$; $p < 0.01$) peut être interprétée de la même manière : plus les individus ont eu des succès passés en utilisant une stratégie prévention, moins les mises en garde contre les erreurs d'omission leur paraissent utiles.

Les relations entre le RFQ et les sous-scores Promotion du RFQ-PF, ajoutées aux fondements théoriques, suggèrent qu'il est cohérent d'utiliser l'échelle Promotion comme un seul score, en particulier si le but est de catégoriser les individus comme étant orientés promotion ou prévention.

4.3.3.2 *Liens PANAS et RFQ-PF*

Concernant le PANAS, le score d'affect positif est fortement corrélé au score Promotion du RFQ-PF ($\rho = 0,47$; $p < 0,01$) mais ne montre pas de corrélation avec le score Prévention. Le score d'affect négatif montre une faible corrélation avec le score Prévention du RFQ-PF ($\rho = 0,21$; $p < 0,05$) mais aucune avec le score Promotion.

Ces résultats peuvent être discutés à la lumière des résultats de Summerville and Roese (2008). En effet, dans une étude s'intéressant à l'indépendance théorique du regulatory focus et de la valence des émotions ressenties, Summerville and Roese (2008) ont montré

que l'échelle Promotion du questionnaire GRFM était corrélée avec l'échelle d'affect positif du PANAS ($r = 0,31$; $p < 0,05$) et que l'échelle Prévention du GRFM était corrélée avec l'échelle d'affect négatif du PANAS ($r = 0,39$; $p = 0,04$). La même étude reporte n'avoir trouvé aucune corrélation entre les scores du RFQ et ceux du PANAS.

Bien qu'il soit difficile de savoir si cet effet est dû à des biais propres aux questionnaires (dûs aux rédacteurs comme aux répondants), il parait important de souligner que la mesure du regulatory focus fournie par le RFQ-PF, comme celle du GRFM, peut être liée à des variations en termes de sensibilité affective. Cette association soulève des questions concernant le fait d'utiliser des questionnaires en auto-attribution comme mesure principale du regulatory focus et les conséquences conceptuelles et méthodologiques doivent être discutées.

4.4 SECONDE ÉTUDE DE VALIDATION

La seconde étude de validation a trois objectifs : 1/ confirmer la structure factorielle du RFQ-PF avec une analyse factorielle confirmatoire (Confirmatory Factor Analysis ou CFA), 2/ étudier plus avant les validités convergentes et discriminantes du RFQ-PF en examinant ses liens avec plusieurs échelles de motivation et de régulation du Soi et 3/ discuter l'incidence de ces liens sur le concept de regulatory focus.

4.4.1 *Méthode*

4.4.1.1 *Matériel*

Les questionnaires utilisés dans cette étude était les suivants :

RFQ-PF La version finale à 18 items du RFQ-PF (disponible en Annexe A) a été utilisée avec la consigne suivante : "Pour chaque proverbe, indiquez dans quelle mesure celui-ci peut s'appliquer à votre façon de voir la vie en général". Les réponses se faisaient sur une échelle de Likert à 7 points allant de 1 ("Pas du tout") à 7 ("Tout à fait").

RFSS Une version traduite du questionnaire RFSS (Regulatory Focus Strategies Scale ; Ouschan et al. (2007)) a été utilisée. Ce questionnaire mesure le regulatory focus, non pas en s'intéressant aux buts poursuivis en tant que tels, mais en évaluant l'adhésion d'un individu à des stratégies représentatives des focus. Ce questionnaire comporte donc 14 items destinés à mesurer l'utilité perçue de stratégies de régulation promotion ou prévention lors de la poursuite d'un but. Une échelle de Likert à 5 points est utilisée pour les réponses.

rses De nombreuses études s'intéressent aux liens entre l'estime de soi et la régulation du Soi (Leonardelli et al., 2007). L'estime de soi est également reliée à la différence entre le soi idéal et le soi actuel, parfois utilisée comme une mesure du focus chronique promotion (Higgins and Tykocinski, 1992). Les personnes présentant un focus chronique promotion devraient donc montrer un score en estime de soi plus élevé. Une version française validée du RSES (Rosenberg's Self-Esteem Scale) a été utilisée (Vallieres and Vallerand, 1990). Ce questionnaire comprend 10 items et utilise une échelle de Likert à 4 points pour les réponses.

bis/bas Le modèle de renforcement de Gray (1990) postule deux systèmes motivationnels jouant sur les comportements et les émotions : un système d'évitement (Behavioral Inhibition System ou BIS) et un système d'approche (Behavioral Activation System ou BAS). Le questionnaire BIS/BAS mesure la sensibilité des individus par rapport à ces deux systèmes. Le score BIS, mesuré par des items faisant référence aux punitions et aux non-gains, devrait être relié au score Prévention. Le BAS est mesuré par trois sous-scores : BAS-Drive, associé à la persévérance dans la poursuite des buts désirés ; BAS-Fun Seeking, par rapport à la sensibilité aux évènements plaisants et aux comportements d'approche associés ; BAS-Reward responsiveness, associé aux réponses positives liées à l'anticipation ou à l'obtention de récompenses. Les sous-scores BAS devraient être reliés au score Promotion. Une version française validée du BIS/BAS a été utilisée (Caci et al., 2007). Ce questionnaire comprend 24 items et utilise une échelle de Likert à 4 points pour les réponses.

spsrq Dans le même esprit que le BIS/BAS, la sensibilité à la récompense devrait être reliée aux aspects promotion alors que la sensibilité à la punition devrait être reliée aux aspects prévention. Les participants ont répondu à une version française validée du SPSRQ (Sensitivity to Punishment and Sensitivity to Reward Questionnaire ; Lardi et al. (2008)). Ce questionnaire comprend 35 items et utilise une échelle de Likert à 4 points.

lot-r L'optimisme correspond à la croyance d'un individu concernant la probabilité d'obtenir des résultats positifs dans le futur. Le questionnaire LOT-R (Life Orientation Test - Revised) permet de mesurer deux sous-scores : l'affirmation de l'optimisme et la négation du pessimisme. Ces deux sous-scores peuvent être additionnés pour obtenir une mesure globale d'optimisme. Le focus promotion étant plus attentif aux conséquences positives, le score Promotion devrait être relié à l'optimisme. À l'inverse, le score Prévention pourrait être relié négativement à l'optimisme car se projeter vers des résultats négatifs (pour mieux les éviter) peut servir la stratégie vigilante caratéristique

du focus prévention. Une version française validée du LOT-R a été utilisée (Sultan and Bureau, 1999). Ce questionnaire comprend 10 items et utilise une échelle de Likert à 5 points.

BFI-FR D'après les définitions théoriques des focus promotion et prévention (Higgins, 1997) et la description des 5 dimensions de la personnalité proposées par le Big Five (Costa and McCrae, 1992a), les hypothèses suivantes peuvent être faites : l'extraversion, en rapport avec l'enthousiasme et l'excitation, et l'ouverture à l'expérience, en rapport avec l'attrait pour la nouveauté, pourraient être reliée au focus promotion. À l'inverse, le névrosisme, en rapport avec l'anxiété et l'agitation, pourrait être relié au focus prévention. Les cinq dimensions du "Big Five" ont été mesurées à l'aide d'une version française validée du BFI (Big Five Inventory - French ; Plaisant et al. (2010)). Ce questionnaire comprend 45 items et utilise une échelle de Likert à 5 points.

4.4.1.2 *Participants*

277 personnes (74 hommes et 203 femmes), âgées de 17 à 78 ans (M âge = 35,4 ; SD âge = 15,2) ont participé à cette étude. Tous les participants ont été contactés via une liste de diffusion. 15 personnes ont été exclues de l'analyse des résultats après que leurs réponses aient été jugées incohérentes (e. g. un questionnaire entier rempli avec la même valeur).

4.4.1.3 *Procédure*

Les questionnaires étaient remplis en ligne par les participants. Sur la page de présentation, il était donné des informations sur les objectifs de l'étude ainsi que sur la procédure et la confidentialité des données. Il était demandé aux participants d'être aussi honnête que possible et de répondre à l'étude individuellement et dans un environnement calme.

Après avoir donné ses informations démographiques, chaque participant répondait aux 7 questionnaires présentés ci-dessus dans un ordre aléatoire : 1/ le RFQ-PF, 2/ le RFSS, 3/ le RSES, 4/le BIS/BAS , 5/ le SPSRQ, 6/ le LOT-R et 7/ le BFI-Fr.

Les données ont ensuite été analysées avec R, version 3.2.4. Les analyses factorielles confirmatoires ont été réalisées avec le package "lavaan", version 0.5-20, et les corrélations ont été calculées avec le package "psych", version 1.5.8.

Tableau 7 – Règles empiriques d'interprétation des indicateurs d'adéquation des modèles produits par les analyses factorielles confirmatoires

	Bonne adéquation	Adéquation acceptable
χ^2 normalisé	< 3	< 5
RMSEA	< 0,03	compris entre 0,03 et 0,07
SRMR	< 0,05	compris entre 0,05 et 0,10
CFI et TLI	> 0,95	compris entre 0,95 et 0,80

D'après Hooper et al. (2008)

4.4.2 *Structure factorielle*

Deux analyses factorielles confirmatoires ont été réalisées pour tester l'adéquation du modèle en 3 facteurs (trouvé dans l'étude précédente) en comparaison du modèle alternatif en 2 facteurs (basé sur la théorie). Avant de lancer les analyses, les coefficients de dissymétrie (*skewness*) et les excès d'aplatissement (*kurtosis* - 3) ont été calculés. Tous se sont révélés être entre -1 et 1 mais de par la nature de nos données (issues d'échelles de Likert), l'hypothèse de normalité (univariée ou multivariée) ne peut être assurée. De ce fait, en suivant In'nami and Koizumi (2013), les CFAs ont été réalisées avec l'estimation du maximum de vraisemblance comme méthode d'inférence pour les paramètres (*Maximum Likelihood parameter estimation method*), avec des erreurs standard robustes et une correction de Satorra-Bentler.

Les indicateurs suivants ont été choisis pour évaluer l'adéquation des modèles par rapport aux données : 1/ le χ^2 normalisé (χ^2/df; df = degrés de liberté), 2/ le RMSEA (Root Mean Square Error of Approximation), et son intervalle de confiance à 90% (IC 90%), 3/ le SRMR (Standardized Root Mean square Residual), 4/ le CFI (Comparative Fit Index) et 5/ le TLI (Tucker-Lewis Index). Les seuils déterminant l'interprétation de ces indicateurs sont présentés dans le Tableau 7.

Les indicateurs d'adéquation des deux CFA réalisées sont présentés en Tableau 8. D'après nos critères d'acceptation, les résultats montrent une meilleure adéquation du modèle à 3 facteurs en comparaison du modèle en 2 facteurs. Même si le modèle en 3 facteurs n'a pas une excellente adéquation, les indicateurs permettent de conclure à une erreur d'approximation raisonnable et donc à une adéquation acceptable.

Tableau 8 – Indicateurs d'adéquation pour les modèles en 2 et 3 facteurs concernant le questionnaire RFQ-PF à 18 items

Modèle	2 facteurs	3 facteurs
χ^2	304	282
df	134	132
χ^2 normalisé	2,27	2,14
RMSEA [IC 90%]	0,070 [0,060 ; 0,079]	0,066 [0,056 ; 0,075]
SRMR	0,087	0,085
CFI	0,78	0,81
TLI	0,75	0,78

4.4.3 *Validités convergente et discriminante*

Trois scores ont été calculés pour le RFQ-PF : un score promotion (PF-Pro), un score prévention (PF-Pre) et l'inclination (PF-Inc = PF-Pro - PF-Pre). Les corrélations entre ces scores et les autres mesures choisies sont présentées dans le Tableau 9.

4.4.3.1 *Corrélation avec le RFSS*

Comme attendu, les scores Promotion et Prévention du RFSS corrèlent plutôt fortement avec les scores PF-Pro et PF-Pre du RFQ-PF (resp. $\rho = 0,47$, $p < 0,001$ et $\rho = 0,45$, $p < 0,001$). De façon logique par rapport à l'absence d'association entre les échelles Promotion et Prévention, le score Promotion (et respectivement Prévention) du RFSS a seulement une légère corrélation négative avec le score PF-Pre (et respectivement PF-Pro) du RFQ-PF. Concernant l'inclination, les scores du RFSS et du RFQ-PF montrent une forte corrélation ($\rho = 0,59$, $p < 0,001$).

4.4.3.2 *Corrélation avec le RSES*

Un motif complémentaire au précédent apparaît avec les corrélations relatives à l'estime de soi. Le score Promotion montre une corrélation modérée avec le score d'estime de soi ($\rho = 0,35$, $p < 0,001$), ainsi qu'une légère corrélation négative avec le score Prévention ($\rho = -0,21$, $p < 0,001$). Ces résultats vont dans le même sens que les résultats de Ouschan et al. (2007) qui montrent une forte corrélation entre l'estime de soi et le score promotion du RFQ de Higgins et une relative indépendance par rapport aux scores du RFSS.

4.4.3.3 *Corrélation avec le LOT-R*

Les corrélations entre le questionnaire sur l'optimisme et le RFQ-PF sont modérées. L'optimisme global corrèle positivement avec le

Tableau 9 – Corrélations de Spearman entre les scores du RFQ-PF et les scores du BFI-Fr, BIS/BAS, LOT-R, RFSS, RSES et SPSRQ

	PF-Pro	PF-Pre	PF-Inc
Ouverture à l'expérience	0,30	-0,29	0,40
Caractère consciencieux	*ns*	*ns*	*ns*
Extraversion	0,21	-0,17	0,25
Agréabilité	*ns*	*ns*	*ns*
Névrosisme	-0,21	0,26	-0,32
BIS	-0,27	0,30	-0,36
BAS-Drive	0,30	*ns*	0,25
BAS-Fun	0,30	-0,30	0,41
BAS-Reward	0,21	*ns*	0,17
Affirmation de l'optimisme	0,43	-0,27	0,48
Négation du pessimisme	0,22	-0,30	0,35
Optimisme global	0,35	-0,32	0,46
RFSS-Pro	0,47	-0,24	0,50
RFSS-Pre	-0,18	0,45	-0,43
RFSS-Inc	0,39	-0,44	0,59
Estime de soi	0,35	-0,21	0,37
Sensibilité aux récompenses	0,22	-0,13	0,24
Sensibilité aux punitions	-0,32	0,25	-0,36
Sensibilité-Inc (récompense – punition)	0,36	-0,28	0,42

Tous les résultats reportés sont significatifs à $p <= 0,05$ *; ns = non significatif*

score PF-Pro ($\rho = 0,35$, $p < 0,001$) et négativement avec le score PF-Pre ($\rho = -0,32$, $p < 0,001$). De façon intéressante, le score PF-Pro a une corrélation plus forte avec l'affirmation de l'optimisme ($\rho = 0,43$, $p < 0,001$) qu'avec la négation du pessimisme ($\rho = 0,22$, $p < 0,001$). Ces résultats en cohérence avec ceux de Grant and Higgins (2003) qui utilise le RFQ mais distincts des résultats avec le RFSS présentés par Ouschan et al. (2007).

4.4.3.4 *Corrélation avec le BIS/BAS*

Les corrélations entre le RFQ-PF et le BIS/BAS ont également la forme attendue : le score PF-Pro est modérément corrélé de façon positive aux trois scores BAS et de façon négative au score BIS ; le score PF-Pre est modérément corrélé de façon positive au score BIS et de façon négative au score BAS-Fun Seeking. Les corrélations entre le RFQ-PF et le BIS/BAS sont plutôt modestes. Mais elles sont d'une amplitude équivalente aux corrélations présentées par Ouschan et al. (2007) pour le RFSS. De plus, Harlow et al. (1997) (cité par Higgins et al. (2001)) rapportent le même type de corrélations pour le RFQ, soit une relation positive entre un fort score Promotion et les échelles BAS-Reward responsiveness et BAS-Fun Seeking et une relation négative entre un fort score Prévention et le score BAS-Fun Seeking.

4.4.3.5 *Corrélation avec le SPSRQ*

Les corrélations entre le RFQ-PF et le SPSRQ sont modestes mais encore une fois, du même ordre que les résultats de Ouschan et al. (2007). Le score RF-Pro est corrélé positivement avec la sensibilité à la récompense et négativement avec la sensibilité à la punition. Inversement, le score RF-Pre est corrélé négativement à la sensibilité à la récompense et positivement avec la sensibilité à la punition. Ces résultats se distinguent néanmoins de ceux d'Ouschan et al. (2007) qui ne trouve pas de lien, ni entre score Promotion du RFSS et sensibilité à la punition, ni entre score Prévention du RFSS et sensibilité à la récompense.

4.4.3.6 *Corrélation avec le BFI-Fr*

Les corrélations entre le RFQ-PF et le BFI-Fr se sont également montrées intéressantes. Comme prévu, le score RF-Pro corrèle positivement avec l'extraversion et l'ouverture à l'expérience. Le score RF-Pro montre également une corrélation négative avec le névrosisme. Les corrélations pour le score RF-Pre offrent un motif inverse. Aucune corrélation significative n'a été trouvée avec l'agréabilité et le caractère consciencieux ni pour le score RF-Pro, ni pour le score RF-Pre.

4.4.4 *Discussion*

Dans l'ensemble, la version finale du RFQ-PF (18 items) montrent de bonnes qualités psychométriques.

Le RFQ-PF montre une structure factorielle à 3 facteurs : un facteur Prévention, un facteur Promotion-Approche (en rapport avec l'efficacité des stratégies d'approche dans la poursuite d'un but) et un facteur Promotion-Omission (en rapport avec des mises en garde contre les erreurs d'omission). Si une structure à 2 facteurs (Promotion et Prévention) était théoriquement attendue, les deux facteurs Promotion montrent une corrélation relativement importante et représentent tous deux des dimensions théoriquement reliées au concept de promotion-focus. Nous pensons qu'il est raisonnable d'utiliser l'ensemble des items de ces deux facteurs comme une seule échelle Promotion. Les deux échelles présentent d'ailleurs une bonne cohérence interne ($\alpha_{\text{Promotion}} = 0,78$; $\alpha_{\text{Prévention}} = 0,77$).

Les validités convergentes et discriminantes ont été étudiées à l'aide de plusieurs questionnaires de régulation du Soi et de motivations. Si les corrélations se sont révélées globalement modestes, les résultats vont cependant dans le sens attendu. Les corrélations les plus fortes se retrouvent entre le RFQ-PF et les deux autres questionnaires de regulatory focus, le RFQ (Higgins et al., 2001) et le RFSS (Ouschan et al., 2007).

Cependant, les études de validation précédentes d'autres questionnaires de regulatory focus sont parfois imprécises. Par exemple, Higgins et al. (2001) rapporte des résultats concernant la validation du RFQ mais l'étude elle-même n'est pas publiée. De la même façon, les seuls indicateurs d'adéquation de leur modèle en 2 facteurs (calculés pour une analyse factorielle confirmatoire) sont le GFI (Goodness-of-Fit Index) et le AGFI (Adjusted Goodness-of-Fit Index). Bien qu'historiquement utilisés, ces indicateurs présentent de forts biais et sont parfois même déconseillés (Hooper et al., 2008).

Il est intéressant de voir que les corrélations pour le score d'inclination étaient généralement les plus importantes. L'inclination caractérise l'aspect marqué de la sensibilité du répondant au focus promotion ou prévention. Plus l'inclination est haute, plus le répondant est sensible au focus promotion ; plus l'inclination est basse, plus le répondant est sensible au focus prévention. Par exemple, le score d'inclination du RFQ-PF présente une corrélation positive avec l'ouverture à l'expérience ($\rho = 0,40$) et une corrélation négative avec le score de névrosisme ($\rho = -0,32$). Cela signifie que plus l'inclination du répondant est haute (i. e. il y a une grande différence entre le score PF-Pro et le score PF-Pre), plus son score d'ouverture est haut et plus son score de névrosisme est bas.

Nos résultats suggèrent donc qu'il est important de prendre en compte ce score d'inclination et de ne pas considérer uniquement les deux scores promotion et prévention indépendamment. Le score d'inclination permet de caractériser la tendance d'un participant à se projeter dans un focus plutôt qu'un autre. Notre questionnaire semble être efficace pour les participants dont les attitudes promotion et prévention sont bien différenciés. Par ailleurs, le RFQ-PF montre des liens aussi bien avec le RFQ qu'avec le RFSS (ce dernier étant présenté comme relativement indépendant du premier). Cela nous encourage à envisager d'étudier plus avant la structure du RFQ-PF pour affiner notre compréhension du construit sous-jacent mesuré.

Malgré ces résultats encourageants, cette étude présente un certain nombre de limitations. Bien que la taille des échantillons utilisés soit suffisante pour les tests statistiques appliqués, le nombre de participants reste relativement faible et manque probablement de représentativité. De plus, nos participants étaient majoritairement des femmes, et bien que cela aille au-delà de la question étudiée ici, il serait intéressant d'intégrer les effets de genre dans le processus de validation du RFQ-PF.

Pour conclure, le RFQ-PF nous paraît avoir les qualités appropriées pour son utilisation dans la suite de nos travaux, en auto-attribution chez l'humain comme en hétéro-attribution pour une entité artificielle. En effet, la consigne est facilement adaptable et nous espérons que l'aspect intuitif et parfois nébuleux des proverbes permettra à nos participants de juger plus facilement la personnalité d'un agent virtuel que les questionnaires classiques.

Après avoir posé la théorie du regulatory focus comme cadre théorique, proposé un cadre de travail, un modèle computationnel ainsi qu'un questionnaire permettant de mesurer le regulatory focus, nos prochaines contributions sont des propositions d'implémentation pour faire transparaitre une personnalité chez un agent artificiel, accompagnées d'études utilisateurs afin d'explorer la perception de la personnalité implémentée et son impact sur l'interaction.

4.5 RÉSUMÉ DU CHAPITRE

Les questionnaires classiquement utilisés pour mesurer le regulatory focus (RFQ (Higgins et al., 2001) ; GRFM (Lockwood et al., 2002)) sont centrés sur l'explicitation des motivations et des buts internes qu'une personne poursuit. Ces mesures nous paraissent difficiles à appliquer en hétéro-attribution dans le but de caractériser la personnalité perçue d'un agent virtuel. De plus, il n'existe pas de version validée en français de ces questionnaires. C'est pourquoi, en nous inspirant de Van Stekelenburg (2006), nous proposons une NOUVELLE ÉCHELLE DE REGULATORY FOCUS BASÉE SUR DES PROVERBES : le RFQ-PF (Regulatory Focus Questionnaire - Proverbs Form).

Le RFQ-PF propose au répondant d'indiquer à quel point certains proverbes peuvent être représentatifs de ses actions et de sa manière de penser. Les proverbes permettent d'appréhender des situations psychologiques variées par des phrases courtes et simples. Le RFQ-PF est destiné à être utilisé aussi bien en auto-attribution qu'en hétéro-attribution. Les proverbes utilisés comme items ont sélectionnés en suivant des étapes de réduction incrémentales : pré-sélection sur la validité apparente ; évaluation par une population tout-venant ; raffinement par des experts ; étude de la structure factorielle.

La version finale du RFQ-PF comprend 18 items, 10 pour l'échelle Prévention et 8 pour l'échelle Promotion. Des analyses factorielles exploratoire et confirmatoire montrent une STRUCTURE EN 3 FACTEURS : un facteur "Prévention", un facteur "Promotion-Approche" (items encourageant l'utilisation de stratégies d'approche) et un facteur "Promotion-Omission" (items mettant en garde contre les erreurs d'omission). Les deux facteurs "Promotion" présentent une corrélation substantielle de 0,46. Nous estimons que les items des deux facteurs "Promotion" peuvent être moyennés et considérés comme une seule échelle Promotion. Les deux échelles ainsi constituées montrent une BONNE COHÉRENCE INTERNE ($\alpha_{\text{Promotion}} = 0,78$; $\alpha_{\text{Prévention}} = 0,77$).

Les VALIDITÉS CONVERGENTE ET DISCRIMINANTE ont été explorées par l'étude des corrélations avec d'autres questionnaires de motivation et de régulation du Soi. Bien que les corrélations trouvées restent modestes, le motif mis en évidence est cohérent par rapport au motif attendu. Les corrélations les plus fortes ont été trouvées pour les échelles mesurant le regulatory focus, comme le RFQ (Higgins et al., 2001) ou le RFSS (Ouschan et al., 2007). Globalement, le RFQ-PF fournit une mesure correcte via un score Promotion et un score Prévention. De plus, le RFQ-PF semble être particulièrement pertinent pour la MESURE DE L'INCLINATION (préférence plus ou moins marquée pour l'un des focus) via la différence entre le score promotion et le score prévention.

Troisième partie

IMPLÉMENTATIONS ET EXPÉRIMENTATIONS

5

IMPLÉMENTATION DIMENSIONNELLE DU REGULATORY FOCUS ET INTERACTION À COURT TERME

Sommaire

5.1 Objectifs **74**
5.2 Méthodologie **75**
5.2.1 Choix d'un cadre applicatif ludique 75
5.2.2 Implémentation des stratégies de jeu 77
5.2.3 Choix d'un agent "contrôle" 82
5.3 Hypothèses générales **82**
5.4 Protocole expérimental **83**
5.4.1 Mesures 83
5.4.2 Participants 83
5.4.3 Protocole 84
5.4.4 Hypothèses opérationnelles 85
5.5 Résultats **85**
5.5.1 Statistiques descriptives 85
5.5.2 Statistiques inférentielles 87
5.6 Discussion **89**
5.6.1 Perception de la personnalité 89
5.6.2 Crédibilité 91
5.6.3 Effet de regulatory-fit 91
5.7 Résumé du chapitre **93**

Dans ce chapitre, nous présenterons une première étude de l'interaction entre un utilisateur et un agent artificiel dont le regulatory focus sera implémenté à l'aide d'une méthode data-driven. Dans un premier temps, nous voulons voir si cette caractéristique de la personnalité peut être transmise par le comportement de l'agent de la même façon qu'un trait de type Big Five. Nous pourrons également étudier la possibilité d'établir un effet de regulatory-fit lorsque l'utilisateur et l'agent artificiel présentent le même regulatory focus.

Nous commencerons par présenter les objectifs généraux de cette étude (section 5.1). Ensuite, le cadre applicatif choisi sera détaillé ainsi que la méthodologie utilisée pour l'implémentation des différents agents utilisés dans cette expérimentation (section 5.2). Après quoi les hypothèses générales seront formulées (section 5.3) et le protocole employé sera décrit (section 5.4). Enfin, nous présenterons les résultats obtenus (section 5.5) que nous discuterons ensuite (section 5.6).

5.1 objectifs

Notre premier objectif est de tester si le regulatory focus d'un agent peut être perçu par un utilisateur au même titre qu'un trait de personnalité du Big Five. En effet, à notre connaissance, aucune étude ne porte sur l'implémentation du regulatory focus chez un agent artificiel. Si la littérature montre qu'il est possible qu'un utilisateur perçoive certains traits du Big Five chez un agent de la façon dont ceux-ci ont été mis en place, il convient d'en faire de même avec le concept de regulatory focus. Cette étude vise donc à reproduire les résultats de l'état de l'art avec le concept de regulatory focus. C'est pourquoi nous utiliserons dans un premier temps les scores de regulatory focus (promotion et prévention) comme des traits, c'est-à-dire sans interroger directement le processus sous-jacent.

Notre second objectif est de voir s'il est possible de produire un effet de regulatory-fit concernant l'interaction entre un utilisateur et un agent artificiel, c'est-à-dire que l'agent et l'interaction sont perçus comme plus appréciables lorsque le focus de l'interaction et celui de l'agent sont orientés dans la même direction. L'effet regulatory-fit a été montré dans différentes situations (cf chapitre 3) mais n'a pas été étudié dans le cadre d'une interaction homme-machine. Pourtant, si la personnalité de l'agent peut amener un effet de regulatory-fit avec l'utilisateur, il est possible que que le sentiment de justesse ainsi créé soit transféré sur l'agent et que l'utilisateur s'implique plus dans la tâche. En d'autres termes, il s'agit d'étudier si un utilisateur ayant un focus chronique promotion (resp. prévention) préfère interagir avec un agent présentant un focus promotion (resp. prévention) lui aussi.

Nous avons également pour objectif de proposer une méthodologie d'implémentation de la personnalité data-driven, i. e. s'appuyant

sur des données réelles. S'appuyer sur une théorie issue de la psychologie permet de proposer une implémentation dite theory-driven en se basant sur l'interprétation qui peut en être faite. Utiliser une théorie socio-cognitive comme celle du regulatory focus, détaillant des structures et des processus, donne une base solide pour imaginer un processus cognitif artificiel. Cependant, les théories psychologiques sont définies de façon à éclairer le fonctionnement de l'être humain et pas pour guider une implémentation informatique. Il n'est pas rare de se heurter à des problématiques propres à l'implémentation qui sont sous-spécifiées au niveau théorique. Cette méthodologie theory-driven repose donc principalement sur l'inspiration des informaticiens et sur leur capacité à transformer la théorie en spécifications fonctionnelles.

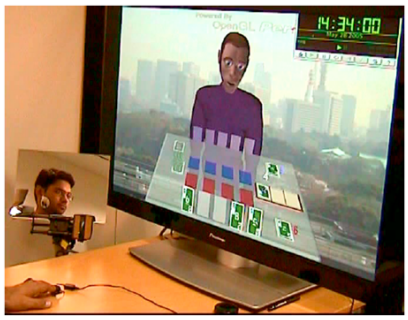
Nous proposons donc d'utiliser une méthodologie data-driven, en recueillant des données utilisateurs et en utilisant des méthodes d'apprentissage automatique afin de créer une première implémentation permettant de diriger le comportement d'un agent artificiel. Cette méthodologie présente plusieurs avantages : cette première implémentation est exempte d'apriori théoriques et de premières études expérimentales peuvent avoir lieu afin d'observer la perception et les effets des comportements ainsi modélisés. De plus, si ces premières études sont concluantes, en choisissant une méthode d'apprentissage automatique dont la sortie reste interprétable, celle-ci peut être utilisée en complément de la théorie pour guider une implémentation plus symbolique mettant en avant les processus à modéliser.

5.2 méthodologie

5.2.1 *Choix d'un cadre applicatif ludique*

Pour tester notre implémentation du regulatory focus chez un agent artificiel, nous choisissons d'utiliser un cadre applicatif ludique : l'interaction consistera en une partie d'un jeu avec un agent. Dans un premier temps, la stratégie utilisée par l'agent lors du jeu sera la seule et unique modalité permettant à l'agent artificiel d'exprimer une personnalité. De nombreux liens ont été fait entre jeu et personnalité, aussi bien en psychologie (Bartle, 1996; Johnson and Gardner, 2010) qu'en informatique (Elliott, 1994; Gratch et al., 2009; Johansson and Verhagen, 2014). Les situations de jeu sont d'ailleurs souvent utilisées en informatique affective pour créer et évaluer les aspects sociaux d'agents ou de robots (cf Figure 10).

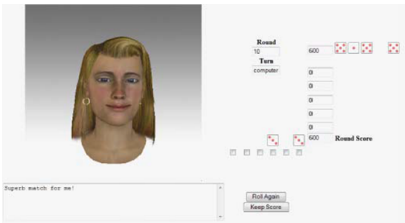
Plusieurs éléments sont à prendre en compte pour choisir un jeu adapté à notre cadre de recherche. En premier lieu, le jeu doit permettre à l'agent d'exprimer son regulatory focus par ses actions dans le jeu, ainsi que par des comportements "expressifs" (comportement



(a) Becker et al. (2005)

(b) Paiva (2011)

(c) Pereira et al. (2012)



(d) Bosse and Zwanenburg (2014)

Figure 10 – Etudes utilisant des jeux pour évaluer des agents ou des robots affectifs

non-verbal) pour de futures études. Comme nous ne prévoyons pas de comportement verbal à ce stade, le jeu choisi ne devra pas nécessiter d'éléments de langage au cours du jeu. D'un point de vue du gameplay, les règles doivent pouvoir être apprises rapidement et le jeu doit présenter un certain potentiel d'engagement pour que les utilisateurs ne s'en désintéressent pas. De plus, le temps d'une partie doit être suffisamment long pour permettre à un utilisateur de se faire un avis sur la personnalité de l'agent mais suffisamment court pour que les temps de passation expérimentale restent raisonnables. Enfin, le matériel de jeu doit pouvoir être accessible à l'agent (i. e. de préférence implémentable comme une interface) et ne pas être trop distractif pour éviter que l'attention des utilisateurs se focalisent dessus.

Nous avons donc sélectionné un jeu de plateau, nommé Can't Stop [1]. Il s'agit d'un jeu de type "stop-ou-encore", i. e. un jeu comportant une décision permettant de s'arrêter (stop), ce qui a pour effet de sécuriser les gains courants mais obligeant à laisser alors la main à l'adversaire, ou de rejouer (encore), ce qui permet de potentiellement gagner plus au risque de perdre les gains déjà acquis.

1. Créé par Sid Sackson, édité en France par Asmodee Editions

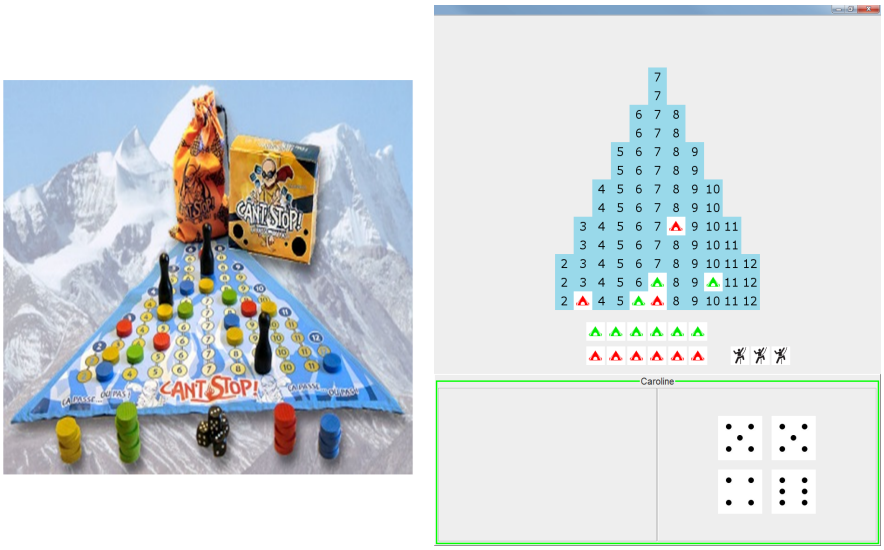


Figure 11 – Plateau de jeu Can't Stop réel (à gauche) et virtuel (à droite)

Utiliser un jeu existant permet de s'assurer que le jeu présente un bon potentiel ludique. Le matériel de jeu est plutôt basique et nous a permis de facilement l'adapter pour en faire un jeu sur ordinateur (Figure 11). Les règles sont assez simples (règle du jeu détaillée en Annexe B) et une partie dure entre 10 et 20 minutes. Ce jeu utilise des lancers de dés et ne permet pas d'anticipation : l'utilisateur ne sera donc pas concentré sur sa stratégie de jeu durant le tour de l'agent et pourra donc observer la stratégie de l'agent. Par ailleurs, du point de vue des perspectives possibles, concernant l'aspect non-verbal, le jeu Can't Stop permettra d'avoir différentes réactions émotionnelles transmissibles par des expressions faciales à différents moments du jeu. Le même jeu pourrait également être réalisé dans le contexte d'un agent pédagogique sur le thème des statistiques et des probabilités. De plus, ce jeu peut se jouer de 2 à 4 joueurs. Cela permet d'imaginer des jeux avec plusieurs compagnons artificiels.

Dans un premier temps, nous ne nous intéressons qu'à l'aspect stratégies de jeu pour essayer de transmettre la personnalité des agents que nous allons implémenter. La stratégie adoptée par l'agent pourra transmettre un certain regulatory focus sur deux points. Le premier point est le choix d'un mouvement, qui permet de rejouer avec une probabilité plus ou moins grande. Le deuxième point est la décision stop-ou-encore, qui permettra de faire la différence entre une stratégie vigilante (stop) ou une stratégie plus intrépide (encore).

5.2.2 *Implémentation des stratégies de jeu*

Comme justifié en section 5.1, plutôt que de commencer par une implémentation symbolique que nous imaginerions en nous appuyant

sur la théorie du regulatory focus (implémentation theory-driven), nous proposons d'utiliser une méthode s'appuyant sur des données de jeu réelles, collectées sur des jeux humain-humain (implémentation data-driven).

5.2.2.1 *Recueil de données*

Quinze participants (13 hommes, 2 femmes ; age M = 29,7 ans, SD = 10,2) ont joué au total 11 parties de Can't Stop. Avant de participer à ces sessions de jeu, chaque participant a répondu au RFQ-PF en auto-attribution afin d'obtenir une mesure de son regulatory focus chronique. Les parties ont eu lieu sur ordinateur avec deux écrans pour chaque joueur. Le premier écran contenait le plateau de jeu commun. Le deuxième écran montrait l'image de l'adversaire filmé par une webcam, afin d'essayer de reproduire un jeu avec un adversaire distant (tel un agent virtuel). Les participants avaient la possibilité de parler l'un avec l'autre au cours du jeu. Les différents écrans sont représentés en Figure 12.

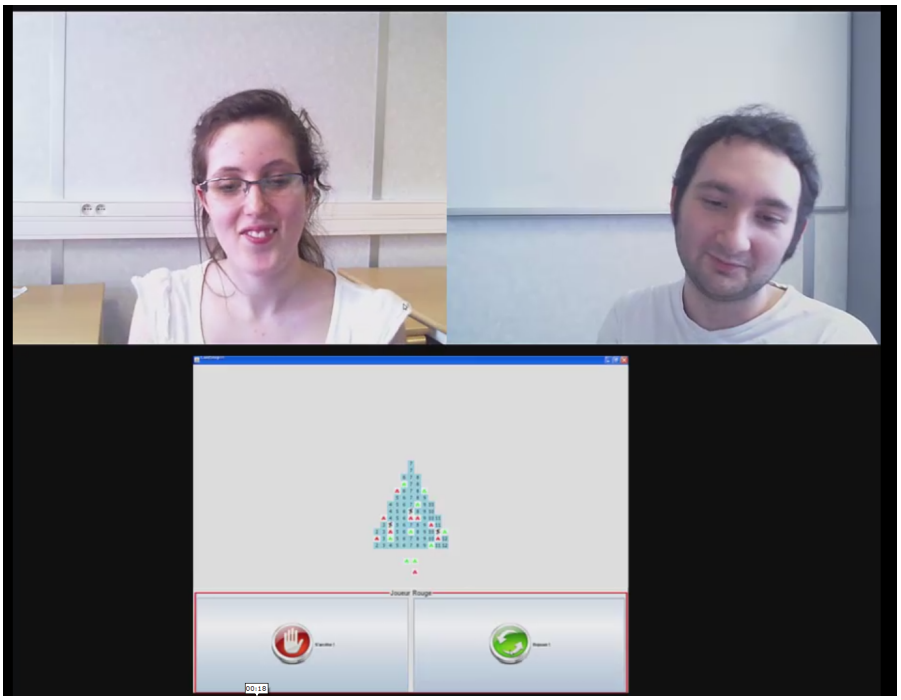


Figure 12 – Ecrans montrant les adversaires filmés par webcams (en haut) et image du plateau de jeu Can't Stop tel que vu par les participants chacun sur leur écran (en bas)

5.2.2.2 *Extraction des stratégies*

Pour extraire les stratégies qui seront utilisées par nos agents artificiels, deux classifieurs ont été entrainés, un pour chaque décision à réaliser dans le jeu : le choix d'un mouvement et la décision stop-

ou-encore. Des caractéristiques ont été calculées en fonction de notre propre analyse du jeu. Les caractéristiques peuvent être en rapport avec l'état global du jeu au moment de la décision (e. g.la différence de score entre les deux joueurs, le nombre de colonnes jouables), avec les mouvements potentiels (e. g.la distance par rapport au haut de la colonne, en absolu et relativement aux autres mouvements possibles) ou avec des statistiques concernant toute la partie (e. g.nombre de tours depuis le dernier décrochage).

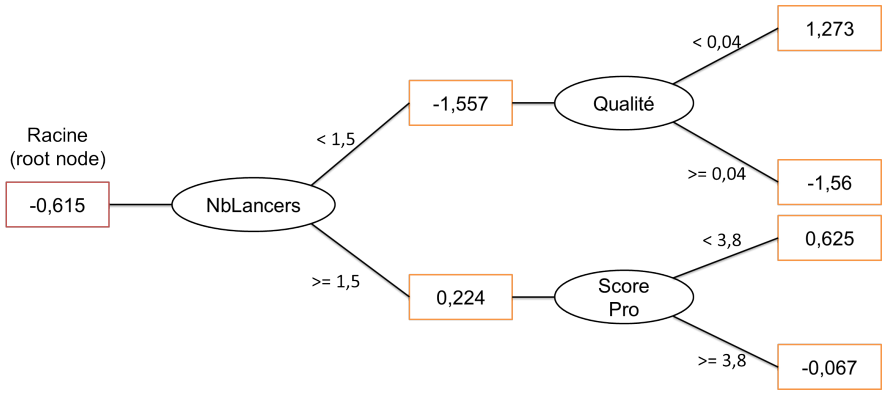
Pour obtenir une modélisation de la personnalité suivant une approche de type trait, les scores de personnalité des participants ont été intégrés en tant que caractéristiques pour la classification.

jeux de données Pour le choix du mouvement, une instance du jeu de données est composée du mouvement effectivement choisi (Chosen Move, CM), comparé avec un des mouvements alternatifs possibles (Alternative Move, AM), des caractéristiques calculées pour cette paire et de la classe-cible "Choisi" ou "NonChoisi". Chaque paire fait l'objet de deux instances, l'une de la forme (CM,AM) avec pour classe-cible "Choisi" et l'autre de la forme (AM,CM) avec pour classe-cible "NonChoisi". Le jeu de données était constitué de 1630 instances et de 108 caractéristiques.
Pour la décision "stop-ou-encore", une instance du jeu de données est composée des caractéristiques calculées au moment de la décision et d'une classe-cible "Continue" si le joueur a choisi de rejouer ou "Stop" si le joueur a choisi de s'arrêter. Le jeu de données était constitué de 670 instances et de 47 caractéristiques (le nombre de caractéristiques est plus important pour le jeu de données concernant le choix du mouvement puisque certaines sont relatives aux mouvements alternatifs).

classifieur adtree Concernant la méthode de classification, nous choisissons d'utiliser un classifieur binaire de type Alternating Decision Tree (*ADTree*) Holmes et al. (2002). Cette méthode a deux avantages principaux : la simplicité de l'arbre de décision et l'efficacité des méthodes de boosting. Le modèle créé est composé de noeuds de décision (prédicats) et de noeuds de prédiction (valeurs numériques). Il en résulte une structure d'arbre de décision combinant plusieurs arbres à un seul niveau (*decision stumps*), classiquement utilisé en boosting. Contrairement aux arbres de décision classiques, les différentes branches ne sont pas mutuellement exclusives. Pour classer une instance, il suffit de parcourir l'ensemble des branches suivant les noeuds de décision en additionnant les valeurs des noeuds de prédiction traversés. Le score final, qui est aussi une mesure de confiance sur la classification, permet de décider la classe à laquelle appartient l'instance.

Ce classifieur est une boite blanche, ce qui permet d'interpréter plus facilement l'arbre en sortie pour pouvoir éventuellement l'utiliser afin de guider une future implémentation symbolique. De plus, la qualité des résultats est équivalente à celles d'autres classifieurs que nous aurions pu utiliser [2]. Le paramètre boosting iterations, qui correspond au nombre de decision stumps utilisés, est ici fixé à 10.

À titre d'exemple, la Figure 13 présente une des branches de l'arbre de décision pour la décision stop-ou-encore avec prise en compte de la personnalité. Le noeud racine contient une première valeur. En fonction du nombre de lancers dans le tour courant, cette valeur est additionnée avec le noeud de prédiction. S'il y a eu moins de 1,5 lancers au moment de la décision, alors on passe au noeud de décision Qualité ; sinon, on observe le score promotion. Par exemple, si l'agent a déjà lancé trois fois dans ce tour et qu'il a un score promotion de 1, le score final à la sortie de cette branche est de $-0,615 + 0,224 + 0,625 = 0,234$. Dans la même situation, un agent ayant un score promotion de 7 aura un score à la sortie de cette branche de $-0,615 + 0,224 - 0,067 = -0,458$. En effectuant le parcours dans l'ensemble des branches de l'arbre, le score final permet à l'agent de choisir s'il s'arrête (score positif) ou s'il rejoue (score négatif).



NbLancers : *nombre de lancers dans le tour courant*
Qualité : *reste à grimper pondéré par la proximité de l'adversaire*
Score Pro *: score promotion-focus*

Figure 13 – Une branche de l'arbre utilisé pour la décision stop-ou-encore montrant les noeuds de prédiction (rectangles) et les noeuds de décision (ovales)

cross-validation Pour estimer l'efficacité de nos classifieurs, une validation croisée à 10 échantillons (*10-fold cross-validation*) a été réalisée. Les résultats de cette validation croisée ont été comparés à ceux d'un classifieur de type Zero-R (utilisé comme un indicateur

2. Les classifications ont été réalisées avec le logiciel Weka (version 3.7)

plancher ; ce classifieur prédit la classe-cible la plus fréquente indépendamment des prédicteurs) et différents classifieurs qui ont finalement présentés des performances équivalentes à celles du classifieur ADTree. À titre d'exemple, nous présentons les résultats d'un classifieur de type Random Forest (avec 10 arbres) dans le Tableau 10. Les variables observées pour juger de la qualité du classifieur sont le pourcentage d'items mal classifiés, le coefficient Kappa et l'aire sous la courbe ROC (*ROC area*). Les performances des classifieurs ADTree sont suffisamment bonnes comparées aux autres classifieurs et le coefficient Kappa est acceptable pour les trois modèles. Les arbres de décision produits par la méthode ADTree sont donc jugés suffisamment logiques et cohérents pour être utilisés. Il est à noter que pour le choix du mouvement, les scores de regulatory focus n'ont pas été sélectionnés comme caractéristique utile dans l'arbre de décision. Par contre, les scores de regulatory focus ont été sélectionnés pour la décision stop-ou-encore et le classifieur montre de meilleurs performances quand ces scores sont pris en compte.

Tableau 10 – Statistiques de la validation croisée à 10 échantillons pour trois classifieurs sur les différents modèles de jeu

Modèle pour le choix du mouvement			
	ZeroR	ADTree	RandomForest
Items mal classifiés	50,3%	16,5%	17,2%
Coeff. Kappa	0	0,67	0,66
ROC area	0,50	0,92	0,91
Modèle pour la décision stop-ou-encore - sans personnalité			
	ZeroR	ADTree	RandomForest
Items mal classifiés	22,5%	23,4%	22,1%
Coeff. Kappa	0	0,25	0,22
ROC area	0,50	0,79	0,77
Modèle pour la décision stop-ou-encore - avec personnalité			
	ZeroR	ADTree	RandomForest
Items mal classifiés	22,5%	19.70%	21,8%
Coeff. Kappa	0	0,33	0,23
ROC area	0,50	0,80	0,76

implémentation Pour les agents avec personnalité, deux arbres ont donc été implémentés : celui représentant le modèle pour le choix du mouvement et celui représentant le modèle pour la décision stop-ou-encore avec personnalité. De façon à reproduire des personnalités marquées, nous avons créé les deux agents suivants :

- l'agent RF-Pro, qui a un score promotion de 7 et un score prévention de 1
- l'agent RF-Pre, qui a un score promotion de 1 et un score prévention de 7

Les deux agents suivent donc les mêmes arbres de décision mais certains noeuds de décision du modèle pour la décision stop-ou-encore utilisent les valeurs des scores promotion ou prévention dans leur prédicat.

5.2.3 *Choix d'un agent "contrôle"*

Dans le but de tester la perception qu'un utilisateur a de la personnalité de nos agents artificiels, nous voulons utiliser un agent "contrôle", donc idéalement un agent qui jouera sans personnalité. Or, en psychologie, il n'existe pas de personne qui n'aurait pas de personnalité. Reste donc à savoir comment cette idée peut être rendue en informatique affective. Concernant la littérature du domaine, il s'avère que cette question méthodologique est assez peu abordée. Il apparait néanmoins que deux types de stratégies sont généralement utilisées en tant que contrôle : la stratégie aléatoire (mais l'absence d'une cohérence comportementale transmet-elle l'absence de personnalité ?) ou une stratégie de type intelligence artificielle traditionnelle (mais l'absence d'une personnalité implémentée en tant que telle est-elle équivalente à l'absence de personnalité ?).

Pour explorer cette question, 2 types d'agent contrôle ont été utilisés dans cette étude :

- L'agent Random, qui choisit ses mouvements de façon aléatoire et qui a une probabilité de 50% d'arrêter son tour
- L'agent Moyen, dont la stratégie suit un ADTree créé sans prendre les scores de regulatory focus en compte dans l'apprentissage

Le comportement de l'agent Random devrait présenter une grande variabilité dans ces comportements donc, même si les utilisateurs y voient un pattern représentant une prsonnalité, ce pattern devrait être différent pour chaque utilisateur. Pour l'agent Moyen, tous les utilisateurs seront confrontés à la même stratégie, dont le regulatory focus ne devrait être ni promotion, ni prévention.

5.3 hypothèses générales

Nous faisons trois hypothèses concernant cette étude :

- H1 : Les différences de personnalité des agents artificiels sont perçues par les participants.
- H2 : La personnalité apporte plus de crédibilité aux agents artificiels.

- H3 : En accord avec la théorie du regulatory fit, les participants préfèrent interagir avec un agent ayant le même regulatory focus que leur propre focus chronique.

5.4 protocole expérimental

5.4.1 *Mesures*

Lors de cette expérimentation, les mesures suivantes ont été utilisées.

questionnaires en auto-attribution Les participants ont répondu pour eux-mêmes au questionnaire de regulatory focus que nous avons conçu, le RFQ-PF, afin de mesurer leur focus chronique. Le RFQ-PF utilise une échelle de Likert à 7 points.

questionnaires en hétéro-attribution Les participants ont répondu à des questionnaires afin de décrire leur perception de l'agent avec lequel ils interagissaient. Concernant la perception du regulatory focus de l'agent, nous avons utilisé le RFQ-PF. Concernant la crédibilité, nous avons utilisé 10 questions du Godspeed Questionnaire (Bartneck et al., 2009) (5 questions à propos de l'appréciation de l'agent et 5 questions à propos de l'intelligence perçue). Nous utilisons ces échelles car Tseng and Fogg (1999) définissent la bienveillance perçue et l'expertise comme deux dimensions clés de la crédibilité. Le Godspeed Questionnaire utilise une échelle de Likert à 5 points.

questionnaires concernant l'interaction Pour évaluer l'interaction, nous utiliserons le concept de flow (Csikszentmihalyi, 1997). L'état de flow se définit par une forte concentration, un plein engagement dans une tâche et une certaine satisfaction dans son accomplissement. Pour évaluer le flow lors de l'interaction avec nos agents, les participants ont répondu au Flow Condition Questionnaire (Schaffer, 2013), composé de 3 questions. Ce questionnaire utilise une échelle de Likert à 5 points.

données de jeu L'ensemble des parties jouées par les participants étaient enregistrées sous forme de fichiers de log.

5.4.2 *Participants*

Vingt participants ont pris part à cette étude (11 hommes et 9 femmes, age M = 30,6 ans, SD = 8,1). Parmi ces participants, 17 étaient de langue maternelle française et 3 étaient bilingues. Nous n'avons pas pré-sélectionné les participants en fonction de leur focus chro-

nique. D'après les résultats du RFQ-PF, 14 participants avaient un focus chronique promotion et 6 un focus chronique prévention.

5.4.3 *Protocole*

Pour démarrer, les participants regardaient un tutoriel vidéo leur apprenant les règles du jeu Can't Stop [3]. Les participants pouvaient naviguer comme ils le voulaient dans la vidéo.

Une fois la vidéo terminée, les participants jouaient une partie d'apprentissage contre l'ordinateur afin de se familiariser avec le jeu et son interface. Le participant était informé que l'ordinateur jouerait de façon aléatoire durant la partie d'apprentissage. L'expérimentateur répondait alors à d'éventuelles questions concernant les règles du jeu.

Ensuite, l'expérimentateur annonçait au participant qu'il/elle allait jouer 4 parties contre différents agents artificiels (Figure 14) et qu'après chaque partie, il/elle devrait répondre à différents questionnaires pour évaluer son interaction avec l'agent contre lequel il/elle venait de jouer.

L'expérimentateur mettait l'accent sur le fait que l'étude cherchait à évaluer la perception de l'agent artificiel et que le fait de gagner ou non les parties n'avait pas d'importance. Bien qu'aucune information sonore ne soit diffusée, les participants portaient un casque audio pour obtenir une certaine isolation phonique leur permettant de se concentrer sur la tâche.

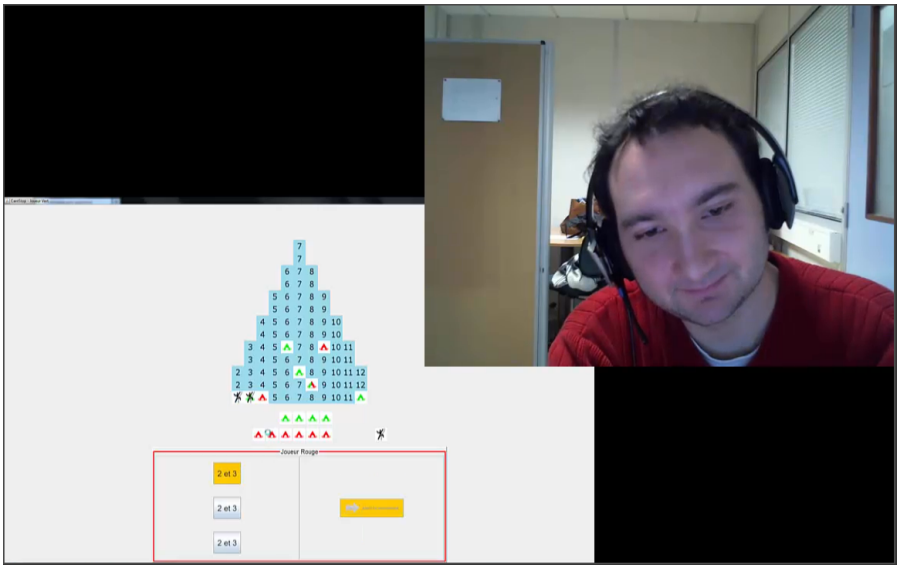


Figure 14 – Un utilisateur jouant à Can't Stop avec un agent artificiel durant le tour de l'agent. En bas à gauche, le plateau de jeu tel que vu par l'utilisateur et en haut à droite, l'enregistrement du visage de l'utilisateur par une webcam

3. Vidéo disponible à l'adresse : https ://www.youtube.com/watch ?v=kQs6JBXotxw

5.4.4 *Hypothèses opérationnelles*

Nous formulons donc les hypothèses opérationnelles suivantes :

- *H1 : Les différences de personnalité des agents artificiels sont perçues par les participants.*
  - Pour l'agent Random, les scores de regulatory focus sont dispersés et l'accord entre les participants est faible.
  - Pour l'agent Moyen, les scores de regulatory focus sont peu dispersés et l'accord entre les participants est bon. La moyenne du score promotion et du score prévention est de 3,5.
  - Pour les agents RF-Pro et RF-Pre, les scores de regulatory focus sont peu dispersés et l'accord entre les participants est bon. La moyenne du score promotion se situe autour de 5,5 pour l'agent RF-Pro (et respectivement concernant le score prévention pour l'agent RF-Pre).

- *H2 : La personnalité apporte plus de crédibilité aux agents artificiels.*
  - Les scores d'appréciation et d'intelligence perçue sont plus hauts pour les agents RF-Pro et RF-Pre que pour les agents Random et Moyen.
  - Le flow est plus important avec les agents RF-Pro et RF-Pre qu'avec les agents Random et Moyen.

- *H3 : En accord avec la théorie du regulatory-fit, les participants préfèrent interagir avec un agent ayant le même regulatory focus que leur propre focus chronique.*
  - Pour les participants avec un focus chronique promotion, les scores d'appréciation et de flow sont plus hauts pour les agents RF-Pro (et respectivement pour les participants avec un focus chronique prévention avec les agents RF-Pre).

5.5 Résultats

5.5.1 *Statistiques descriptives*

Nous observerons la moyenne (M) et l'écart-type (SD), ainsi que le coefficient de variation des quartiles (CQV ; $(Q3 - Q1)/(Q1 + Q3)$), qui est une statistique de dispersion permettant des comparaisons (Bonett, 2006) et le coefficient de Finn [4] (Finn, 1970), comme un indicateur de l'accord entre les participants. Ces indicateurs permettent une description plus fine de la dispersion que la simple observation des écarts-types. Les résultats complets sont présentés dans le Tableau 11.

Pour résumer, les résultats importants sont les suivants :

4. R package *irr*, version 0.84

Tableau 11 – Statistiques descriptives pour les différents scores collectés
Pro Sc. = score promotion ; Pre Sc. = score prévention ; Lik. = appréciation ; Perc. Int. = intelligence perçue ; CQV = Coefficient de variation des quartiles

		Random	**Moyen**	**RF-Pro**	**RF-Pre**
Pro Sc.	M	3,53	5,3	5,26	3,09
	SD	1,47	1,44	1,28	1,22
	CQV	32%	15%	14%	20%
	Finn coeff.	0,46	0,48	0,59	0,63
Pre Sc.	M	4,51	3,18	2,91	5,58
	SD	1,71	1,4	1,33	0,73
	CQV	30%	37%	25%	8%
	Finn coeff.	0,27	0,51	0,56	0,87
Lik.	M	3,3	3,22	3,02	3,51
	SD	0,78	0,67	0,95	0,75
	CQV	19%	13%	18%	14%
Perc. Int.	M	2,63	2,94	3,11	3,85
	SD	0,6	1,16	0,98	0,68
	CQV	18%	33%	27%	11%
Flow	M	3,55	3,67	3,72	3,67
	SD	0,66	0,68	0,61	0,73
	CQV	14%	9%	9%	13%

- Dispersion des données : Le CQV pour l'agent Random est en moyenne 1,5 fois plus élevé que pour les autres agents.

- Accord inter-participants : Le coefficient de Finn de l'agent Random est le plus faible pour le score promotion comme pour le score prévention.

- Scores de personnalité
  - Score promotion : l'agent RF-Pro et l'agent Moyen on un haut score promotion (resp. $M = 5,26$; $SD = 1,28$ et $M = 5,30$; $SD = 1,44$) contrairement à l'agent RF-Pre et à l'agent Moyen (resp. $M = 3,09$; $SD = 1,22$ et $M = 3,53$; $SD = 1,47$).
  - Score prévention : l'agent RF-Pre et l'agent Random on un haut score prévention (resp. $M = 5,58$; $SD = 0,73$ et $M = 4,51$; $SD = 1,71$) contrairement à l'agent RF-Pro et à l'agent

Moyen (resp. $M = 2,91$; $SD = 1,33$ et $M = 3,18$; $SD = 1,40$).

- Flow : les scores sont au-dessus de 3,5 (sur une échelle à 5 points) pour tous les agents.

- Crédibilité
  - Appréciation : les scores sont autour de 3 (sur une échelle à 5 points) pour tous les agents.
  - Intelligence perçue : les scores vont de 2,63 (agent Random) à 3,85 (agent RF-Pre).

5.5.2 *Statistiques inférentielles*

Pour l'analyse des différences entre les différents agents, des statistiques non-paramétriques ont été utilisées car la normalité des données ne pouvait être assurée au vu du nombre de participants. Pour l'analyse principale, le test de Friedman a été appliqué et en tant que test post-hoc, des tests des rangs signés de Wilcoxon (avec une correction de Holm pour ajuster la p-value).

score promotion (Figure 15) Le test de Friedman montre une différence significative pour le score promotion ($\chi^2(3) = 23,44$; $p < 0,001$). Le score promotion de l'agent RF-Pro est significativement plus haut que celui de l'agent RF-Pre ($p \leqslant 0,001$) et de l'agent Random ($p \leqslant 0,01$). Il n'y a pas de différence significative entre l'agent RF-Pro et l'agent RF-Pre.

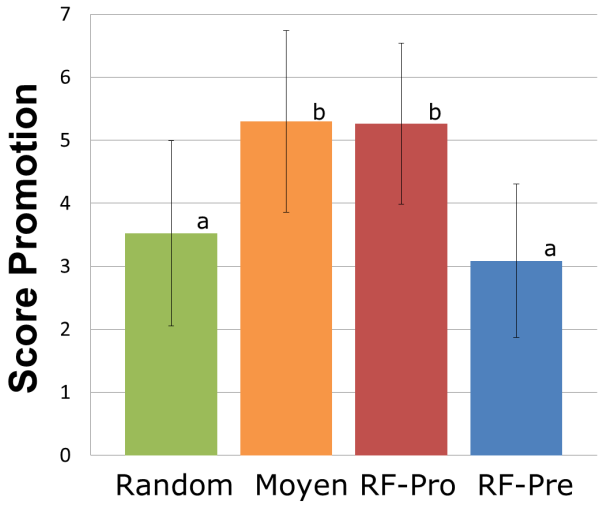


Figure 15 – Score promotion pour les différents agents.
Les valeurs avec la même lettre ne sont pas significativement différentes à p < 0,05 ; Barre d'erreur = SD

score prévention (Figure 16) Le test de Friedman montre également une différence significative pour le score prévention ($\chi^2(3) = 23,28$; $p < 0,001$). Le score prévention de l'agent RF-Pre est significativement plus haut que celui de l'agent RF-Pro ($p \leqslant 0,001$) et de l'agent Moyen ($p \leqslant 0,01$). Il n'y a pas de différence significative entre l'agent RF-Pre et l'agent Random.

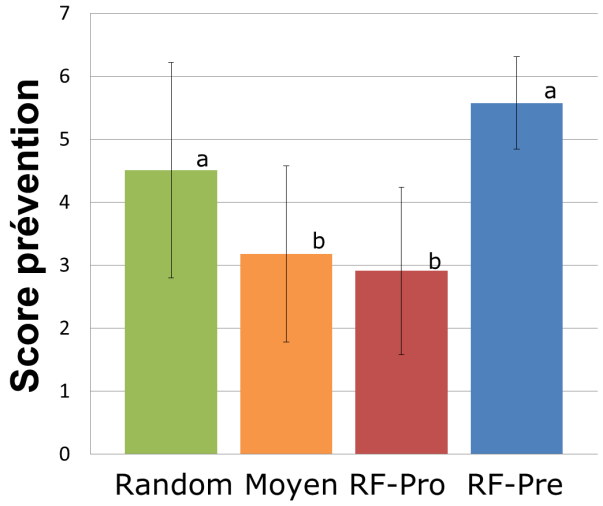
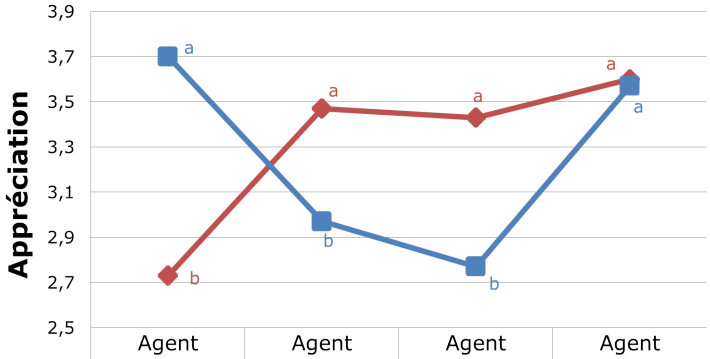


Figure 16 – Score prévention pour les différents agents.
Les valeurs avec la même lettre ne sont pas significativement différentes à p < 0,05 ; Barre d'erreur = SD

crédibilité Il existe une différence significative pour l'intelligence perçue ($\chi^2(3) = 15,18$; $p = 0,002$) : l'agent RF-Pre est perçue comme significativement plus intelligent que l'agent Random, l'agent Moyen et l'agent RF-Pro (resp. $p \leqslant 0,01$; $p \leqslant 0,05$; $p \leqslant 0,05$). Il n'y a pas de différence significative entre les trois autres agents. Il n'y a pas de différence significative concernant l'appréciation ($\chi^2(3) = 2,24$; n.s.).

effet de regulatory-fit Pour l'analyse des interactions, seuls les participants ayant les focus chroniques les plus marqués (i. e. les plus grandes différences entre leur score promotion et leur score prévention) ont été pris en compte car ils sont ceux qui sont les plus enclins à montrer un effet de regulatory-fit. De ce fait, ces analyses ont été réalisées sur un sous-ensemble des données comprenant les 6 participants les plus promotion et les 6 participants les plus prévention. De façon à obtenir un test robuste bien que nos données soient non-paramétriques, des tests ajustés sur rangs transformés (*adjusted rank transform test*, Leys and Schumann (2010)) ont été utilisés.

Les tests ont montré une interaction significative entre le focus chronique des participants et le type d'agent pour le score d'appréciation ($F(3,30) = 4,532$; $p < 0,01$; $\eta^2 partiel = 0,31$). Comme le



	Agent Random	Agent Moyen	Agent RF-Pro	Agent RF-Pre
Participants promotion	2,73	3,47	3,43	3,6
Participants prévention	3,7	2,97	2,77	3,57

Figure 17 – Interaction entre le focus chronique des participants et la personnalité de l'agent pour le score d'appréciation.
Les valeurs avec la même lettre ne sont pas significativement différentes à $p < 0{,}05$

montre la Figure 17, les participants avec un focus chronique prévention ont trouvé l'agent Random et l'agent RF-Pre (tous deux perçus comme prévention) plus appréciables que les agents Moyen et RF-Pro (tous deux perçus promotion). Les participants avec un focus chronique promotion ont trouvé l'agent Random moins appréciable que les autres agents. Aucune autre interaction significative n'a été trouvée.

5.6 Discussion

5.6.1 *Perception de la personnalité*

Notre première hypothèse (*Les différences de personnalité des agents artificiels sont perçues par les participants*) est en partie validée.

D'une part, les résultats ont montré que les agent RF-Pro et RF-Pre ont été respectivement perçus comme ayant un focus promotion et prévention. Nos stratégies de jeu implémentées de façon data-driven ont donc transmis le regulatory focus de l'agent de la façon prévue. Le concept de regulatory focus et son implémentation data-driven semblent donc donner des résultats équivalents à ceux des études en informatique affective sur la personnalité.

D'autre part, les résultats ont également montré que les agents Random et Moyen ont été respectivement perçus comme ayant un focus prévention et promotion. Dans un premier temps, cette perception peut être expliquée en analysant les stratégies de jeu de ces deux agents.

Concernant l'agent Random, ce dernier a une probabilité de 50% d'arrêter son tour à chaque décision. La probabilité P de s'arrêter après i décisions est donc $P_i = \frac{1}{2^i}$. Les tours courts sont donc plus probables que les tours longs. En conséquence les utilisateurs ont été plus probablement exposés à des tours courts lors de leur partie avec cet agent. Ces arrêts rapides ont pu être interprétés comme une stratégie plutôt orienté sur la sécurité, amenant les participants à percevoir l'agent Random comme plutôt prévention.

Concernant l'agent Moyen, nous avons analysé l'arbre de décision utilisé par cet agent. Cet arbre présente un noeud de décision concernant une caractéristique relative à la situation sur le plateau de jeu, dont la valeur est positive si l'adversaire est mieux placé sur le plateau et négative sinon. Lorsque cette caractéristique est supérieure à 1, le noeud de prédiction associé a une valeur de -1,61 (rappelons que si le score final est négatif, l'agent choisit de rejouer). De ce fait, dès que l'adversaire a une meilleur situation que l'agent dans le jeu, l'agent a une forte inclination à rejouer, ce qui encourage les tours longs. De plus, bien que nous n'ayons pas les probabilités associées à chacun des 81 scores finaux possibles, il s'avère que seul 2 d'entre eux amène l'agent Moyen à arrêter son tour. Des tours longs peuvent être vus comme une stratégie plutôt audacieuse, plus représentative du focus promotion.

D'un point de vue plus général, l'attribution d'une personnalité marquée aux agents "contrôle" peut être expliquée par ce que Shermer (2012) appelle *patternicity* et *agenticity*, soit la tendance à percevoir des motifs récurrents significatifs dans du bruit non-signifiant [5] et la tendance à voir un sens, une intention et une volonté dans ces motifs [6]. Ces tendances ont également pu être renforcées par le temps d'interaction relativement court. En effet, l'interaction dans cette étude est restée assez brève (une seule partie avec chaque agent). Cette interaction étant à court-terme, les utilisateurs sont plus dans la formation d'impression que dans le véritable jugement de personnalité. Il est possible que, si les utilisateurs avaient pu interagir plusieurs fois avec les agents, la cohérence des stratégies des agents sur le long-terme (ou l'incohérence dans le cas de l'agent Random) aurait pu modifier la perception de leur personnalité.

Les résultats de l'agent Random et de l'agent Moyen mettent en lumière la difficulté de contrôler la perception de la personnalité chez des agents artificiels. En effet, si nos stratégies data-driven transmettent le regulatory focus attendu, des stratégies tout autres permettent également aux utilisateurs de percevoir une personnalité chez l'agent. Si une méthodologie telle que celle que nous proposons nous

5. "the tendency to find meaningful patterns in meaningless noise"
6. "the tendency to infuse [these] patterns with meaning, intention and agency"

semble une meilleure base pour s'assurer de transmettre la personnalité espérée par les concepteurs, nous voyons que les études perceptives ne permettent pas vraiment de juger de la validité de l'implémentation elle-même.

5.6.2 *Crédibilité*

Notre deuxième hypothèse (*La personnalité apporte plus de crédibilité aux agents artificiels*) n'est pas validée.

Pour l'appréciation, les scores sont au dessus de la moyenne pour tous les agents. Ce pourrait être le reflet d'un léger biais positif envers les agents, dû à une tendance naturelle de l'esprit à se focaliser sur les côtés positifs à un niveau inconscient (*person positivity bias*, Sears (1983)). Toujours est-il que nous n'avons pas trouvé de différences globales entre nos agents concernant leur appréciation. Les participants ont d'ailleurs relaté des difficultés à évaluer les items relatifs à cette dimension. La brièveté de l'interaction peut être mise en cause, tout comme le fait que des stratégies de jeu seules ne permettent que difficilement de communiquer un concept aussi social.

Concernant l'intelligence perçue, les scores sont plus dispersés. L'agent RF-Pre a été noté comme significativement plus intelligent que les autres agents. Et bien que l'agent RF-Pro ait été noté en moyenne comme plus intelligent que l'agent Random et l'agent Moyen, cette différence n'est pas statistiquement significative. Etant donné le faible nombre de participants, il est difficile de savoir si cette non-significativité est dûe à un manque de données ou à une réelle différence transmise par les stratégies des agents. Il est également possible que la capacité de l'agent à gagner une partie (ou au minimum à se montrer un adversaire ayant un certain niveau de jeu) soit reliée à la perception de l'intelligence de l'agent par l'utilisateur. À ce propos, nous pouvons noter que l'agent RF-Pre a gagné 47% de ses parties, quand l'agent RF-Pro en a gagné 42%. De son côté, l'agent Moyen n'a gagné que 32% de ses parties, quand l'agent Random n'a que 5% de réussites.

5.6.3 *Effet de regulatory-fit*

Notre troisième hypothèse (*En accord avec la théorie du regulatory fit, les participants préfèrent interagir avec un agent ayant le même regulatory focus que leur propre focus chronique*) est en partie validée.

Nous avons trouvé une interaction entre le focus chronique des sujets et la personnalité des agents sur le score d'appréciation : les participants avec un focus prévention ont trouvé l'agent RF-Pre et l'agent Random plus appréciables que les agents RF-Pro et Moyen. Les agents RF-Pre et Random ont été perçus comme prévention, il

semble donc que nous ayons un effet de regulatory fit pour les participants prévention. Cependant, cet effet n'a pu être montré que pour les participants ayant un focus chronique clairement différencié. De plus, ce résultat n'est pas répliqué pour les participants avec un focus promotion : s'ils ont bien trouvé les agents RF-Pro et Moyen (tous deux perçus comme promotion) plus appréciables que l'agent Random, ils ont également trouvé l'agent RF-Pre aussi appréciable que les deux autres.

L'ensemble de ces résultats semblent montrer que le regulatory focus est un concept perceptible, au même titre que les dimensions du Big Five, par un utilisateur interagissant avec une machine. Un effet de regulatory fit a également pu être mis en évidence chez les participants ayant un focus chronique prévention. Cependant, l'implémentation proposée ici est dimensionnelle et ne permet pas d'informer le concept de regulatory focus lui-même, du point de vue du processus psychologique. De plus, les interactions se limitent à une partie de Can't Stop avec chaque agent. Il n'est donc pas possible de questionner le développement de la relation potentielle entre l'utilisateur et l'agent. C'est pourquoi notre contribution suivante est constituée d'une seconde implémentation, orientée socio-cognitive, et d'une étude utilisateur à sessions répétées.

5.7 résumé du chapitre

Avec cette première étude, nous avons cherché à repliquer les résultats de l'état de l'art avec le concept de regulatory focus. Nous avons conçu une première implémentation dimensionnelle du regulatory focus pour des agents artificiels et nous avons demandé à des utilisateurs d'évaluer la personnalité de ces agents après une interaction ludique (jeu de plateau de type stop-ou-encore nommé Can't Stop).

Pour cette implémentation dimensionnelle, nous avons proposé une méthodologie data-driven. Nous avons donc demandé à des participants de jouer à Can't Stop entre eux. Avec ces logs de jeu et les scores des participants au RFQ-PF, nous avons extrait automatiquement des stratégies de jeu à l'aide d'un classifieur ADTree (Alternating Decision Tree). Les scores promotion et prévention sont utilisés par le modèle de prise de décision (concernant le choix stop-ou-encore). Ces arbres de décision ont alors été implémentés pour diriger le comportement d'agents artificiels.

De cette façon, nous avons créé 4 agents artificiels : un agent RF-Pro avec un score promotion de 7 et un score prévention de 1 ; un agent RF-Pre avec un score promotion de 1 et un score prévention de 7 ; un agent Moyen, utilisant un modèle également issu d'un ADTree mais sans prise en compte des scores de personnalité ; un agent Random, ayant une stratégie aléatoire.

20 participants ont interagi avec les 4 agents (en jouant une partie de Can't Stop contre chaque) avant de remplir des questionnaires concernant la personnalité perçue de l'agent et sa crédibilité (sous la forme de deux échelles : appréciation et intelligence perçue). La stratégie de jeu perçue était la seule modalité disponible pour juger l'agent.

Les résultats ont montré qu'à l'instar des études utilisant le Big Five, il était possible de percevoir le regulatory focus chez un agent artificiel (agent RF-Pro perçu comme promotion ; agent RF-Pre perçu comme prévention). Cependant, nos agents "contrôle" ont également été perçu comme possédant un regulatory focus (agent Moyen perçu comme promotion ; agent Random perçu comme prévention) bien que les indicateurs de dispersion soient plus importants pour ces agents.

Un effet de regulatory fit (congruence entre le focus chronique du participant et le focus perçu chez l'agent) a pu être en partie démontré chez les participants ayant une inclination clairement marquée : les participants ayant un focus prévention ont trouvé les agents RF-Pre et Random plus appréciables que les autres. Néanmoins, cet effet n'a pu être montré pour les participants promotion.

6

IMPLÉMENTATION SOCIO-COGNITIVE DU REGULATORY FOCUS ET INTERACTIONS RÉPÉTÉES

Sommaire

6.1 Objectifs **96**
6.2 Méthodologie **99**
6.2.1 Implémentation du comportement des agents 99
6.2.2 Choix d'un agent "contrôle" 105
6.3 Hypothèses générales **109**
6.3.1 Concernant la perception de l'agent 110
6.3.2 Concernant la perception de l'interaction 110
6.3.3 Concernant la relation avec l'agent 110
6.4 Protocole expérimental **110**
6.4.1 Mesures 110
6.4.2 Participants 113
6.4.3 Protocole 113
6.4.4 Hypothèses opérationnelles 116
6.5 Résultats **117**
6.5.1 Validité des échelles 118
6.5.2 Données de jeu 119
6.5.3 Perception de l'agent artificiel 119
6.5.4 Perception de l'interaction 123
6.5.5 Perception de la relation 125
6.6 Discussion **136**
6.6.1 Perception de l'agent 137
6.6.2 Concernant la perception de l'interaction 139
6.6.3 Concernant la relation avec l'agent 140
6.7 Résumé du chapitre **142**

Dans ce chapitre, nous présenterons une deuxième étude de l'interaction entre un utilisateur et un agent virtuel présentant un certain regulatory-focus. Notre objectif est d'étendre et d'approfondir les résultats de notre étude précédente, mais en apportant différentes modifications au regard des limites de celle-ci.

En premier lieu, les objectifs généraux de cette étude seront présentés (section 6.1). Nous détaillerons la méthodologie employée pour l'implémentation des différents agents utilisés dans cette expérimentation (section 6.2) et les hypothèses que nous formulons (section 6.3). Nous présenterons ensuite les participants et le protocole employé (section 6.4). Enfin, nous décrirons les résultats obtenus (section 6.5) que nous discuterons ensuite (section 6.6).

6.1 objectifs

L'objectif principal de cette expérimentation est d'étudier la perception du regulatory-focus d'un agent artificiel non-stéréotypique lors de sessions d'interaction répétées et l'impact du regulatory focus sur l'interaction. Nous conservons le même cadre applicatif que précédemment : une interaction ludique au travers du jeu Can't Stop.

Nous avons vu qu'il était possible de transmettre la personnalité de l'agent via une implémentation data-driven de stratégies de jeu (chapitre 5). Cependant, nous avions opté dans un premier temps pour une approche de type trait, en utilisant les scores promotion et prévention comme caractéristiques pour la classification. En d'autres termes, un seul arbre de décision avait été implémenté et les scores promotion et prévention étaient présents dans les noeuds de décision (i. e. prédicats du type : si score promotion < X alors ...). Chaque score de regulatory-focus influe donc sur la décision finale mais les différences entre les processus promotion et prévention ne sont pas mises en avant. Pour nous inscrire plus clairement dans une approche socio-cognitive, l'utilisation des données pour l'implémentation des stratégies de jeu doit donc être revue.

Par ailleurs, nous avons utilisé des paramètres extrêmes pour la personnalité de nos agents (score promotion à 7 et score prévention à 1 pour l'agent RF-Pro ; score promotion à 1 et score prévention à 7 pour l'agent RF-Pre). Or, dans la population humaine, de tels scores sont extrêmement rares. Les personnes présentent des scores plus équilibrés, une personnalité plus nuancée. Pour que le comportement de l'agent artificiel se rapproche du comportement humain, nous devons proposer une implémentation plus subtile du regulatory focus dans les stratégies de jeu.

Une nouvelle implémentation des stratégies de jeu pour les agents RF-Pro et RF-Pre est donc proposée. Celle-ci est détaillée en sous-sous-section 6.2.1.1. L'étude présentée dans ce chapitre nous permet-

tra donc de voir si ces nouvelles stratégies permettent également de transmettre le regulatory-focus de l'agent.

Les stratégies de jeu sont une première modalité pour essayer de transmettre la personnalité. Cependant, dans le cadre d'une interaction humain-humain, la personnalité d'un individu est évaluée de façon globale. De plus, dans l'étude précédente, plusieurs participants ont évoqué des difficultés à évaluer l'agent, jugeant l'interaction trop limitée au sens où ils avaient plus l'impression de jouer "contre l'ordinateur" que contre un individu. Il nous parait donc intéressant d'étudier l'apport d'une modalité visuelle, c'est-à-dire d'incarner l'agent artificiel dans un agent virtuel visible par l'utilisateur.

Nous ajoutons donc à notre dispositif un agent virtuel situé dans un environnement 3D, affiché à l'écran à la manière de l'image renvoyée par une webcam lors d'une conversation Skype (plan rapproché poitrine, cf Figure 18). Nous utilisons un des agents de la plateforme MARC (Multimodal Affective and Reactive Characters) développée au LIMSI-CNRS (Courgeon et al., 2014).

Figure 18 – Agent virtuel MARC utilisé lors de la deuxième étude

En plus d'incarner l'agent artificiel, cet agent virtuel sera capable de réagir au cours du jeu avec des expressions faciales exprimant des émotions. Ces expressions faciales seront calculées par un modèle émotionnel de type cognitif, adapté à chaque focus. La définition de ce modèle émotionnel est décrite en sous-sous-section 6.2.1.2. Nous aurons donc une expression multiple de la personnalité, car celle-ci sera exprimée de deux manières : par la stratégie de jeu et par les expressions faciales de l'agent.

Dans cette étude, nous pourrons donc évaluer l'impact de la présence d'un agent virtuel sur l'évaluation de l'agent et de l'interaction par les participants, ainsi que l'apport d'une deuxième voie d'expression de la personnalité.

La brièveté de l'interaction était une limite importante de notre expérimentation précédente. En effet, dans notre présentation de la

personnalité, l'accent est mis sur l'importance d'obtenir des comportements stables et cohérents au cours du temps. Or, il est difficile d'évaluer cet aspect lorsque les sujets interagissent une unique fois avec nos agents. À défaut d'une interaction sur le long-terme, nous voulons maintenant permettre aux participants d'interagir plusieurs fois avec le même agent. Par ces sessions répétées, nous essayons donc de nous éloigner de la formation d'impression pour nous approcher d'un jugement de personnalité.

Cette deuxième étude est également l'occasion de se rapprocher du cadre applicatif du projet MoCA en utilisant une population proche de celle visée par le projet : des collégiens en classe de 4ème.

Bien que peu d'études se soient intéressées au regulatory focus chez le jeune adolescent, Zhao and Pechmann (2007) ont trouvé un effet de regulatory fit entre le focus chronique d'adolescents (entre 14 et 15 ans) et l'orientation de messages pour une campagne anti-tabac. Il parait donc intéressant d'explorer les liens entre regulatory focus et adolescence, d'autant plus que le regulatory focus montre des liens importants avec l'éducation et se développe dès l'enfance (Keller, 2008). À notre connaissance, aucune étude n'utilise cette population pour évaluer le regulatory focus perçu chez autrui.

Cependant, l'adolescence reste une période de transition dans le développement de la personnalité. Durant cette phase de maturation, certains traits de personnalité changent, d'autres se stabilisent et cela à des vitesses différentes suivant les individus (Klimstra et al., 2009). Cette période développementale particulière affecte la perception qu'un adolescent a de sa propre personnalité comme de celle des autres (Branje et al., 2003). Il sera donc important de garder à l'esprit la spécificité de la population étudiée dans l'interprétation des résultats de cette étude.

Nos participants seront donc des adolescents qui interagiront avec un même agent lors de sessions répétées. Nous présenterons la tâche comme certes compétitive mais également comme collaborative, afin de nous rapprocher de la notion de compagnon.

Nous nous intéressons toujours à la possible présence d'un effet de regulatory-fit entre l'humain et l'agent au cours de l'interaction, c'est-à-dire que l'agent et l'interaction sont perçus comme plus appréciables lorsque le focus de l'interaction et la personnalité de l'agent sont orientés dans la même direction. Pour tester cet effet de fit, nous avons précedemment considéré que les participants percevaient l'interaction au travers de leur propre focus chronique. Cependant, l'utilisation du focus chronique des participants a présenté certaines limites, comme celle de la répartition des participants. En effet, le focus promotion est plus répandu dans les sociétés occidentales, culturellement plus individualistes et mettant en avant la réussite par la compétition (Uskul et al., 2009). De plus, notre population étant jeune, la

mesure en auto-attribution du regulatory-focus risque d'être moins cohérente que celle des adultes (Soto et al., 2008) et donc plus difficile à exploiter.

Pour ces raisons, nous choisissons ici d'induire les différents focus lors des différentes sessions de jeu. Le but de cette induction est de créer un contexte promotion-focus (resp. prévention-focus) pour l'ensemble des participants et d'étudier la perception des agents dans ces différents contextes afin de tester l'effet fit. L'idée est donc de voir si un agent avec un focus promotion (resp. prévention) est plus apprécié dans un contexte promotion (resp. prévention) que dans un contexte prévention (resp. promotion). Pour mettre en place cette induction de façon la moins intrusive possible tout en impliquant nos jeunes participants, nous présenterons l'expérimentation sous forme d'un tournoi, avec un système de points. Même si l'expérimentation reste relativement contrôlé, cet habillage offre un aspect plus écologique. Le protocole est présenté plus avant en section 6.4.

6.2 méthodologie

6.2.1 *Implémentation du comportement des agents*

6.2.1.1 *Implémentation des stratégies de jeu*

Concernant l'implémentation des stratégies de jeu, nous ré-utilisons la même approche data-driven basée sur des Alternating Decision Trees (ADTrees). Nous apportons cependant des modifications concernant les données et la façon d'utiliser celles-ci pour l'apprentissage, décrites ci-dessous.

utilisation des scores de regulatory-focus Dans l'implémentation précédente, nous avons utilisé les scores promotion et prévention des individus comme une caractéristique pour la classification. Il s'agit donc d'une approche comparable à l'utilisation d'un trait, où la valeur même du paramètre va faire varier la décision dans une branche de l'arbre. Cette méthode nous a permis de représenter des stratégies prenant en compte le regulatory-focus mais elle ne permet pas de mettre en regard les caractéristiques spécifiques utilisées par des personnes orientées promotion par rapport à celles utilisées par des personnes orientées prévention.

Pour essayer de répondre à cette problématique, nous avons séparé notre population en deux, en fonction du focus chronique des individus (promotion ou prévention). Nous avons alors entrainé indépendamment nos classifieurs sur chacune de ces populations. En conséquence, nous obtenons 4 arbres de décision : 2 arbres représentant les comportements promotion (l'un pour le choix d'un mouvement

et l'autre pour le choix stop-ou-encore) et de la même façon, 2 arbres représentant les comportements prévention.

données d'apprentissage Pour apprendre les stratégies, nous utilisons les données de jeu des joueurs humains de l'expérimentation précédente (chapitre 5), soit 80 parties de Can't Stop. 6 parties n'ont pas pu être exploitées pour l'apprentissage suite à des problèmes techniques. Parmi ces 74 parties, 23 parties correspondaient à des personnes avec un focus chronique prévention et 51 à des personnes avec un focus chronique promotion. Nous avons également ajouté des caractéristiques à notre jeu de données, principalement pour mieux prendre en compte la différence entre des caractéristiques propres aux colonnes impliquées dans le tour en cours et des caractéristiques calculées sur l'ensemble du plateau.

Pour le choix du mouvement, une instance du jeu de données est composée du mouvement effectivement choisi (Chosen Move, CM), comparé avec un des mouvements alternatifs possibles (Alternative Move, AM), des caractéristiques calculées pour cette paire et de la classe-cible "Choisi" ou "NonChoisi". Chaque paire fait l'objet de deux instances, l'une de la forme (CM,AM) avec pour classe-cible "Choisi" et l'autre de la forme (AM,CM) avec pour classe-cible "NonChoisi".

Pour la décision "stop-ou-encore", une instance du jeu de données est composée des caractéristiques calculées au moment de la décision et d'une classe-cible "Continue" si le joueur a choisi de rejouer ou "Stop" si le joueur a choisi de s'arrêter.
Au final, les jeux de données utilisés étaient de la forme suivante :

1. Jeux de données pour le choix du mouvement
   - 155 caractéristiques
   - Population promotion : 3754 instances
   - Population prévention : 1754 instances
2. Jeux de données pour le choix stop-ou-encore
   - 88 caractéristiques
   - Population promotion : 1525 instances
   - Population prévention : 703 instances

profondeur des arbres Pour obtenir la même qualité de représentation de nos données que dans l'expérimentation précédente et affiner la complexité des stratégies qui seront utilisées par nos agents, le paramètre *boosting iterations* de l'algorithme de classification est défini à 25.

Nous avons ensuite implémenté ces arbres de décision afin de créer les stratégies de nos agents RF-Pro et RF-Pre.

6.2.1.2 *Implémentation d'un modèle émotionnel*

Notre agent virtuel exprime une émotion par le biais d'une expression faciale dans le cadre des 7 événements principaux du jeu :

A. Pendant le tour du joueur humain :
   - après le résultat du lancer de dés,
   - après le choix du mouvement par le joueur,
   - après le choix stop-ou-encore ou la perte de tour du joueur.

B. Pendant son propre tour
   - après le résultat du lancer de dés,
   - après le choix du son mouvement,
   - après son choix stop-ou-encore (s'il s'arrête ou s'il perd son tour).

C. À la fin de la partie.

utilisation du modèle cpm Pour calculer les expressions faciales exprimées par notre agent virtuel, nous nous inspirons du modèle CPM (Componential Process Model) de Scherer (2001). D'un point de vue théorique, ce modèle est intéressant car il met l'accent sur la composante cognitive de l'émotion. Ceci nous permet de créer deux processus distincts pour représenter nos deux focus promotion et prévention. (Courgeon et al., 2014) ont également utilisé ce modèle pour modéliser le processus émotionnel dans le cadre d'un jeu de plateau. Cette modélisation était spécifique au jeu utilisé. Par contre, la plateforme MARC propose une fonctionnalité générique, qui permet d'envoyer directement des valeurs de checks (avec un formalisme EML (*Emotional Markup Language*)) qui sont alors utilisées pour créer une animation faciale dynamique séquentielle. Cette fonctionnalité est donc facilement réutilisable mais le processus émotionnel doit être ré-implémenté pour notre cas applicatif et de façon à intégrer la notion de regulatory focus.

Le modèle CPM définit l'émotion comme le résultat d'un processus d'évaluation cognitive : pour réagir face à un évènement, différents critères relatifs à l'évènement sont évalués pour déterminer l'émotion résultante. Le fait que l'accent soit mis sur la composante cognitive de l'émotion est très intéressante car ceci nous permet de créer deux processus distincts pour représenter nos deux focus promotion et prévention. Dans le modèle CPM, les critères d'évaluation sont nommés Stimulus Evaluation Checks (SECs) ou plus simplement, checks (Scherer and Sangsue, 2004). Ces checks peuvent être regroupés en 4 grandes catégories : la pertinence, le rapport aux buts, le potentiel de maîtrise et l'accord avec les standards. Plusieurs checks sont ainsi définis pour ces différentes catégories. En plus de proposer un fonctionnement de l'évaluation cognitive ayant lieu au cours du processus émotionnel, le modèle CPM met en regard l'évaluation de ces checks avec l'expressivité, en particulier celle du visage (Scherer, 2001).

Parmi les checks implémentés dans MARC, nous choisissons d'utiliser un sous-ensemble de 4 éléments qui sont les plus pertinents dans le cadre d'un jeu et qui sont calculables à partir des éléments disponibles de la situation. Les checks choisis sont les suivants :

1. Nouveauté : l'événement a-t-il été anticipé ? Ce check prend ses valeurs entre -1 (événement attendu) et 1 (événement inattendu)
2. Agrément intrinsèque : l'événement est-il plaisant en lui-même ? Ce check prend ses valeurs entre -1 (événement déplaisant) et 1 (événement plaisant)
3. Rapport aux buts : cet événement va-t-il dans le sens de mes buts ? Ce check prend ses valeurs entre -1 (événement entravant la réalisation de mes buts) et 1 (événement facilitant la réalisation de mes buts)
4. Potentiel de maîtrise : suis-je capable de faire face ou de m'adapter à cet évènement ? Ce check prend ses valeurs entre -1 (impossibilité de faire face) et 1 (totalement gérable)

Une fois les valeurs de checks calculées, celles-ci sont envoyées à la plateforme MARC sous forme d'EML via une connexion UDP et la plateforme génère une animation temps réel en fonction de ces valeurs.

modélisation des checks Pour définir les algorithmes permettant de calculer chacun de ces checks, nous nous appuyons à la fois sur la définition théorique du regulatory-focus mais également sur les arbres de décision que nous avons calculés. En effet, si nous partons du principe que les caractéristiques du jeu sélectionnées pour la stratégie promotion (resp. prévention) sont des éléments importants pour les personnes présentant un tel focus, il semble pertinent d'utiliser les mêmes caractéristiques pour déterminer l'évaluation cognitive d'un événement du jeu. Nous interprétons donc les arbres de décision promotion et prévention dans le cadre de la théorie du regulatory focus pour guider notre implémentation des différents checks. D'une part, il semble que l'arbre promotion fasse ressortir des caractéristiques liées à la position du joueur relativement au sommet des colonnes ("est-ce que je m'approche du gain de telle ou telle colonne ?"). Nous pouvons l'interpréter comme un intérêt pour le gain d'une colonne, dans l'absolu. D'autre part, l'arbre prévention utilise plus de caractéristiques liées à la position du joueur certes par rapport au sommet des colonnes mais ceci relativement à la position de l'adversaire ("suis-je en train de perdre du terrain par rapport à mon adversaire ?"). Nous pouvons l'interpréter comme un intérêt non pas pour le gain en lui-même, mais pour la non-perte par rapport à l'adversaire. À titre d'exemples, les calculs de 4 checks sont présentés dans le Tableau 12 et le Tableau 13.

Tableau 12 – Présentation des choix de modélisation d'une partie des checks du modèle émotionnel durant le tour de l'humain

Check	Différence pre/pro	Calcul	Explication	Basé sur
Après le résultat des dés pour l'humain				
Rapport aux buts	Calcul différent	Prévention		Interprétation de la théorie du regulatory-focus
		fonction du nombre de mouvements possibles sur des colonnes où il y a des camps de l'agent	Si l'humain a la possibilité d'avancer sur les mêmes colonnes que l'agent, c'est une perte potentielle pour l'agent	
		Promotion		
		fonction du nombre de mouvements possibles sur des colonnes où il y a des camps dans le dernier tiers de la colonne	Si l'humain a la possibilité d'avancer sur une colonne bien avancée, c'est un potentiel non-gain pour l'agent	
Après le choix d'un mouvement par l'humain				
Nouveauté	Même principe	fonction du score pour chaque mouvement possible (sortie de l'arbre de décision propre au focus de l'agent)	Le choix de l'humain est d'autant plus "nouveau" qu'il est éloigné de ce que l'agent aurait lui-même choisi	Utilisation des arbres de décision

Tableau 13 – Présentation des choix de modélisation d'une partie des checks du modèle émotionnel durant le tour de l'agent

Check	Différence pre/pro	Calcul	Explication	Basé sur
Après le résultat des dés pour l'agent				
Nouveauté	Aucune	fonction de la probabilité des combinaisons (sur une fenêtre glissante de 4 coups)	Plus les combinaisons de dés sont improbables, plus l'événement est "nouveau"	Logique
Après le choix d'un mouvement par l'agent				
Rapport aux buts	Calcul différent	Prévention		Interprétation des arbres de décision
		fonction de la probabilité de pouvoir rejouer au coup suivant et de la qualité de placement* par rapport au début du tour	Plus la probabilité de pouvoir rejouer est grande et plus l'agent améliore la qualité de son placement, plus la situation satisfait le but de non-perte de l'agent	
		Promotion		
		fonction du reste à grimper sur les colonnes jouées par rapport au début du tour	Plus l'agent se rapproche du haut des colonnes, plus la situation satisfait le but de gain de l'agent	

* La qualité de placement est définie comme le reste à grimper pondéré par la proximité de l'adversaire

comportement non-verbal de l'agent virtuel Le comportement non-verbal de l'agent virtuel est finalement défini de la façon suivante : lorsque les dés sont lancés, l'agent tourne la tête et regarde la partie inférieure droite de l'écran où se situe l'affichage des dés et des mouvements. Lorsque l'affichage des mouvements est terminé, l'agent relève la tête et réalise une expression faciale. Lors du tour de jeu de l'humain, l'agent réalise une expression faciale après chaque décision prise par le joueur (choix du mouvement / décision stop-ou-encore). Lors de son tour de jeu, l'agent prend quelques secondes avant de choisir un mouvement / décider de continuer ou non son tour. Il lève alors légèrement la tête et oriente son regard vers le haut afin de mimer la réflexion. Il revient ensuite face caméra et réalise une expression faciale. Une expression faciale est également jouée lorsqu'un des deux joueurs gagne la partie. Certains de ses comportements sont illustrés en Figure 19.

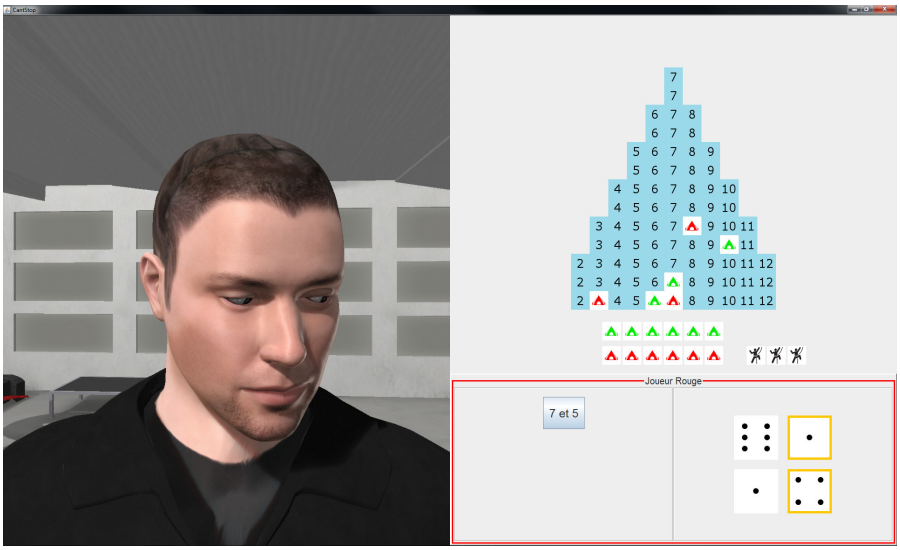
6.2.1.3 *Résumé de l'implémentation*

Nous avons donc réalisé une application permettant à un utilisateur de jouer à Can't Stop contre un agent artificiel, avec ou sans agent virtuel. La prise de décision a été implémentée de manière data-driven, à l'aide d'un classifieur (ADTree) entrainé sur des données de jeu humaines. Le modèle émotionnel a été implémenté de façon mixte (theory-driven et data-driven) : les algorithmes de calculs de checks CPM ont été définis en s'appuyant sur la théorie du regulatory-focus et sur l'interprétation des ADTrees générés pour la prise de décision.

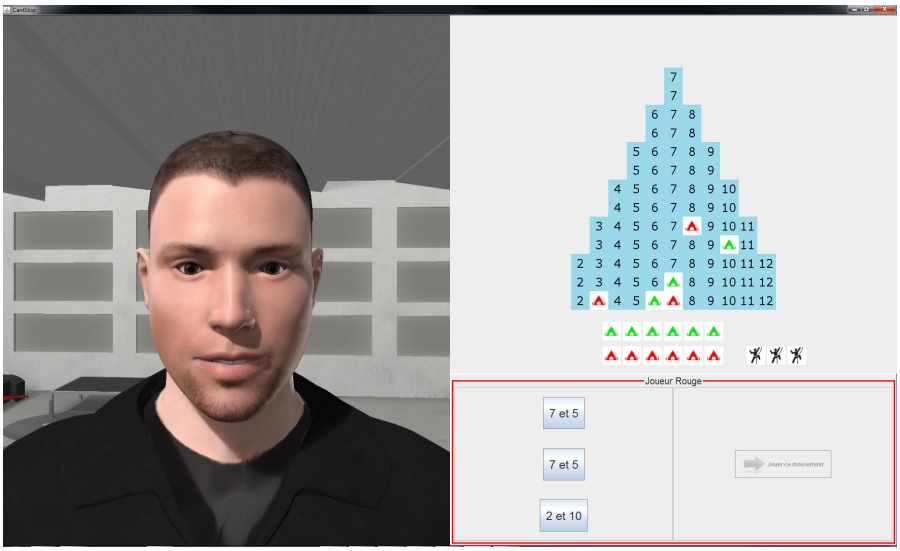
L'architecture de l'application est représentée en Figure 20. L'application génère des événements à certains moments du jeu (e. g.action de l'utilisateur, décision à prendre par l'agent...). Le cas échéant, les informations concernant la situation de jeu sont alors transmises au module de prise de décision de l'agent, qui prend une décision en fonction de l'ADTree propre à son regulatory-focus. Lorsqu'un agent virtuel MARC est présent, les informations concernant la situation de jeu sont transmises au module émotionnel qui calculent les valeurs des checks CPM en fonction des règles propre à son regulatory-focus. Les checks calculés sont alors envoyés à la plateforme MARC qui se charge de réaliser une animation faciale en temps réel en fonction de ces valeurs.

6.2.2 *Choix d'un agent "contrôle"*

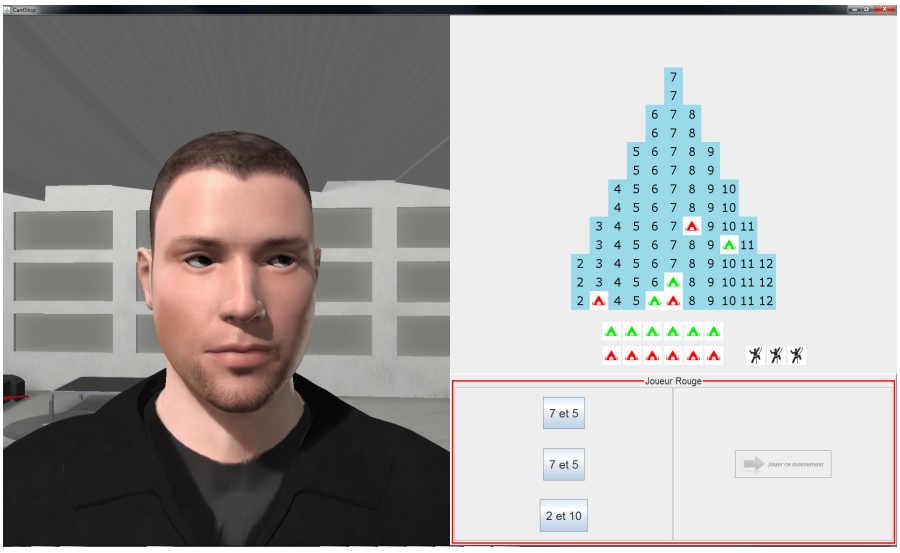
La question sur le choix d'un agent "contrôle" pertinent reste ouverte. Dans l'étude précédente, nous avions utilisé un agent jouant de façon aléatoire (agent Random), ainsi qu'un agent jouant avec une stratégie "dépersonnalisée" (agent Moyen). Les résultats ainsi obtenus montrent qu'une personnalité sera perçue chez l'agent, que cette personnalité ait été implémentée à dessein ou non. Par ailleurs, ces deux



(a) L'agent MARC observe l'affichage des mouvements

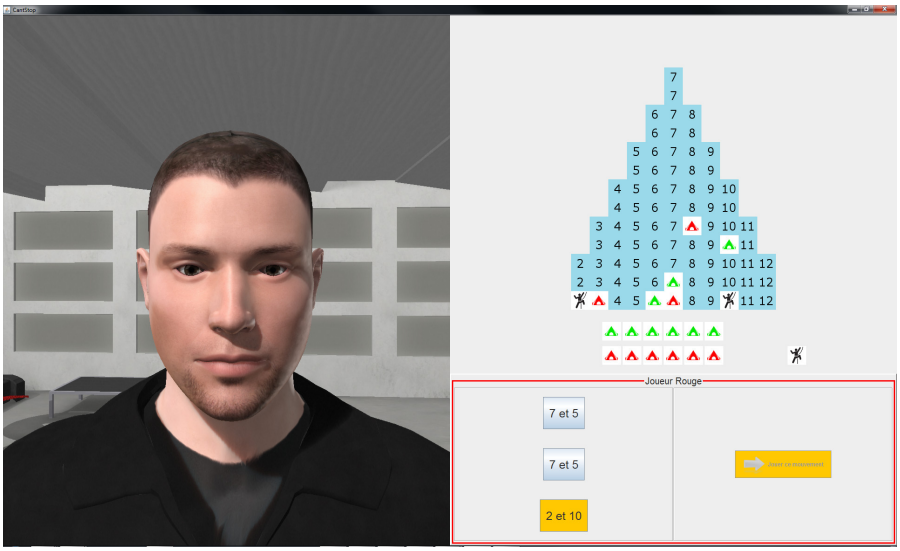


(b) L'agent MARC réalise une expression faciale après l'affichage des mouvements

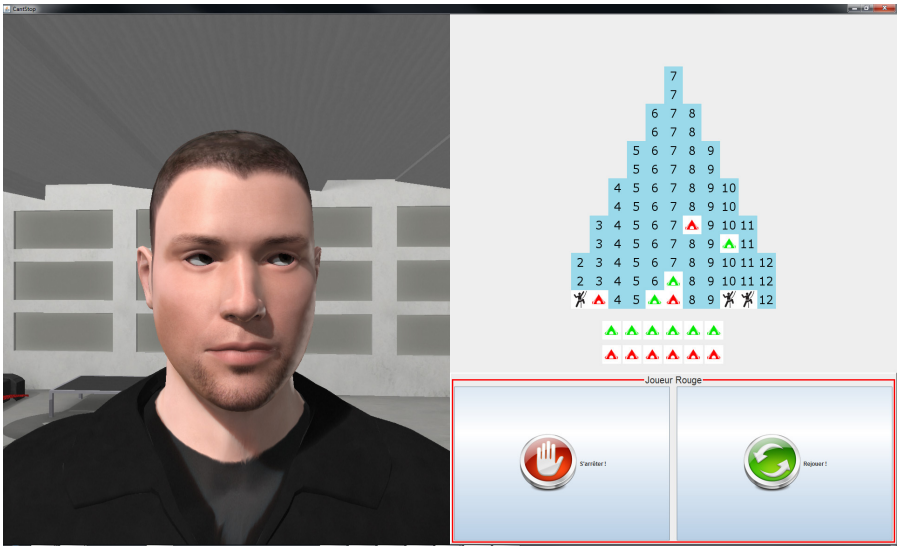


(c) L'agent MARC réfléchit au choix du mouvement

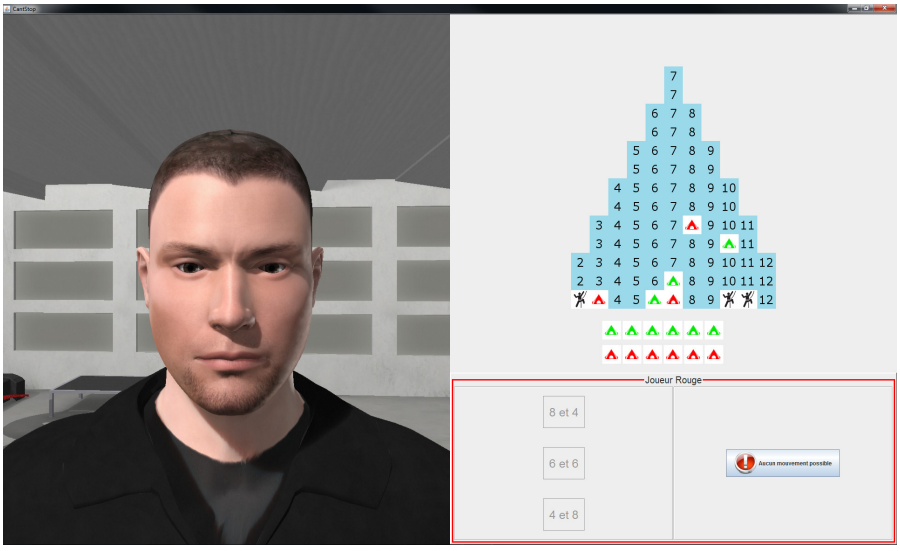
Figure 19 – Comportements non-verbaux de l'agent virtuel durant une phase de jeu (1/2)



(d) L'agent MARC choisit un mouvement



(e) L'agent MARC réfléchit au choix stop-ou-encore



(f) L'agent MARC réalise une expression faciale après avoir perdu son tour

Figure 19 – Comportements non-verbaux de l'agent virtuel durant une phase de jeu (2/2)

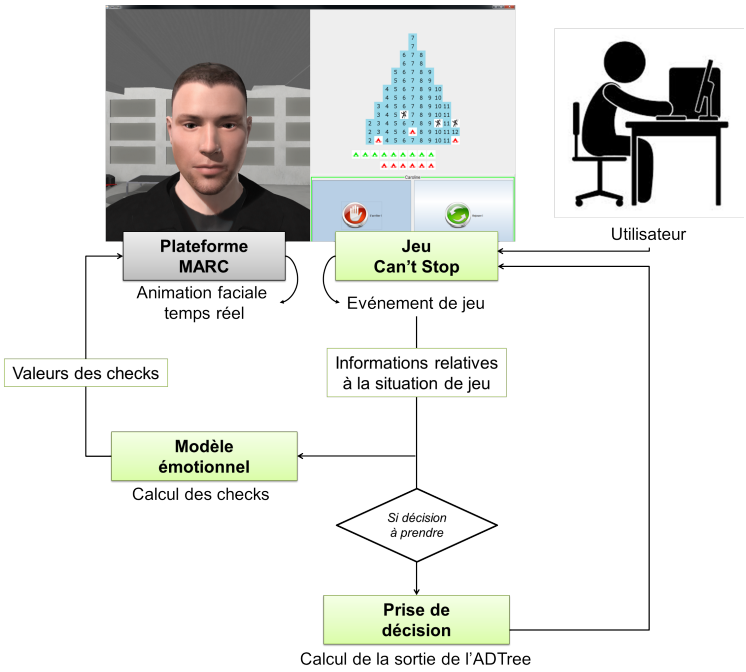


Figure 20 – Architecture de l'application Can't Stop (en vert) en interaction avec la plateforme MARC

agents ont montré des stratégies inférieures en termes de qualité de jeu (cf. le pourcentage de parties gagnées). L'implémentation des stratégies de jeu "sans personnalité" crée donc une différence en termes de compétence de l'agent par rapport au jeu. Cette caractéristique également peut influencer la perception de l'agent et donc être une variable parasite dans notre protocole. D'un point de vue méthodologique, la question des caractéristiques d'un bon agent "contrôle" reste ouverte.

Il est cependant intéressant (pour ne pas dire nécessaire) de comparer les résultats obtenus pour nos agents avec regulatory-focus aux résultats d'agents ne possédant pas cette implémentation spécifique. Afin de limiter le nombre de conditions expérimentales, nous n'utiliserons ici qu'un seul agent "contrôle".

Nous ne retenons pas l'agent Random afin de maintenir l'intérêt de nos participants pour le jeu. En effet, nous mettons en place des sessions répétées, il faut donc maintenir l'engagement des participants sur la durée. La stratégie de l'agent Random ne présente que peu d'intérêt et pourrait lasser certains joueurs. De plus, les participants jouant avec un agent Random auraient une plus forte probabilité de gagner leurs parties (95% de parties gagnées par l'humain contre l'agent Random lors de l'étude précédente). L'expérimentation est présentée comme un tournoi et les participants peuvent communiquer entre eux. Certains participants pourraient donc se sentir lésés

si cette différence est perçue et se désengager en réponse à ce déséquilibre. Cette limite s'applique aussi à l'agent Moyen. Même s'il a été plus performant que l'agent Random, l'agent Moyen a été un adversaire peu stimulant (32% de parties gagnées par l'agent). De plus, nous proposons ci-dessous une nouvelle implémentation des stratégies de jeu pour nos agents RF-Pro et RF-Pre, moins stéréotypique. Ces agents devraient donc maintenant être plus proches d'une personne réelle, avec une personnalité plus subtile et donc plus "moyenne". L'agent Moyen était une base de comparaison intéressante par rapport à nos agents précédents, à la personnalité extrême. Par rapport à nos nouveaux agents, l'aspect "moyen" de l'agent contrôle perd de son sens et de son intérêt.

Nous avons souligné l'importance de la cohérence des comportements au cours du temps comme un des intérêts de l'ajout d'une personnalité. Nous proposons donc d'explorer une nouvelle piste en utilisant comme agent contrôle un agent versatile, changeant. Sa stratégie sera raisonnée, afin d'avoir un adversaire stimulant, mais sa stratégie ne présentera pas la même stabilité que celle des agents Promotion et Prévention. Cet agent contrôle présente plusieurs intérêts pour nous. D'une part, il est intéressant de voir comment, en comparaison des agents RF-Pre et RF-Pro, des changements de "personnalité" au cours du jeu peuvent affecter la perception de l'agent comme de l'interaction. D'autre part, cet agent nous permet d'explorer l'importance de la stabilité implémentée au travers de deux modalités d'expression de la personnalité.

Cet agent sera désigné par le suite par le terme "Agent Switch". Concernant l'agent Switch, afin d'obtenir l'aspect changeant évoqué plus haut, celui-ci choisira de façon aléatoire d'utiliser la stratégie promotion ou la stratégie prévention, et ce à chaque décision qu'il aura à prendre. Autrement dit, lorsque l'application enverra un événement de jeu pour qu'une décision soit prise par l'agent, l'agent Switch calculera la sortie de l'ADTree Promotion ou de l'ADTree Prévention, ADTree choisi aléatoirement à chaque décision. Lorsqu'un agent virtuel sera présent, son expression faciale sera déterminée de la même manière que pour sa stratégie de jeu. À chaque expression qu'il devra faire, il choisira aléatoirement les calculs de checks Promotion ou Prévention.

6.3 hypothèses générales

Pour cette étude, en nous basant sur notre implémentation et notre expérience précédente, nous formulons les hypothèses générales suivantes.

6.3.1 *Concernant la perception de l'agent*

- H1 : Les deux regulatory-focus implémentés sont perçues comme tels par les utilisateurs.
- H2 : La présence visuelle d'un agent virtuel améliore la perception de la personnalité grâce à la cohérence des informations fournies (stratégies de jeu et expressions faciales).
- H3 : L'implémentation d'un regulatory-focus cohérent apporte plus de crédibilité aux agents.
- H4 : Il existe un effet de regulatory fit lorsque la personnalité de l'agent et le contexte induit sont les mêmes.

6.3.2 *Concernant la perception de l'interaction*

- H5 : L'interaction est mieux perçue lorsque l'agent a une personnalité cohérente.
- H6 : Un agent virtuel amoindrit le sentiment d'interagir uniquement "avec un ordinateur".

6.3.3 *Concernant la relation avec l'agent*

- H7 : Les sessions répétées permettent aux participants de se construire une image de la personnalité de l'agent allant plus loin que la formation d'impression.
- H8 : Les sessions répétées permettent aux participants de commencer à construire une relation avec leur agent.

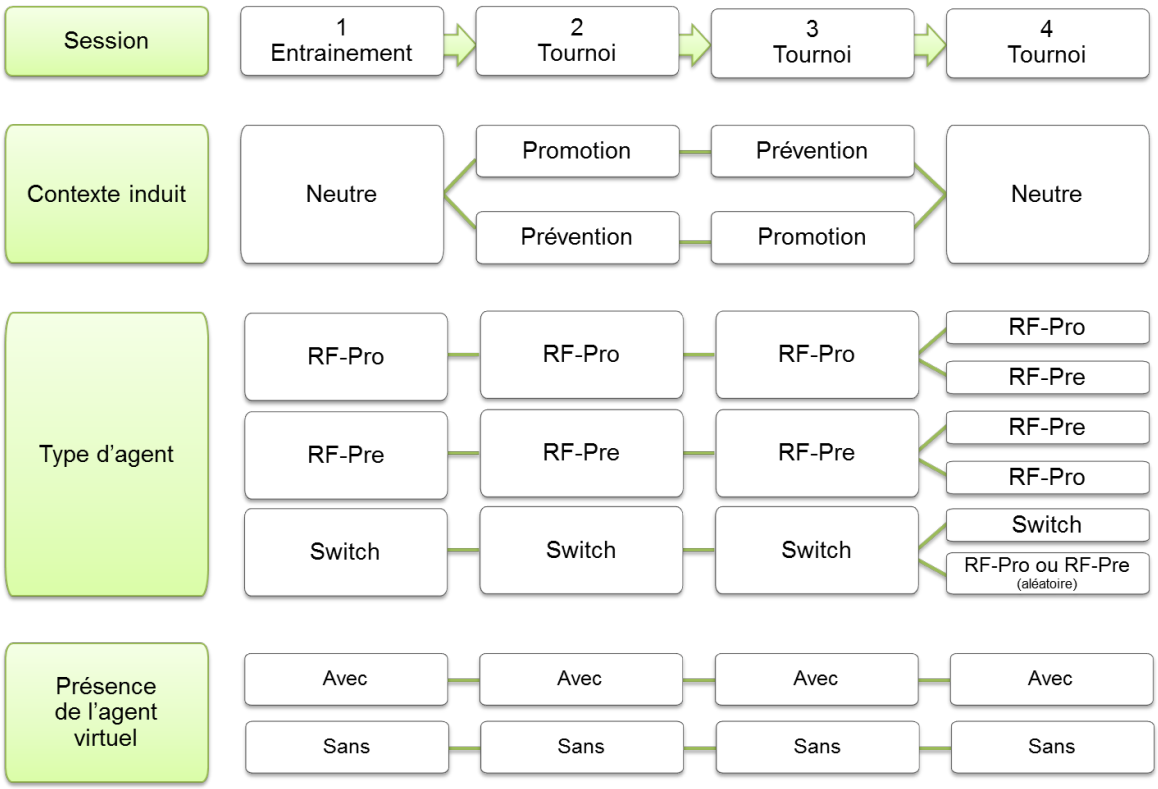
6.4 protocole expérimental

Notre design expérimental est résumé par la Figure 21. Les variables indépendantes de notre protocole et leurs variations au cours des sessions sont présentées en Figure 21a. Les variables dépendantes en fonction des sessions sont présentées en Figure 21b. Les mesures, les participants et le déroulement temporel du protocole sont détaillés ci-après.

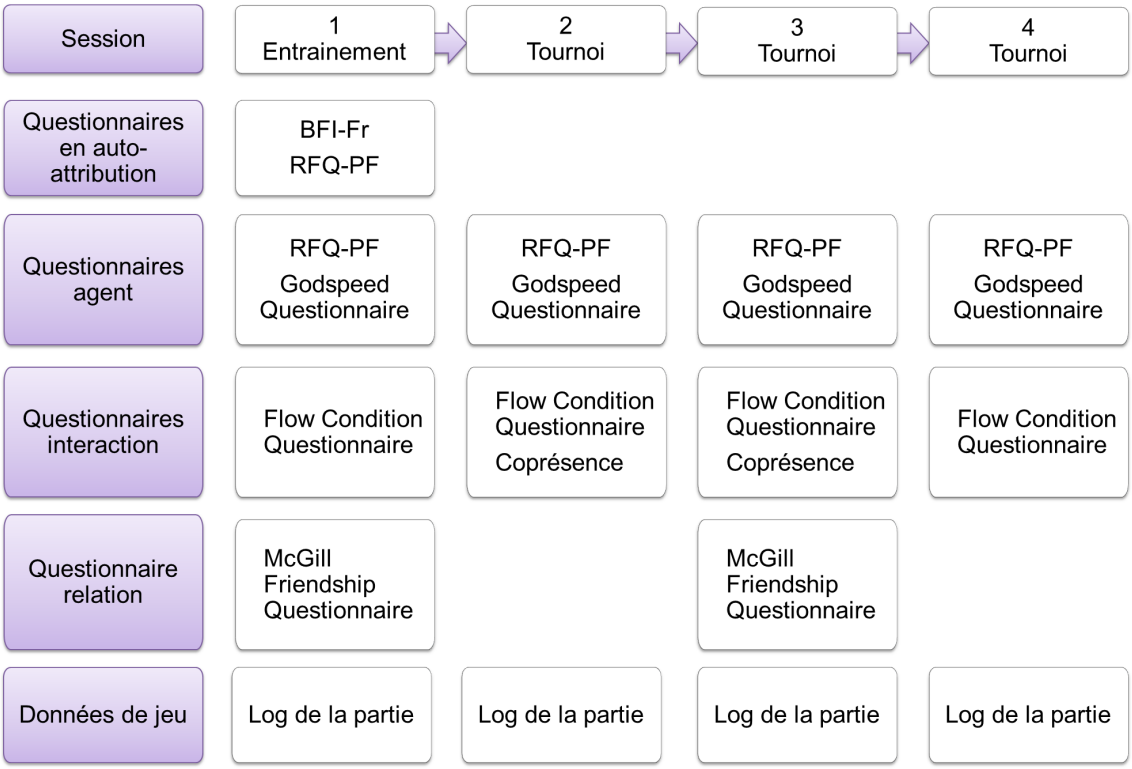
6.4.1 *Mesures*

Lors de cette expérimentation, les mesures suivantes ont été utilisées.

questionnaires en auto-attribution Les participants ont répondu pour eux-mêmes à deux questionnaires de personnalité : le RFQ-PF, afin de mesurer leur focus chronique et le BFI-Fr (Plaisant et al., 2010), mesurant les cinq dimensions de la personnalité telles



(a) Variables indépendantes



(b) Variables dépendantes

Figure 21 – Conditions (variables indépendantes) et mesures (variables dépendantes) de notre protocole à interactions répétées

que définies par le modèle du Big 5. Le RFQ-PF utilise une échelle de Likert à 7 points. Le BFI-Fr utilise une échelle de Likert à 5 points.

questionnaires en hétéro-attribution Les participants ont répondu à des questionnaires afin de décrire leur perception de l'agent avec lequel ils interagissaient. Concernant la perception du regulatory-focus de l'agent, nous avons utilisé le *RFQ-PF*. Concernant la crédibilité, nous avons utilisé les même questions que dans l'étude précédente, soit 10 questions du *Godspeed Questionnaire* (Bartneck et al., 2009) (5 questions à propos de l'appréciation de l'agent et 5 questions à propos de l'intelligence perçue). Le Godspeed Questionnaire utilise une échelle de Likert à 5 points.

questionnaires concernant l'interaction Les participants ont répondu à différents questionnaires pour évaluer leur perception de l'interaction : une *échelle de coprésence*, proposé par Bailenson et al. (2006), composé de 4 questions et le *Flow Condition Questionnaire* (Schaffer, 2013), composé de 3 questions. Ces deux questionnaires utilisent une échelle de Likert à 5 points.

questionnaire concernant la relation Afin d'essayer de mesurer l'évolution de la relation entre les participants et leurs agents, nous avons utilisé le *McGill Friendship Questionnaire* (Mendelson and Aboud, 1999). Ce questionnaire permet de mesurer à quel degré une personne (désignée lors de la consigne) remplit les six fonctions de l'amitié définies par les auteurs. Ces six fonctions sont :

- la Stimulation (*Stimulating companionship*), qui réfère au fait de faire des choses sympathiques et amusantes ensemble,
- l'Aide (*Help*), qui réfère au fait que l'ami apporte des conseils, des informations, des biens physiques ou tout autre forme d'aide permettant de satisfaire des besoins ou des buts du répondant,
- l'Intimité (*Intimity*), qui réfère au fait que l'ami offre un contexte où le répondant peut exprimer ouvertement et sincèrement des informations personnelles,
- la Fiabilité (*Reliable Alliance*), qui réfère au fait que le répondant peut compter sur la disponibilité et la loyauté d'un ami, et ce durablement dans le temps,
- la Self Validation (*Self-Validation*), qui réfère au fait que le répondant perçoit l'ami comme quelqu'un qui est à l'écoute, qui rassure, encourage, ce qui permet au répondant de maintenir une bonne image de lui-même.
- la Sécurité Emotionnelle (*Emotional Security*), qui réfère au fait que l'ami offre au répondant réconfort et confiance dans des situations nouvelles ou anxiogènes.

Ce questionnaire comporte 30 questions et a déjà été utilisé dans le cadre d'interactions homme-machine (Leite et al., 2010; Petisca et al., 2015). Nous avons mis en place la consigne suivante : "Imaginez que votre agent virtuel soit disponible dans votre vie quotidienne et qu'il ait plus de capacités. Dans ce cas, je pense qu'il pourrait...". Ce questionnaire utilise une échelle de Likert à 8 points.

données de jeu L'ensemble des parties jouées par les participants étaient enregistrées sous forme de fichiers de log.

6.4.2 *Participants*

Cette étude a eu lieu au collège Alain Fournier, à Orsay (91400), avec les 5 classes de 4ème en collaboration avec la professeur de technologie. 126 élèves ont accepté de participer (66 filles et 60 garçons). Un formulaire de consentement a également été signé par leurs parents. L'âge moyen des participants est de 13 ans (ecart-type = 0,4). Pour les passations, les élèves étaient répartis en groupes de 11 à 15 personnes.

6.4.3 *Protocole*

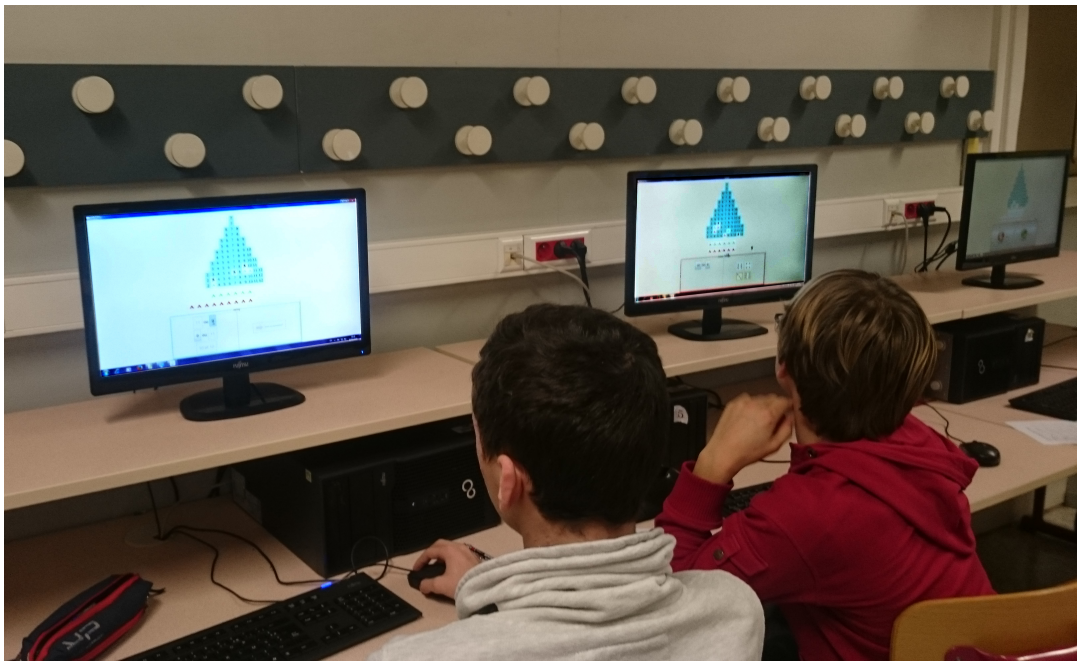
L'étude a été présentée aux participants comme un tournoi de Can't Stop se déroulant en 4 sessions. La durée entre les sessions était de 3 à 7 jours. Le tournoi se jouait en équipe, chaque équipe étant composée d'un participant et de son agent. Un participant garde le même agent durant tout le tournoi. Le participant jouait des parties de Can't Stop contre son agent, mais des points étaient attribués à l'équipe en fonction des actions des deux membres (participant et agent). L'agent artificiel était donc présenté comme étant à la fois un partenaire et un adversaire. La moitié des participants se voyaient attribuer un partenaire visible, i. e. avec un agent virtuel (Figure 22a) et l'autre moitié un partenaire "invisible", i. e. sans agent virtuel (Figure 22b).

6.4.3.1 *Déroulement des sessions*

session 1 Lors de la première session, l'expérimentateur expliquait le principe général du tournoi aux participants. Suite à cette présentation, les participants commençaient par répondre aux deux questionnaires de personnalité (RFQ-PF et BFI-Fr) en auto-attribution ainsi qu'aux questions relatives aux données démographiques. L'expérimentateur pouvait répondre aux questions concernant le vocabulaire. Une fois ces questionnaires achevés, les participants regardaient un tutoriel vidéo leur apprenant les règles du jeu Can't Stop [1]. Ils pouvaient naviguer dans la vidéo et/ou solliciter l'expérimentateur

1. Vidéo disponible à l'adresse : https ://www.youtube.com/watch ?v=kQs6JBXotxw

(a) Participants jouant avec un agent virtuel MARC



(b) Participants jouant avec un agent artificiel sans représentation physique

Figure 22 – Participants jouant à Can't Stop contre un agent durant une session expérimentale

en cas de questions. Après le tutoriel, les participants jouaient une première partie contre leur agent. Cette partie était présentée comme un entraînement afin de bien comprendre le jeu et de commencer à découvrir son partenaire artificiel. Après cette partie, le participant répondait aux questionnaires concernant sa perception de l'agent et de l'interaction ainsi qu'au McGill Friendship Questionnaire.

sessions 2 et 3 Lors des deuxième et troisième sessions, l'expérimentateur expliquait le système de points en vigueur pour la session. Ces deux sessions correspondent aux inductions des deux regulatory-focus. Les deux inductions ont été contrebalancées. La moitié des groupes effectuait l'induction Promotion en session 2 et l'induction Prévention en session 3. L'autre moitié effectuait l'induction Prévention en session 2 et l'induction Promotion en session 3. L'induction Promotion (resp. Prévention) était effectuée à l'aide des points attribués aux équipes, qui mettaient en avant l'aspect gain/non-gain (resp. perte/non-perte) et encourageaient les participants à avoir une stratégie plus intrépide (resp. plus prudente). Après avoir joué sa partie, le participant répondait aux questionnaires concernant sa perception de l'agent et de l'interaction. Uniquement lors de la session 3, le participant répondait à nouveau au McGill Friendship Questionnaire. Ce questionnaire n'est pas rempli lors de la deuxième session pour laisser le temps aux participants d'interagir avec l'agent avant de donner de nouveau leur point de vue sur la relation.

session 4 La quatrième session ne comporte pas d'induction mais permet de tester si les participants sont capables de "reconnaître" leur partenaire de jeu, avec lequel ils ont interagi lors des 3 précédentes sessions. Cette session était présentée comme une session "mystère". On annonçait aux participants que la moitié des joueurs avaient changé de partenaire artificiel. Ce changement n'était pas visible au niveau de l'interface pour les participants dont le partenaire était visible (même agent virtuel) mais l'expérimentateur expliquait que la moitié des participants jouaient malgré tout avec un partenaire différent de d'habitude. Le système de points était adapté pour essayer de n'induire aucun focus particulier et de pousser les participants à répondre sérieusement lorsqu'on leur demanderait s'ils pensaient avoir changé d'agent. Après avoir joué sa partie, le participant répondait aux questionnaires concernant sa perception de l'agent et de l'interaction, ainsi qu'aux questions "Je pense avoir changé de partenaire de jeu cette semaine" (réponse Oui/Non) et "À quel point êtes-vous certain d'avoir ou de ne pas avoir changé de partenaire" (échelle de Likert à 5 points).

6.4.3.2 *Consignes*

Les consignes étaient énoncées sous forme de différents systèmes de points à gagner lors du tournoi. Ces systèmes étaient les suivants :

- Induction Promotion
  - +1 point si l'humain gagne la partie
  - 0 point si l'agent gagne la partie
  - Points bonus : autant de points en plus que de lancers de dés dans le tour le plus long de l'humain *ou* de l'agent
- Induction Prévention
  - -1 point si l'humain perd la partie
  - 0 point si l'agent perd la partie
  - Points malus : autant de points en moins que de décrochages (tours perdus) pour l'humain *et* pour l'agent
- Semaine "mystère"
  - +1 point si l'humain gagne la partie
  - -1 point si l'humain perd la partie
  - +2 points pour une bonne réponse à la question "Je pense avoir changé de partenaire"
  - -2 points pour une mauvaise réponse à la question "Je pense avoir changé de partenaire"

6.4.4 *Hypothèses opérationnelles*

Nous formulons donc les hypothèses opérationnelles suivantes.

6.4.4.1 *Concernant la perception de l'agent*

- *H1 : Les deux regulatory-focus implémentés sont perçues comme tels par les utilisateurs*
  - Les agents RF-Pro ont un haut score promotion
  - Les agents RF-Pre ont un haut score prévention
  - Les scores promotion et prévention des agents Switch sont les plus dispersés
  - Les participants interagissant avec des agents RF-Pro et RF-Pre sont plus capables de reconnaître leur agent lors de la quatrième session que ceux interagissant avec un agent Switch.

- *H2 : La présence visuelle d'un agent virtuel améliore la perception de la personnalité grâce à la cohérence des informations fournies (stratégies de jeu et expressions faciales)*
  - Les agents RF-Pro et RF-Pre avec MARC ont une inclination plus marquée que les agents RF-Pro et RF-Pre sans MARC
  - Les scores des agents Switch avec MARC sont plus dispersés que pour les agents Switch sans MARC

- *H3 : L'implémentation d'un regulatory-focus cohérent apporte plus de crédibilité aux agents*
  - Les agents RF-Pro et RF-Pre ont des scores plus hauts en appréciation et en intelligence perçue que les agents Switch.
- *H4 : Il existe un effet de regulatory fit lorsque la personnalité de l'agent et le contexte induit sont les mêmes*
  - En contexte Promotion, les agents RF-Pro ont des scores plus hauts en appréciation et en intelligence perçue.
  - En contexte Prévention, les agents RF-Pre ont des scores plus hauts en appréciation et en intelligence perçue.

6.4.4.2 *Concernant la perception de l'interaction*

- *H5 : L'interaction est mieux perçue lorsque l'agent a une personnalité cohérente*
  - Le flow est plus important avec un agent avec personnalité que sans.

- *H6 : Un agent virtuel amoindrit le sentiment d'interagir uniquement "avec un ordinateur"*
  - Le score de coprésence est plus important avec un agent virtuel que sans.

6.4.4.3 *Concernant la relation avec l'agent*

- *H7 : Les sessions répétées permettent aux participants de se construire une image de la personnalité de l'agent allant plus loin que la formation d'impression*
  - Les scores de personnalité sont plus nuancés au fur et à mesure des sessions : la différence entre le score promotion et le score prévention se réduit avec le temps.

- *H8 : Les sessions répétées permettent aux participants de commencer à construire une relation avec leur agent*
  - Les scores au questionnaire de McGill augmentent entre la session 1 et la session 3.

6.5 résultats

Les analyses statistiques ont été réalisées avec R version 3.2.4 Revised. Nous avons réalisé des analyses de la variance pour design mixte (mixed-design ANOVAs)[2] comme analyses principales, et des tests de Student[3] pour les comparaisons paire à paire (avec une correction de Holm pour l'évaluation de la p-value).

2. R package *ez*, version 4.3
3. R package *stats*, version 3.2.4

Cette section présente les résultats statistiques, qui seront discutés en regard des hypothèses à la section 6.6.

6.5.1 *Validité des échelles*

Les échelles que nous utilisons ne sont pas spécifiques à la tranche d'âge de nos participants et n'ont pas toutes fait l'objet d'une validation psychométrique. Pour ces deux raisons, nous avons vérifié la cohérence de ces échelles en calculant leurs coefficients alpha de Cronbach[4] et en effectuant des analyses factorielles[4] lorsque nécessaire.

rfq-pf en auto-attribution L'échelle promotion et l'échelle prévention du RFQ-PF en auto-attribution montrent une cohérence interne acceptable ($\alpha_{\text{Promotion}} = 0,71$; $\alpha_{\text{Prévention}} = 0,68$) mais l'analyse factorielle révèle une structure incohérente avec beaucoup d'items ayant une faible saturation ($< 0,4$) et/ou présentant des saturations croisées sur un autre facteur. De ce fait, les scores reportés par les participants pour ce questionnaire ne seront pas exploités.

rfq-pf en hétéro-attribution L'analyse factorielle du RFQ-PF en hétéro-attribution montre une structure en deux facteurs claire, les items promotion saturant sur un facteur et les items prévention saturant sur le deuxième. Deux items prévention (Pre03 et Pre09) ont cependant une saturation croisée sur le facteur promotion. Ces items ont été supprimés du calcul des scores. Les deux échelles ainsi constituées présentent une bonne cohérence interne ($\alpha_{\text{Promotion}} = 0,79$; $\alpha_{\text{Prévention}} = 0,78$).

godspeed questionnaire Les deux échelles utilisées présentent la structure factorielle attendue et ont toutes deux une très bonne cohérence interne ($\alpha_{\text{Agréabilité}} = 0,9$; $\alpha_{\text{Intelligenceperçue}} = 0.85$).

flow condition questionnaire L'échelle à trois items du Flow Condition Questionnaire ne présente pas de cohérence interne ($\alpha = 0,31$). Il apparaît que les deux premiers items ("J'ai trouvé cette partie stimulante" et "Je me suis senti à même de relever les challenges que j'ai pu rencontrer") sont correctement reliés entre eux ($r(437) = 0,43, p < 0,001$) mais sont indépendants du troisième item ("J'ai été facilement distrait, déconcentré durant la partie") (resp. $r(437) = 0,02, p = 0.72$ et $r(437) = 0,01, p = 0.67$). Les deux premiers items ont donc été moyennés pour calculer un score que nous nommons *Flow - Intérêt* et le troisième item est utilisé seul comme un score que nous nommons *Flow - Distraction*.

4. R package *psych*, version 1.5.8

coprésence L'échelle de coprésence présente une cohérence interne acceptable ($\alpha = 0,67$) et est conservée en l'état.

mcgill friendship questionnaire Le questionnaire concernant la relation entre le participant et son agent ne présente pas la structure en 6 facteurs attendue. En effet, les items correspondant au facteur Sécurité Emotionnelle ne saturent pas sur un facteur commun mais sont distribués entre différents facteurs. Ces items ne seront donc pas pris en compte. Ces résultats sont cohérents avec Mendelson and Aboud (1997), qui montrent également une structure en 5 facteurs. Concernant les facteurs restants, nous retirons deux items (1 pour le facteur Aide et 1 pour le facteur Fiabilité) dont la saturation est insuffisante. Les échelles résultantes ont une cohérence interne satisfaisante ($\alpha_{\text{Stimulation}} = 0,89$; $\alpha_{\text{Aide}} = 0.87$; $\alpha_{\text{Intimité}} = 0,85$; $\alpha_{\text{Fiabilité}} = 0,88$; $\alpha_{\text{Selfvalidation}} = 0,88$).

6.5.2 *Données de jeu*

Dans un premier temps, nous notons que les consignes ont été prises en compte par les participants. Le nombre moyen de lancers de dés par tour (par les participants) est plus important lors de l'induction Promotion que lors de l'induction Prévention ($F(1,112) = 10,4, p < 0,01, \eta^2 = 0,03$). Le nombres de décrochages (tour perdu par les participants) est plus faible lors de l'induction Prévention que lors de l'induction Promotion ($F(1,112) = 6,89, p < 0,01, \eta^2 = 0,03$).

Par ailleurs, il semble que les trois types d'agents aient eu des comportements plutôt équivalents en termes de performances dans le jeu. Les agents RF-Pre ont gagné 66% de leurs parties, les agents RF-Pro ont gagné 44% de leurs parties et les agents Switch 50%.

6.5.3 *Perception de l'agent artificiel*

Les résultats concernant la perception de l'agent artificiel sont présentés pour les trois premières sessions, en fonction du type d'agent, de la présence de l'agent virtuel et du contexte d'induction. La semaine d'entraînement est considérée comme une semaine neutre. 17 participants ont été retirés des analyses car ils n'ont pas participé à l'ensemble des trois sessions.

6.5.3.1 *Perception du regulatory focus*

Nous commencons par présenter les statistiques descriptives concernant les scores de regulatory focus des agents. Nous observerons la moyenne et l'écart-type, ainsi que le coefficient de variation des quartiles (CQV ; $(Q3 - Q1)/(Q1 + Q3)$), qui est une statistique de dispersion permettant des comparaisons (Bonett, 2006) et le coefficient de

Finn[5] (Finn, 1970), comme un indicateur de l'accord entre les participants. Ces deux derniers indicateurs sont complémentaires à l'observation des écarts-types et nous permettent d'explorer la dispersion des données plus finement.

Les résultats complets sont présentés en Annexe C dans le Tableau 18 pour le score promotion et dans le Tableau 19 pour le score prévention.

score promotion Concernant le score promotion, il n'y a pas de différence en fonction du type d'agent. Les scores sont plutôt supérieurs à la moyenne pour tous les agents (entre 3,71 et 4,77 sur une échelle de 7).

Dans la condition avec agent virtuel, la dispersion des scores pour les agents RF-Pro et RF-Pre est stable ; dans cette condition, l'agent Switch présente des coefficients de Finn plus bas que les agents RF-Pro et RF-Pre.

Dans la condition sans agent virtuel, la dispersion des scores pour l'agent RF-Pre est 1,5 fois plus élevée lors de l'induction Prévention que lors de l'induction Promotion ; à l'inverse, la dispersion des scores pour l'agent RF-Pro est 2 fois plus élevée lors de l'induction Promotion que lors de l'induction Prévention.

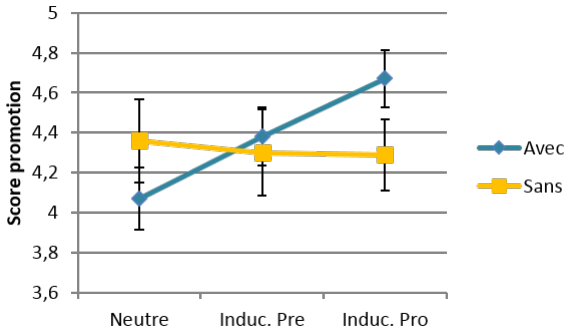
L'analyse de la variance n'a pas montré d'effet principal significatif. Il existe un faible effet d'interaction entre le contexte d'induction et la présence d'un agent virtuel sur le score promotion, $F(2, 206) = 3,39, p < 0,05, \eta^2 = 0,01$. Cependant, les tests post-hoc n'ont pas permis de mettre en évidence de différences significatives. Les résultats sont présentés en Figure 23.
Il semble que les différents agents soient perçus comme plus promotion dans le contexte de l'induction Promotion, mais uniquement pour les participants interagissant avec un agent virtuel.

score prévention Concernant le score prévention, il n'y a pas non plus de différence en fonction du type d'agent. Les scores sont également plutôt supérieurs à la moyenne pour tous les agents, dans les mêmes proportions que pour les scores promotion (entre 3,71 et 4,73 sur une échelle de 7). Les analyses de la variance n'ont montré aucune différence significative.

La dispersion des scores pour l'agent RF-Pre est 2 fois plus élevée lors de l'induction Prévention que lors de l'induction Promotion dans les conditions avec et sans agent virtuel.

Dans la condition avec agent virtuel, le coefficient de Finn pour l'agent RF-Pre est plus haut dans le contexte de l'induction Promotion

5. R package *irr*, version 0.84



(a) Interaction entre la présence d'un agent virtuel et le contexte d'induction (*Barre d'erreur = erreur type*)

		Neutre	Induction Pre	Induction Pro
Avec MARC	moyenne	4,07	4,38	4,67
	écart-type	1,2	1,2	1,2
Sans MARC	moyenne	4,36	4,3	4,29
	écart-type	1,4	1,5	1,2

(b) Statistiques du score promotion en fonction de la présence d'un agent virtuel et du contexte d'induction

FIGURE 23 – Graphique et statistiques de l'effet d'interaction entre la présence d'un agent virtuel et le contexte d'induction pour le score promotion

que dans celui de l'induction Prévention. À l'inverse, le coefficient de Finn pour l'agent RF-Pro est plus haut dans le contexte de l'induction Prévention que dans celui de l'induction Promotion.

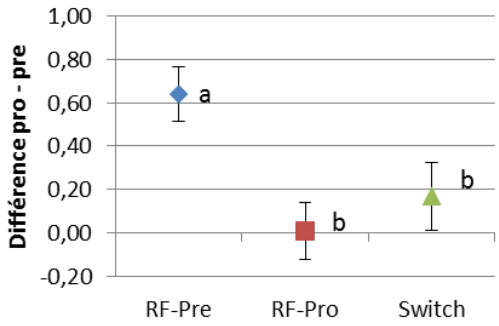
Globalement, les coefficients de Finn les plus bas (0,53 dans la condition avec agent virtuel ; 0,45 dans la condition sans agent virtuel) correspondent à l'agent Switch.

INCLINATION L'inclination perçue, i. e. la différence entre le score promotion et le score prévention, a été calculée. Une inclination positive correspond donc à une perception de l'agent comme plus promotion que prévention (et ce d'autant plus que le score s'éloigne de zéro) et inversement pour une inclination négative.

En moyenne, les scores d'inclination sont centrés autour de zéro pour tous les agents dans les différentes conditions (minimum$= -0,36$; maximum$= 0,90$).

L'analyse de la variance a montré un léger effet du type d'agent sur l'inclination, $F(2,103) = 3,41, p < 0,05, \eta^2 = 0,03$. Les résultats sont présentés en Figure 24.

L'agent RF-Pre a été perçu comme significativement plus orienté promotion que l'agent RF-Pro et l'agent Switch (resp. $p < 0,01$ et $p <$



(a) Effet du type d'agent sur l'inclination
Les valeurs avec la même lettre ne sont pas significativement différentes à p < 0,05 ; Barre d'erreur = erreur type

	RF-Pre	RF-Pro	Switch
moyenne	0,64	0,01	0,17
écart-type	1,3	1,5	1,6

(b) Statistiques de l'inclination en fonction du type d'agent

Figure 24 – Graphique et statistiques de l'effet principal du type d'agent sur l'inclination perçue

$0,05$). La différence entre l'agent RF-Pro et l'agent Switch n'est pas significative.

6.5.3.2 *Perception de la crédibilité*

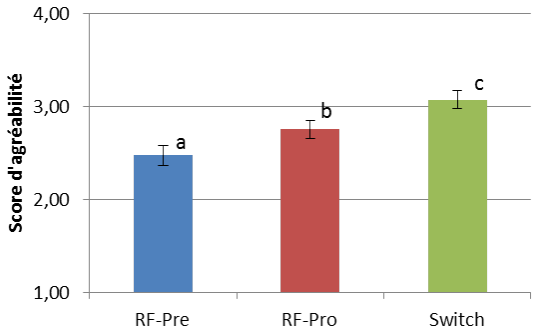
Appréciation Les statistiques descriptives concernant les scores d'appréciation sont présentées en Annexe C - Tableau 20. Globalement, les scores sont centrés autour de 2,5 (sur une échelle de 5 ; minimum= $2,27$; maximum= $3,27$).

L'analyse de la variance a montré un effet modéré du type d'agent sur l'appréciation, $F(2,103) = 5,18, p < 0,01, \eta^2 = 0,05$. Les résultats sont présentés en Figure 25.

L'agent RF-Pre a été perçu comme significativement moins agréable que l'agent RF-Pro et l'agent Switch (resp. $p < 0,05$ et $p < 0,001$) et l'agent RF-Pro a été perçu comme significativement moins agréable que l'agent Switch ($p < 0,05$).

Intelligence perçue Les statistiques descriptives concernant les scores d'intelligence perçue sont présentées en Annexe C - Tableau 21. Globalement, les scores sont centrés autour de 3 (sur une échelle de 5 ; minimum= $2,73$; maximum= $3,56$). Aucune différence significative n'a pu être montrée.

Par ailleurs, l'appréciation de l'agent et l'intelligence perçue sont assez fortement correlées ($r(327) = 0,47, p < 0,001$).



(a) Effet du type d'agent sur l'appréciation
Les valeurs avec la même lettre ne sont pas significativement différentes à p < 0,05 ; Barre d'erreur = erreur type

	RF-Pre	RF-Pro	Switch
moyenne	2,47	2,75	3,07
écart-type	1,1	1,1	0,9

(b) Statistiques de l'appréciation en fonction du type d'agent

FIGURE 25 – Graphique et statistiques de l'effet principal du type d'agent sur l'appréciation perçue

6.5.3.3 *Perception du changement*

39 participants n'ont pu réaliser la session 4, celle du changement d'agent, à cause de modifications du planning expérimental par le collège. Les analyses ont donc été réalisées sur les 87 participants restants.

Les fréquences sont présentées en Annexe C dans le Tableau 17. Il y a significativement moins d'erreurs (39%) que de bonnes réponses (61%) ($\chi^2(2, N = 87) = 4,15, p < 0,05$). Les autres différences ne sont pas significatives.

6.5.4 *Perception de l'interaction*

Les résultats concernant la perception de l'interaction sont présentés en fonction du type d'agent, de la présence de l'agent virtuel et de la session de passation (variable temporelle). Comme pour la perception de l'agent, 17 participants ont été retirés des analyses car ils n'ont pas participé à l'ensemble des trois sessions.

6.5.4.1 *Flow*

FLOW - INTÉRÊT Les statistiques descriptives concernant les scores de flow pour la partie Intérêt sont présentées en Annexe C - Tableau 22. Globalement, les scores sont centrés autour de 3,5 (sur une

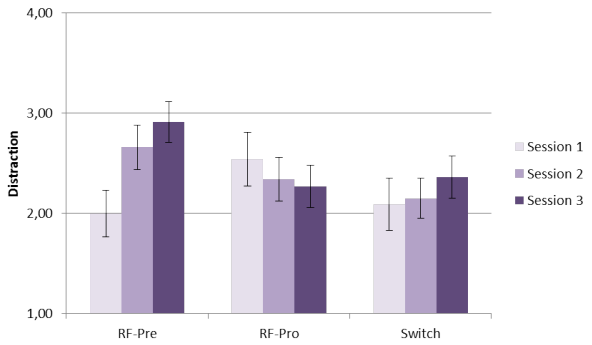
échelle de 5 ; minimum= $3,04$; maximum= $3,88$). Aucune différence significative n'a pu être montrée.

L'intérêt pour la partie et l'appréciation de l'agent ont une corrélation modérée ($r(327) = 0,40, p < 0,001$).

flow - distraction Les statistiques descriptives concernant les scores de flow pour la partie Distraction sont présentées en Annexe C - Tableau 23. En moyenne, les scores sont centrés autour de 2,5 (sur une échelle de 5 ; minimum= $1,90$; maximum= $3,07$) mais sont plutôt dispersés (CQV moyen= 47%).

L'analyse de la variance n'a pas montré d'effet principal significatif. Il existe une faible interaction entre le type d'agent et la session de passation sur la distraction, $F(4,206) = 2,46, p = 0,02, \eta^2 = 0,02$. Les tests post-hoc n'ont pas permis de montrer de différences significatives. Les résultats sont présentés en Figure 26.

Il semble que la distraction soit plus importante lors de l'interaction avec un agent RF-Pre durant les sessions 2 et 3 que durant la session 1, alors que la distraction semble plus stable pour les agents RF-Pro et Switch.



(a) Interaction entre le type d'agent et la session (*Barre d'erreur = erreur type*)

		Session 1	Session 2	Session 3
RF-Pre	moyenne	2,00	2,66	2,91
	écart-type	1,4	1,6	1,5
RF-Pro	moyenne	2,54	2,34	2,27
	écart-type	1,4	1,4	1,3
Switch	moyenne	2,09	2,15	2,36
	écart-type	1,2	1,2	1,2

(b) Statistiques du score de distraction en fonction du type d'agent et de la session

Figure 26 – Graphique et statistiques de l'effet d'interaction entre le type d'agent et la session pour le score de distraction

6.5.4.2 *Coprésence*

Les statistiques descriptives concernant les scores de coprésence sont présentées en Annexe C - Tableau 24. Les scores sont centrés autour de 3 (sur une échelle de 5 ; minimum= $2,28$; maximum= $3,10$). Aucune différence significative n'a pu être montrée.

La coprésence est modérément corrélée à l'intelligence perçue ($r(218) = 0,40, p < 0,001$) et à l'intérêt pour la partie ($r(218) = 0,40, p < 0,001$).

6.5.5 *Perception de la relation*

Le questionnaire de McGill rempli lors de la session 1 permet d'obtenir une idée des attentes initiales des sujets concernant l'agent. En effet, à ce moment, les participants n'avaient eu qu'une seule interaction avec l'agent. De ce fait, leurs réponses reflètent plus la relation qu'ils supposent possible avec un agent artificiel imaginaire que la relation qu'ils pensent pouvoir entretenir avec l'agent artificiel qui leur est concrètement présenté.

Pour analyser les résultats, un partitionnement en k-moyennes des participants a donc été réalisé afin de voir si l'évolution de la perception de la relation diffère en fonction des attentes initiales. Le partitionnement a été fait sur les 5 scores du questionnaire de McGill (amplitude : de 0 à 8 ; plus le score est important, plus la fonction de l'amitié en question parait réalisable par l'agent).

Les participants ont été divisés en 3 clusters. Les scores moyens pour chaque cluster sont présentés dans le Tableau 14. Le premier cluster ($n = 45$) est interprété comme un groupe ayant des attentes modérées sur la relation avec un agent (*Gr-Mod*) ; le deuxième cluster ($n = 18$) est interprété comme un groupe ayant des attentes basses (*Gr-Low*), la relation leur paraissant globalement improbable ; le troisième cluster ($n = 43$) est interprété comme un groupe ayant des attentes élevées (*Gr-High*), la relation leur paraissant potentiellement forte.

Tableau 14 – Moyennes des 5 scores du questionnaire McGill pour les trois clusters constitués

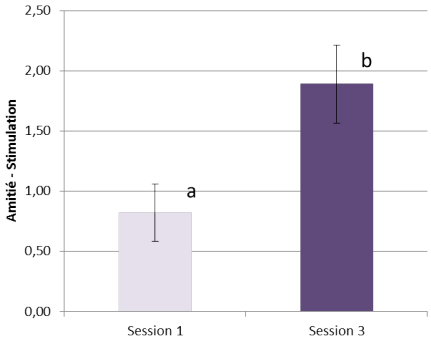
	Gr-Mod ($n = 45$)	Gr-Low ($n = 18$)	Gr-High ($n = 43$)
Stimulation	3,41	0,82	5,89
Aide	3,60	1,58	6,20
Intimité	2,44	0,43	3,98
Fiabilité	3,15	0,56	5,72
Self validation	4,47	0,91	6,03

6.5.5.1 *Stimulation*

Le score Stimulation réfère à l'évaluation de la capacité pour l'agent et l'utilisateur de faire des choses sympathiques et amusantes ensemble.

gr-mod Dans le cluster aux attentes modérées, aucun effet significatif n'a été montré sur le score Stimulation (Session 1 : $M = 3,41$; $SD = 1,31$; Session 3 : $M = 2,92$; $SD = 1,93$).

gr-low Dans le cluster aux attentes basses, l'analyse de la variance a montré un effet de la session sur le score Stimulation, $F(1,12) = 6,28, p < 0,05, \eta^2 = 0,23$. Les agents sont perçus comme capables d'être des compagnons plus stimulants en Session 3 qu'en Session 1 ($p = 0,01$). Les résultats sont présentés en Figure 27.



(a) Effet du temps sur le score Stimulation
Les valeurs avec la même lettre ne sont pas significativement différentes à p < 0,05 ; Barre d'erreur = erreur type

	Session 1	Session 3
moyenne	0,82	1,89
écart-type	1,01	1,37

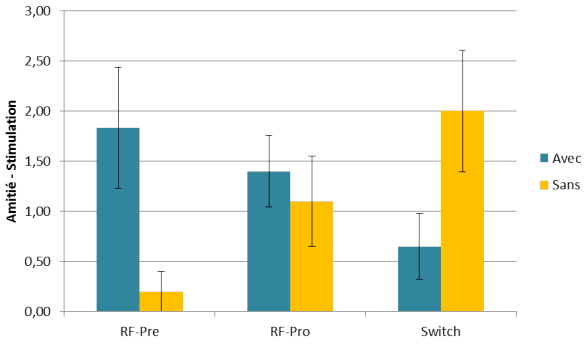
(b) Statistiques du score Stimulation en fonction de la session

Figure 27 – Graphique et statistiques de l'effet de la session sur le score Stimulation pour le Gr-Low

L'analyse de la variance a également montré un effet d'interaction entre le type d'agent et la présence de l'agent virtuel, $F(2,12) = 4,51, p < 0,05, \eta^2 = 0,25$. Les tests post-hoc n'ont pas permis de montrer de différences significatives.

L'agent RF-Pre parait pouvoir être un compagnon plus stimulant lorsqu'il est accompagné d'un agent virtuel MARC. Au contraire, l'agent Switch parait pouvoir être un compagnon plus stimulant lorsqu'il

n'a pas de représentation physique. Les résultats sont présentés en Figure 28.



(a) Interaction entre le type d'agent et la présence d'un agent virtuel (*Barre d'erreur = erreur type*)

		Avec MARC	Sans MARC
RF-Pre	moyenne	1,83	0,20
	écart-type	1,47	0,40
RF-Pro	moyenne	1,40	1,10
	écart-type	1,13	0,90
Switch	moyenne	0,65	2,00
	écart-type	0,66	1,71

(b) Statistiques du score Stimulation en fonction du type d'agent et de la présence d'un agent virtuel

FIGURE 28 – Graphique et statistiques de l'effet d'interaction sur le score Stimulation pour le Gr-Low

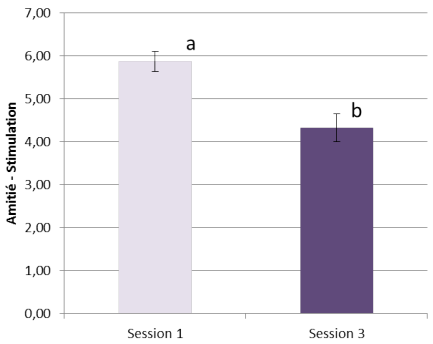
GR-HIGH Dans le cluster aux attentes fortes, l'analyse de la variance a montré un effet de la session sur le score Stimulation, $F(1,37) = 30,6, p < 0,001, \eta^2 = 0,14$.

Les agents sont perçus comme moins capables d'être des compagnons stimulants en Session 3 qu'en Session 1 ($p < 0,001$). Les résultats sont présentés en Figure 29.

6.5.5.2 *Aide*

Le score Aide (*Help*) évalue la capacité perçue de l'agent à apporter des conseils, des informations ou tout autre forme d'aide permettant de satisfaire des besoins ou des buts du participant.

GR-MOD Dans le cluster aux attentes modérées, aucun effet significatif n'a été montré sur le score Aide (Session 1 : $M = 3,60$; $SD = 1,42$; Session 3 : $M = 3,36$; $SD = 1,83$).



(a) Effet du temps sur le score Stimulation
Les valeurs avec la même lettre ne sont pas significativement différentes à p < 0,05 ; Barre d'erreur = erreur type

	Session 1	Session 3
moyenne	5,86	4,32
écart-type	1,77	2,12

(b) Statistiques du score Stimulation en fonction de la session

Figure 29 – Graphique et statistiques de l'effet de la session sur le score Stimulation pour le Gr-High

gr-low Dans le cluster aux attentes basses, aucun effet significatif n'a été montré sur le score Aide (Session 1 : $M = 1,58$; $SD = 1,62$; Session 3 : $M = 2,12$; $SD = 1,70$).

gr-high Dans le cluster aux attentes fortes, l'analyse de la variance a montré un effet de la session sur le score Aide, $F(1,37) = 22,9, p < 0,001, \eta^2 = 0,20$.

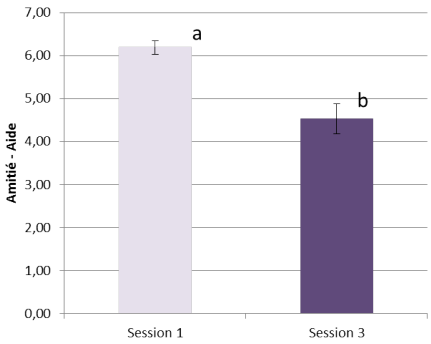
Les agents sont perçus comme moins capables de fournir de l'aide en Session 3 qu'en Session 1 ($p < 0,001$). Les résultats sont présentés en Figure 30.

6.5.5.3 *Intimité*

Le score Intimité (*Intimity*) évalue la capacité perçue de l'agent à offrir un contexte où le participant peut exprimer ouvertement et sincèrement des informations personnelles. Il est à noter que le score moyen Intimité est le plus bas (par rapport aux 4 autres dimensions) pour les 3 clusters.

gr-mod Dans le cluster aux attentes modérées, l'analyse de la variance a montré un effet de la présence de l'agent virtuel sur le score Intimité, $F(1,42) = 4,60, p < 0,05, \eta^2 = 0,06$.

Les participants interagissant sans agent virtuel se sentent plus à même de partager des choses personnelles avec leur agent que les



(a) Effet du temps sur le score Aide
Les valeurs avec la même lettre ne sont pas significativement différentes à p < 0,05 ; Barre d'erreur = erreur type

	Session 1	Session 3
moyenne	6,19	4,53
écart-type	1,01	2,27

(b) Statistiques du score Aide en fonction de la session

Figure 30 – Graphique et statistiques de l'effet de la session sur le score Aide pour le Gr-High

participants interagissant avec un agent virtuel ($p = 0,01$). Les résultats sont présentés en Figure 31.

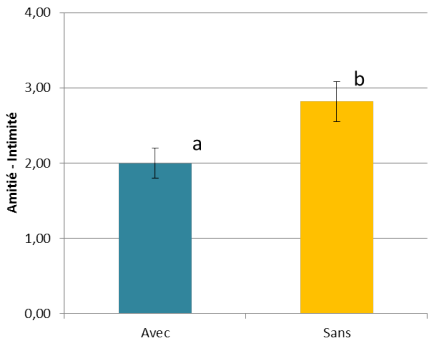
gr-low Dans le cluster aux attentes basses, aucun effet significatif n'a été montré sur le score Intimité (Session 1 : $M = 0,43$; $SD = 0,57$; Session 3 : $M = 1,18$; $SD = 1,38$).

gr-high Dans le cluster aux attentes fortes, l'analyse de la variance a montré un effet de la semaine, $F(1,37) = 7,00, p = 0,01, \eta^2 = 0,04$ ainsi qu'un léger effet d'interaction entre le type d'agent, la présence de l'agent virtuel et la session $F(2,37) = 3,29, p < 0,05, \eta^2 = 0,03$. Les tests post-hoc n'ont pas permis de montrer de différences significatives.

Il semble qu'en présence d'un agent virtuel, les scores Intimité diminuent entre les deux sessions pour les agents RF-Pre et RF-Pro mais pas pour l'agent Switch ; en l'absence d'agent virtuel, le pattern s'inverse : les scores Intimité diminuent entre les deux sessions pour l'agent Switch mais pas pour les agents RF-Pre et RF-Pro. Les résultats sont présentés en Figure 32.

6.5.5.4 *Fiabilité*

Le score Fiabilité (*Reliable Alliance*) évalue dans quelle mesure le participant pense pouvoir compter sur la disponibilité et la loyauté de l'agent, et ce durablement dans le temps.



(a) Effet de la présence d'un agent virtuel sur le score Intimité
Les valeurs avec la même lettre ne sont pas significativement différentes à p < 0,05 ; Barre d'erreur = erreur type

	Avec MARC	Sans MARC
moyenne	2,00	2,82
écart-type	1,53	1,59

(b) Statistiques du score Aide en fonction de la session

Figure 31 – Graphique et statistiques de l'effet de la session pour le score Aide pour le Gr-High

gr-mod Dans le cluster aux attentes modérées, l'analyse de la variance a montré un effet d'interaction entre le type d'agent, la présence de l'agent virtuel et la session $F(2,42) = 3,35, p < 0,05, \eta^2 = 0,07$. Les tests post-hoc n'ont pas permis de montrer de différences significatives.

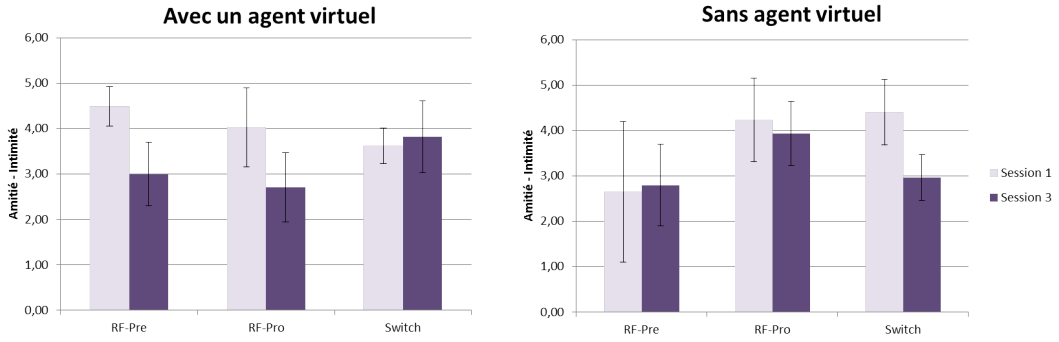
Il semble qu'en présence d'un agent virtuel, le score Fiabilité diminue pour l'agent RF-Pre et augmente pour l'agent RF-Pro ; en l'absence d'agent virtuel, le score Fiabilité augmente pour l'agent RF-Pre et l'agent Switch mais diminue pour l'agent RF-Pro. Les résultats sont présentés en Figure 33

gr-low Dans le cluster aux attentes basses, l'analyse de la variance a montré un effet de la session sur le score Fiabilité, $F(1,12) = 10,77, p < 0,01, \eta^2 = 0,36$.

Les agents sont perçus comme capables d'être des compagnons plus loyaux et disponibles en Session 3 qu'en Session 1 ($p = 0,01$). Les résultats sont présentés en Figure 34.

L'analyse de la variance a également montré un effet d'interaction entre le type d'agent et la présence de l'agent virtuel, $F(2,12) = 11,32, p < 0,01, \eta^2 = 0,40$. Les tests post-hoc n'ont pas permis de montrer de différences significatives.

Les agents RF-Pre et RF-Pro paraissent pouvoir être des compagnons plus fiables lorsqu'ils sont accompagné d'un agent virtuel MARC.

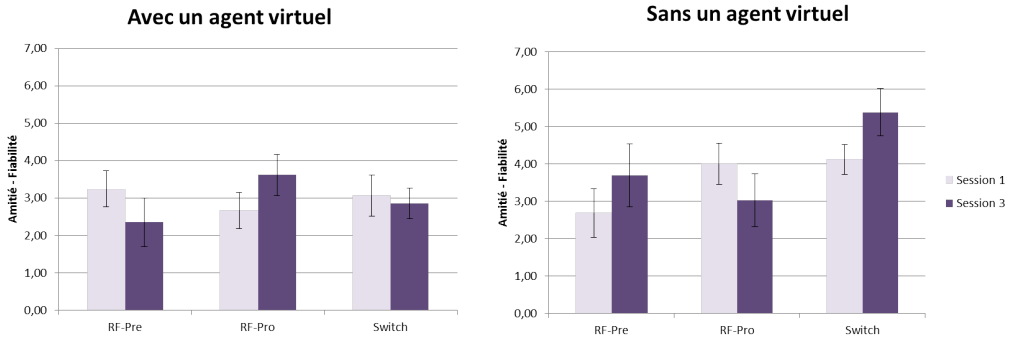


(a) Interaction entre la session, le type d'agent et la présence d'un agent virtuel sur le score Intimité (*Barre d'erreur = erreur type*)

		Session 1	Session 3
Avec MARC			
RF-Pre	moyenne	4,49	3,00
	écart-type	1,14	2,37
RF-Pro	moyenne	4,03	2,71
	écart-type	2,30	2,02
Switch	moyenne	3,62	2,82
	écart-type	1,16	2,36
		Session 2	Session 3
Sans MARC			
RF-Pre	moyenne	2,65	2,80
	écart-type	3,09	0,98
RF-Pro	moyenne	4,23	3,94
	écart-type	2,44	1,86
Switch	moyenne	4,40	2,96
	écart-type	2,16	1,53

(b) Statistiques du score Intimité en fonction de la session, du type d'agent et de la présence d'un agent virtuel

Figure 32 – Graphique et statistiques de l'effet d'interaction sur le score Intimité pour le Gr-High

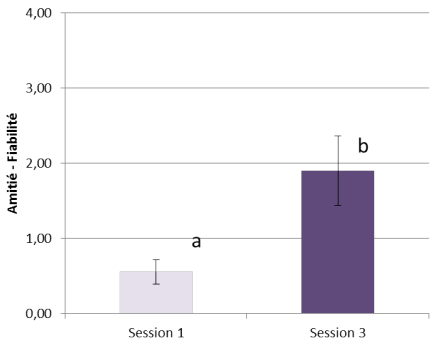


(a) Interaction entre la session, le type d'agent et la présence d'un agent virtuel sur le score Fiabilité (*Barre d'erreur = erreur type*)

		Session 1	Session 3
Avec MARC			
RF-Pre	moyenne	3,25	2,36
	écart-type	1,58	2,15
RF-Pro	moyenne	2,67	3,62
	écart-type	1,69	1,91
Switch	moyenne	3,07	2,86
	écart-type	1,45	1,06
		Session 2	Session 3
Sans MARC			
RF-Pre	moyenne	2,69	3,69
	écart-type	1,20	2,37
RF-Pro	moyenne	4,00	3,03
	écart-type	1,00	2,00
Switch	moyenne	4,12	5,38
	écart-type	0,88	0,88

(b) Statistiques du score Fiabilité en fonction de la session, du type d'agent et de la présence d'un agent virtuel

Figure 33 – Graphique et statistiques de l'effet d'interaction sur le score Fiabilité pour le Gr-Mod



(a) Effet du temps sur le score Fiabilité
Les valeurs avec la même lettre ne sont pas significativement différentes à p < 0,05 ; Barre d'erreur = erreur type

	Session 1	Session 3
moyenne	0,56	1,90
écart-type	0,16	0,46

(b) Statistiques du score Fiabilité en fonction de la session

Figure 34 – Graphique et statistiques de l'effet de la session sur le score Fiabilité pour le Gr-Low

Au contraire, l'agent Switch parait pouvoir être un compagnon plus fiable lorsqu'il n'a pas de représentation physique. Les résultats sont présentés en Figure 35.

gr-high Dans le cluster aux attentes fortes, l'analyse de la variance a montré un effet de la session sur le score Fiabilité, $F(1,37) = 10,1, p < 0,01, \eta^2 = 0,13$.

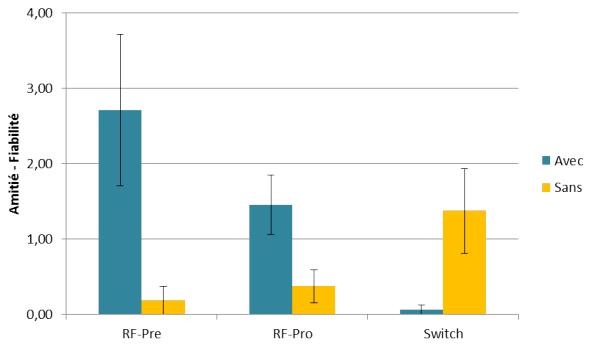
Les agents sont perçus moins fiables en Session 3 qu'en Session 1 ($p < 0,01$). Les résultats sont présentés en Figure 36.

L'analyse de la variance a également montré un effet d'interaction entre le type d'agent et la présence de l'agent virtuel, $F(2,37) = 5,41, p < 0,01, \eta^2 = 0,12$. Les tests post-hoc n'ont pas permis de montrer de différences significatives.

L'agent RF-Pre parait pouvoir être un compagnon plus fiable lorsqu'il est accompagné d'un agent virtuel MARC. Par contre, l'agent RF-Pro parait pouvoir être un compagnon plus fiable lorsqu'il n'a pas de représentation physique. Les résultats sont présentés en Figure 37.

6.5.5.5 *Self validation*

Le score de Self Validation (*Self-Validation*) évalue dans quelle mesure l'agent est perçu quelqu'un qui pourrait être à l'écoute, encourageant le participant à maintenir une bonne image de lui-même.

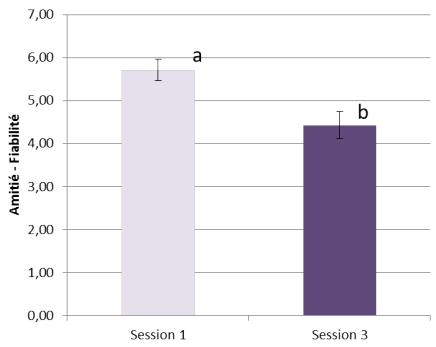


(a) Interaction entre le type d'agent et la présence d'un agent virtuel (*Barre d'erreur = erreur type*)

		Avec MARC	Sans MARC
RF-Pre	moyenne	2,71	0,19
	écart-type	2,46	0,38
RF-Pro	moyenne	1,45	0,38
	écart-type	1,25	0,43
Switch	moyenne	0,06	1,38
	écart-type	0,13	1,59

(b) Statistiques du score Fiabilité en fonction du type d'agent et de la présence d'un agent virtuel

Figure 35 – Graphique et statistiques de l'effet d'interaction sur le score Fiabilité pour le Gr-Low

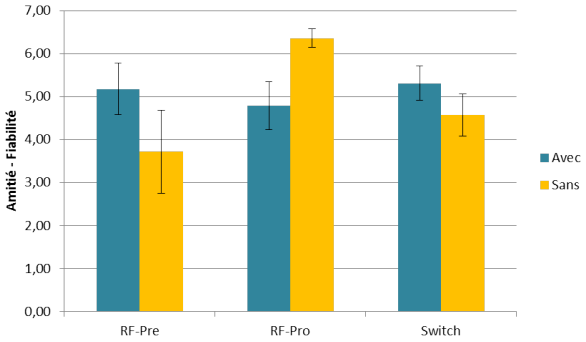


(a) Effet du temps sur le score Fiabilité
Les valeurs avec la même lettre ne sont pas significativement différentes à $p < 0,05$; Barre d'erreur = erreur type

	Session 1	Session 3
moyenne	5,71	4,43
écart-type	0,25	0,32

(b) Statistiques du score Fiabilité en fonction de la session

Figure 36 – Graphique et statistiques de l'effet de la session sur le score Fiabilité pour le Gr-High



(a) Interaction entre le type d'agent et la présence d'un agent virtuel (*Barre d'erreur = erreur type*)

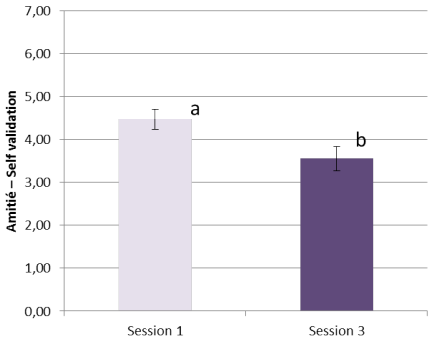
		Avec MARC	Sans MARC
RF-Pre	moyenne	5,18	3,72
	écart-type	2,26	2,72
RF-Pro	moyenne	4,79	6,36
	écart-type	2,07	0,80
Switch	moyenne	5,31	4,57
	écart-type	1,70	2,06

(b) Statistiques du score Fiabilité en fonction du type d'agent et de la présence d'un agent virtuel

Figure 37 – Graphique et statistiques de l'effet d'interaction sur le score Fiabilité pour le Gr-High

gr-mod Dans le cluster aux attentes modérées, l'analyse de la variance a montré un effet de la session sur le score Self validation, $F(1,42) = 14,2, p < 0,001, \eta^2 = 0,07$.

Les agents sont perçus moins capables d'aider le participant à maintenir une bonne image de lui-même en Session 3 qu'en Session 1 ($p = 0,01$). Les résultats sont présentés en Figure 38.



(a) Effet du temps sur le score Self validation
Les valeurs avec la même lettre ne sont pas significativement différentes à p < 0,05 ; Barre d'erreur = erreur type

	Session 1	Session 3
moyenne	4,47	3,55
écart-type	1,62	1,95

(b) Statistiques du score Self validation en fonction de la session

Figure 38 – Graphique et statistiques de l'effet de la session sur le score Self validation pour le Gr-Mod

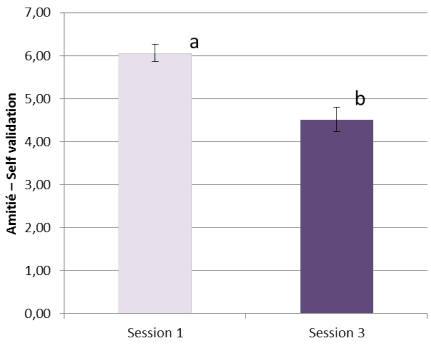
gr-low Dans le cluster aux attentes basses, aucun effet significatif n'a été montré sur le score Self validation (Session 1 : $M = 0,91$; $SD = 1,19$; Session 3 : $M = 1,78$; $SD = 1,59$).

gr-high Dans le cluster aux attentes fortes, l'analyse de la variance a montré un effet de la session sur le score Self validation, $F(1,37) = 22,6, p < 0,001, \eta^2 = 0,21$.

Les agents sont perçus moins capables d'aider le participant à maintenir une bonne image de lui-même en Session 3 qu'en Session 1 ($p < 0,01$). Les résultats sont présentés en Figure 39.

6.6 discussion

Sur nos 8 hypothèses de départ, 5 sont infirmées, 1 n'est ni confirmée ni infirmée et 2 sont partiellement validées. Ces résultats contras-



(a) Effet du temps sur le score Self validation
Les valeurs avec la même lettre ne sont pas significativement différentes à p < 0,05 ; Barre d'erreur = erreur type

	Session 1	Session 3
moyenne	6,06	4,51
écart-type	1,33	1,85

(b) Statistiques du score Self validation en fonction de la session

Figure 39 – Graphique et statistiques de l'effet de la session sur le score Self validation pour le Gr-High

tés demandent à être interprétés et questionnent aussi bien la conception que l'évaluation de systèmes informatiques dotés de personnalité. Cette étude apporte également un éclairage intéressant sur les attentes des adolescents concernant la relation avec un compagnon artificiel et les évolutions de ces attentes après les sessions d'interaction.

6.6.1 *Perception de l'agent*

6.6.1.1 *Perception du regulatory focus*

Notre hypothèse H1 (*Les deux regulatory-focus implémentés sont perçus comme tels par les utilisateurs*) n'est pas validée et notre hypothèse H2 (*La présence visuelle d'un agent virtuel améliore la perception de la personnalité grâce à la cohérence des informations fournies*) est en partie validée.

En effet, concernant les scores promotion et prévention notés en hétéro-attribution, les agents RF-Pro (resp. RF-Pre) n'ont pas une score plus élevé en promotion (resp. prévention) que les autres agents. Au contraire, en observant l'inclination, il s'avère que les agents RF-Pre ont été perçus comme plus inclinés promotion que les autres agents.

En explorant les données de jeu, il s'avère que les tours les plus longs ont été réalisés par les agents RF-Pre (en moyenne, nombre de lancers de dés dans le tour le plus long = 7,7) par rapport aux agents RF-Pro (5,2) et aux agents Switch (5,1). Les agents RF-Pre ont donc pu être perçus comme prenant plus de risques dans leur stratégie.

Nous pouvons expliquer cette différence en analysant les arbres de décision propres à chaque focus. En effet, comme nous l'avons dit plus haut, il apparait que l'arbre Prévention fasse ressortir plus de caractéristiques prenant en compte la position de l'adversaire. Lorsque l'agent RF-Pre a beaucoup d'avance sur son adversaire, sa position dominante semble lui permettre de prendre des risques : même si l'agent perd son tour, il retourne à sa situation initiale qui lui était déjà profitable, ce qui représente donc une non-perte. Or, nous avons pu observer que beaucoup de nos participants avaient une stratégie (tours plutôt courts et éparpillement des camps sur diverses colonnes) assez différente des adultes sur lesquels les stratégies des agents sont basées. Il est donc possible qu'un certain nombre d'adolescents ne se soient pas montrés des adversaires très menaçants pour l'agent RF-Pre, qui a adapté son jeu en fonction.

Globalement, le consensus autour de la personnalité des agents Switch est plus bas que pour les agents RF-Pre et RF-Pro. Cet effet semble être modéré par la présence d'un agent virtuel pour le score promotion. Il est possible que la stratégie de jeu tantôt promotion, tantôt prévention de l'agent Switch ait été perçue comme une forme d'adaptation au jeu par les participants et donc considéré comme une forme de personnalité à part entière.

La perception du changement n'est d'ailleurs pas significativement moins bonne pour les agents Switch que pour les agents RF-Pre et RF-Pro. Cependant, il semble que l'incohérence des expressions faciales, tantôt promotion, tantôt prévention apporte plus de confusion. Cet effet n'est cependant pas clairement démontré et mériterait d'être exploré plus avant.

La préférence des participants pour l'agent Switch questionne la méthodologie à adopter lors de l'évaluation de ce type de dispositifs. Les études existantes proposent généralement deux personnalités clairement différenciées (typiquement introverti/extraverti (Lee et al., 2006; Tapus et al., 2008)). Entre ces deux personnalités extrêmes, les utilisateurs peuvent avoir une préférence. Mais en présence d'une troisième option plus nuancée, cette préférence resterait-elle la même ?

6.6.1.2 *Perception de la crédibilité*

Nos hypothèse H3 (*L'implémentation d'un regulatory-focus cohérent apporte plus de crédibilité aux agents*) et H4 (*Il existe un effet de regulatory fit lorsque la personnalité de l'agent et le contexte induit sont les mêmes*) ne sont pas vérifiées.

Concernant la crédibilité de l'agent, l'ensemble des scores des trois agents sont supérieurs à la moyenne. Les trois agents implémentés semblent donc relativement égaux en termes de crédibilité. Et contrairement à notre hypothèse, c'est l'agent Switch qui est perçu comme le plus appréciable. Il est possible que les variations comportementales de l'agent Switch, qui le rendent moins prédictible, soient un atout dans le cadre d'une interaction ludique. Il serait intéressant de voir s'il en est de même dans une activité plus sérieuse.

Bien qu'aucun effet de regulatory fit n'ait pu être montré, l'induction d'un contexte particulier en termes de regulatory-focus semble avoir un effet sur la perception de l'agent. En effet, nous avons montré que les participants interagissant avec un agent virtuel trouvaient leurs agents plus promotion en contexte Promotion, ce qui n'est pas le cas pour les participants interagissant uniquement avec un agent artificiel. La présence d'une incarnation physique joue donc sur la perception de la personnalité, peut-être par une interprétation des expressions faciales en fonction du contexte.

Il est également à noter que les scores de personnalité des agents avec personnalité semblent plus dispersés en contexte fit (induction Prévention pour l'agent RF-Pre et induction Promotion pour l'agent RF-Pro) qu'en contexte non-fit. Le fait que l'agent soit en contexte fit joue donc probablement un effet différent en fonction des participants, peut-être en fonction de leur focus chronique. Il est possible qu'un effet de regulatory fit existe en prenant en compte le tryptique focus du participant, focus de l'agent et focus du contexte. Van-Dijk and Kluger (2004) encouragent d'ailleurs à explorer une triple interaction du même type dans le cadre de leurs travaux sur l'impact du feedback sur la motivation (focus chronique $\times$ focus situationnel $\times$ focus du feedback). Néanmoins, en l'absence de données cohérentes sur le focus chronique des participants, nous ne pouvons aller plus loin dans l'interprétation de ce résultat.

6.6.2 *Concernant la perception de l'interaction*

Nos hypothèses H5 (*L'interaction est mieux perçue lorsque l'agent a une personnalité cohérente*) et H6 (*Un agent virtuel amoindrit le sentiment d'interagir uniquement "avec un ordinateur"*) ne sont pas vérifiées.

Globalement, les participants se sont montrés intéressés par la tâche mais avec une tendance à se montrer distraits au cours de celle-ci. La distraction augmente d'ailleurs avec le temps pour les participants interagissant avec des agents RF-Pre. Cet effet peut être expliqué par le fait que les agents RF-Pre ont joué des tours plus longs que les autres agents. Nous avons pu observer qu'un certain nombre de participants avaient du mal à rester concentrés sur le jeu lors du tour de

l'agent et cet effet a dû être d'autant plus important que l'agent jouait longtemps.

Le sentiment de coprésence ne semble pas être augmenté par la présence d'un agent virtuel ou par l'ajout du regulatory focus. Les scores étaient plutôt moyens dans toutes les conditions. Si la présence de l'agent virtuel semble avoir d'autres effets, ceux-ci ne sont pas mesurés par ce questionnaire.

Nous notons que le score de coprésence est corrélé à l'intelligence perçue et à l'intérêt pour la partie. Ces deux caractéristiques (percevoir l'ordinateur comme "intelligent" et être engagé dans l'interaction) sont peut-être deux variables plus importantes que la présence d'une incarnation physique pour le sentiment de coprésence chez notre population (possiblement plus habituée aux interactions homme-machine et/ou dématérialisées du fait de leur tranche d'âge).

6.6.3 *Concernant la relation avec l'agent*

Notre hypothèse H7 (*Les sessions répétées permettent aux participants de se construire une image de la personnalité de l'agent allant plus loin que la formation d'impression*) n'est ni infirmée ni confirmée car la personnalité des agents est déjà perçue comme très nuancée dès la première session.

Aucune inclination n'est clairement marquée (pas de différence > 1). En l'état, il est difficile de savoir si cet effet est dû aux comportements des agents proprement dits ou au questionnaire utilisé. De plus, il n'est pas dit que les participants aient pu dépasser la formation d'impression au sens où l'interaction avec l'agent était toujours la même (jouer une partie de Can't Stop). Il serait intéressant de voir si le fait d'interagir avec l'agent dans différents contextes (par exemple, une tâche ludique et une tâche pédagogique) permettrait une perception plus complexe de la personnalité de l'agent.

Notre hypothèse H8 (*Les sessions répétées permettent aux participants de commencer à construire une relation avec leur agent*) est en partie vérifiée. En effet, il y a clairement une évolution de la perception de la relation possible entre le participant et l'agent du fait des sessions répétées.

Cependant, cette évolution est différente en fonction des attentes initiales des participants. La plupart des scores du questionnaire d'amitié tendent à baisser pour le groupe ayant de fortes attentes initiales et à monter pour le groupe ayant de faibles attentes. Il semble donc que les participants qui idéalisaient la relation possible avec un agent artificiel revoient leurs attentes à la baisse, tandis que ceux qui pensaient la relation quasi-impossible soient plus prêts à l'imaginer après avoir interagi plusieurs fois avec leur agent.

Quant à savoir pourquoi certains pensaient la relation impossible, la question reste ouverte : est-ce parce que l'agent leur paraissait être un inconnu ou parce qu'une relation avec un dispositif technologique ne leur paraissait pas concevable ? Il apparait en tout cas que les participants tendent vers une visualisation d'une relation "moyenne" avec leur agent. Une vision peut-être plus en accord avec la technologie que nous leur proposons, soit une machine qui n'est pas à la hauteur des capacités sociales humaines mais qui n'est pas non plus pas totalement dénuée d'affectivité ?

Par ailleurs, la nature de la relation imaginée peut être mise en avant ici. Par exemple, les scores Intimité sont les plus bas pour tous les groupes. Confier des choses personnelles à un agent artificiel ne semble donc pas faire partie des fonctions attendues pour un compagnon artificiel, du moins lorsque celui-ci est présenté dans une interaction ludique telle que la nôtre.

De plus, si nous observons le groupe aux attentes modérées, le score Intimité est plus bas pour les participants interagissant avec un agent virtuel que pour ceux n'en ayant pas. L'absence d'incarnation permet-elle de se projeter dans une relation plus personnelle, sans jugement de la part de l'agent ou bien la présence d'une expression émotionnelle (d'autant plus que celle-ci n'est pas empathique) amoindrit-elle la capacité de l'agent à apparaitre comme un compagnon potentiellement à l'écoute ?

Il apparait également que la présence ou l'absence d'un agent virtuel joue sur d'autres paramètres de cette relation. Par exemple, pour le groupe ayant des faibles attentes, les agents dotés d'une personnalité semblent être des compagnons potentiellement plus stimulants et plus fiables lorsqu'ils sont accompagnés d'une incarnation.

Cet effet pourrait être dû au fait que le participant peut alors expérimenter le fait que la machine exprime quelque chose d'émotionnel. Cette interprétation peut être confortée par l'inversion de cet effet pour les agents Switch (plus stimulants et plus fiables en l'absence d'un agent virtuel). L'incohérence du pattern émotionnel exprimé par ces agents pourrait expliquer ce résultat.

6.7 résumé du chapitre

Dans cette deuxième étude, une implémentation socio-cognitive du regulatory focus a été proposée, influençant non seulement les stratégies de jeu mais également le modèle émotionnel d'un agent virtuel capable d'afficher des expressions faciales (MARC, Courgeon et al. (2014)). Nous avons également mis en place une expérimentation à sessions répétées pendant lesquelles les participants (adolescents de 13 ans) jouaient à Can't Stop plusieurs fois avec le même agent.

Pour cette implémentation socio-cognitive, nous avons réutilisé la même méthodologie data-driven que dans le chapitre 5, basée sur des classifieurs ADTree. Plutôt que d'utiliser les scores promotion et prévention comme caractéristiques pour la classification, nous avons séparé notre population en fonction de leur focus chronique et nous avons entrainé deux classifieurs différents : un pour la population promotion et un pour la population prévention. Le modèle émotionnel fait l'objet d'une implémentation symbolique qui est inspirée du modèle CPM (Scherer, 2001). Les modèles produits par les classifieurs, aussi bien que la théorie du regulatory focus, nous ont guidés quant aux choix des variables à prendre en compte pour l'évaluation cognitive.

De cette façon, nous avons créé 3 agents artificiels : un agent RF-Pro ; un agent RF-Pre ; un agent Switch, qui utilise aléatoirement les modèles des agents RF-Pro ou RF-Pre à chaque décision de jeu / expression faciale à réaliser.

126 participants ont interagi (jeu Can't Stop) avec un même agent pendant 4 séances d'1 heure. La moitié des participants interagissaient avec un agent artificiel (stratégies de jeu uniquement) et l'autre moitié avec un agent virtuel (stratégies de jeu et personnage expressif). De façon à évaluer la perception du changement, lors de la quatrième séance, la moitié des participants jouaient avec un agent ayant une personnalité différente tandis que l'autre moitié jouaient avec leur partenaire habituel.

Les résultats ont montré que si les personnalités perçues sont plus nuancées que dans la première étude, les agents RF-Pre et RF-Pro n'ont pas été perçus comme attendu. De même pour la crédibilité, les différences trouvées (Switch > RF-Pro > RF-Pre) ne vont pas dans le sens de l'hypothèse du regulatory fit. Cependant, lors de la session de perception du changement, 60% des participants ont été capables de déterminer correctement s'ils avaient changé de partenaire de jeu ou non. L'exploration des résultats concernant le questionnaire portant sur la relation nous a permis de déterminer 3 groupes ayant des attentes différentes concernant l'agent (attentes faibles, modérées et fortes). Les sessions répétées semblent montrer que tous les participants tendent à évoluer vers des attentes modérées.

Quatrième partie

CONCLUSION

7 CONCLUSION ET PERSPECTIVES

Sommaire

7.1 Résumé du travail de thèse et perspectives à court-terme **146**

7.1.1 Première contribution : Perseed 146

7.1.2 Deuxième contribution : le questionnaire RFQ-PF 147

7.1.3 Troisième contribution : implémentation dimensionnelle du regulatory focus et étude en interaction unique 148

7.1.4 Quatrième contribution : implémentation socio-cognitive du regulatory focus et étude en interaction à sessions répétées 149

7.2 Réflexions générales et perspectives à long-terme **150**

7.2.1 Personnalité et informatique 150

7.2.2 Problématiques méthodologiques 151

7.2.3 Articuler psychologie et informatique 153

7.1 résumé du travail de thèse et perspectives à court-terme

Ce travail de thèse s'intéresse à la problématique de la création d'une personnalité artificielle pour un agent virtuel. L'étude de l'existant a montré que les travaux en informatique affective se rattachant à la génération de personnalité (soit la synthèse de comportements amenant un utilisateur à attribuer une certaine personnalité à un dispositif technologique) sont dominés par l'utilisation du modèle psychologique du Big Five (Vinciarelli and Mohammadi, 2014b).

Le modèle du Big Five définit la personnalité suivant 5 traits principaux. Cependant, ce modèle reste d'obédience descriptive et ne donne pas d'informations quand aux structures et processus mis en jeu dans l'apparition des comportements. C'est pourquoi une approche plus socio-cognitive, qui s'intéresse justement à l'explication de la personnalité et de ses processus, a été proposée. De plus, dans une vision pluridisciplinaire, l'implémentation d'un concept psychologique suivant cette approche devrait pouvoir informer en retour ledit concept.

Un processus psychologique spécifique a été choisi pour aborder la question de la personnalité : la régulation du Soi. Les structures et les processus concernant le Soi sont souvent mis en avant dans les théories socio-cognitives de la personnalité. De plus, ces processus sont centraux pour l'individualité. Notre modélisation s'appuie principalement sur la théorie du regulatory focus de Higgins (1997) qui distingue deux types de stratégies de régulation (promotion-focus et prévention-focus) et qui propose le concept de regulatory fit (effet dû à l'adéquation entre le focus d'un utilisateur et celui de l'interaction à laquelle il participe).

Ce travail de thèse essaie donc de concilier deux buts : étudier l'apport de la modélisation du regulatory focus pour l'interaction homme-machine et informer le concept de regulatory focus par notre implémentation.

7.1.1 *Première contribution : Perseed*

En nous appuyant sur la théorie du regulatory focus (Higgins, 1997) et le framework de la self-régulation proposé par Morf (2006), nous avons proposé le modèle Perseed pour la modélisation computationnel du regulatory focus comme composante de la personnalité. Perseed représente à la fois un cadre conceptuel et un modèle computationnel. Le cadre conceptuel Perseed propose donc un ensemble de concepts pour guider les développeurs dans la caractérisation des éléments utiles lors de la modélisation d'un processus de régulation

du Soi. Le modèle computationnel Perseed constitue une spécification fonctionnelle pour la mise en place du regulatory focus comme processus de régulation du Soi, sous la forme d'un module externe destiné à être connecté à une architecture cognitive.

Si le cadre conceptuel a en partie guidé la suite de notre travail, le modèle computationnel n'a pas été implémenté en tant que tel. A l'intérieur de notre cadre applicatif (jeu Can't Stop), il serait intéressant d'utiliser les données de notre implémentation socio-cognitive (interprétation des arbres de décision) pour implémenter une version symbolique du module de prise de décision. Les comportements produits par des agents symboliques par rapport aux agents machine-learning pourraient être comparés et des expérimentations pourraient étudier si cela entraine des différences dans la perception des agents par des utilisateurs. Il pourrait également être intéressant de reprendre une application fonctionnelle, déjà implémentée avec une architecture cognitive particulière et d'y implémenter Perseed sous la forme d'un plug-in afin d'étudier l'impact de l'ajout du regulatory focus par rapport à l'application existante. Les différents points de vue sur le Soi proposés dans le modèle n'ont pas été exploités pour le moment dans nos implémentations, explorer cet aspect pourrait également être une perspective intéressante à court-terme.

7.1.2 *Deuxième contribution : le questionnaire RFQ-PF*

Face à l'absence de questionnaire français validé et pour pouvoir mesurer le regulatory focus en hétéro-attribution chez un agent virtuel, un questionnaire français a été créé et validé psychométriquement. Inspiré par le travail de Van Stekelenburg (2006), le RFQ-PF (Regulatory Focus Questionnaire - Proverbs Form) propose de mesurer le regulatory focus en utilisant des proverbes. Ces proverbes ont été sélectionnés via différentes étapes impliquant des utilisateurs tout-venant aussi bien que des experts. Plusieurs études ont ensuite permis d'explorer la structure factorielle ainsi que les validités convergentes et discriminantes du RFQ-PF.

Bien qu'encourageantes, ces premières études demandent à être complétées pour mieux cerner le construit mesuré par le RFQ-PF. De plus, la fiabilité par test-retest pourrait être testée. Les effets du genre et de la culture pourraient également être explorés. Par ailleurs, nos études de validation examinent les caractéristiques du RFQ-PF en auto-attribution. Or nous utilisons également ce questionnaire en hétéro-attribution, qui plus est pour des agents artificiels. Notre expérimentation impliquant des adolescents (cf. chapitre 6) a montré une différence entre l'administration du questionnaire en auto- et en hétéro-attribution. En effet, lorsque les adolescents ont dû décrire leur propre personnalité via le RFQ-PF, les résultats n'ont pu être exploi-

tés car la structure factorielle était loin de la structure en 3 facteurs découverte au cours des études de validation. Par contre, lorsque les adolescents ont décrit la personnalité de l'agent via le RFQ-PF, une structure factorielle claire en 2 facteurs (un facteur "Prévention", un facteur "Promotion") est apparue. Il est difficile ici de faire la parte des choses entre ce qui est dû à l'âge des participants et ce qui pourrait s'expliquer par le fait d'être ou non soi-même la cible du questionnaire. D'autres études donc seraient utiles pour mieux comprendre comment les utilisateurs appréhendent ce questionnaire lorsqu'il est utilisé en hétéro-attribution.

7.1.3 *Troisième contribution : implémentation dimensionnelle du regulatory focus et étude en interaction unique*

Après avoir sélectionné un jeu de plateau (de type stop-ou-encore) permettant de mettre en évidence des différences dans les stratégies de jeu selon le focus considéré, une première implémentation dimensionnelle a été proposée. En utilisant une méthodologie data-driven, des données provenant de jeu entre humains ont été utilisées pour entrainer un classifieur de type ADTree (Alternating Decision Tree). Les scores promotion et prévention ont été utilisés par le modèle de prise de décision (concernant le choix stop-ou-encore). Ces arbres de décision ont alors été implémentés pour diriger le comportement d'agents artificiels.

Une expérimentation a alors été mise en place pour comparer la perception des utilisateurs de 4 agents artificiels (un agent promotion, un agent prévention et deux agents "contrôle" : un agent moyen et un agent aléatoire). Chaque participant a joué une partie contre chaque agent et a reporté la personnalité et la crédibilité perçues chez l'agent après chaque partie. Si les agents promotion et prévention ont été perçus comme tels, les agents "contrôle" ont également été perçus comme présentant un regulatory focus marqué. Un effet de regulatory fit est également apparu pour les participants prévention-focus.

Cette première étude suggère que le regulatory focus est un concept pertinent pour l'informatique affective. Le regulatory focus peut être perçu par un utilisateur et amener un effet positif de regulatory fit. D'un point de vue applicatif, un tel effet pourrait être particulièrement utile dans le cadre de l'informatique persuasive. Présenter un message orienté dans le regulatory focus qui convient à l'utilisateur que l'on essaie de convaincre peut accroitre le pouvoir persuasif de la machine. De plus, notre méthodologie data-driven et l'utilisation de techniques d'apprentissage automatique se sont montrées pertinentes pour permettre de diriger les comportements de nos agents et colorer ceux-ci pour que les participants attribuent un certain focus aux agents.

7.1.4 *Quatrième contribution : implémentation socio-cognitive du regulatory focus et étude en interaction à sessions répétées*

La première étude dimensionnelle ne permettait pas d'informer les processus sous-jacents au regulatory focus, c'est pourquoi la deuxième étude utilise une approche orientée socio-cognitive. A l'aide des données des participants recueillies lors de la première étude, nous avons entrainé de nouveaux arbres de décision différenciés selon les focus. De plus, la première étude ne proposait qu'une courte interaction avec l'agent. Dans cette deuxième étude, les participants, des collégiens, ont pu interagir plusieurs fois avec le même agent afin de pouvoir interroger le développement de la relation entre le participant et son agent. Un agent virtuel, capable de réaliser des expressions faciales dirigées par un modèle émotionnel, a également été ajouté au dispositif afin d'étudier l'effet combiné d'une expression multiple de la personnalité (stratégies de jeu et expressions faciales).

Les résultats sont plus contrastés que dans la première étude. Comme attendu, les participants perçoivent la personnalité des agents comme plus nuancée mais les agents promotion et prévention n'ont pas été clairement différenciés. Cependant, l'analyse des logs et l'interprétation des arbres de décision ont pu donner des pistes d'explication pour ces résultats : l'agent prévention, par sa plus grande prise en compte du jeu de l'adversaire, a particulièrement adapté son jeu à celui des adolescents. Il semble donc que notre implémentation des stratégies de jeu par apprentissage automatique permette de prendre en compte la versatilité du concept de regulatory focus malgré l'absence de représentation explicite des motivations.

L'étude du développement de la relation a mis en avant l'importance de prendre en compte les attentes préalables des utilisateurs. En effet, des éléments comme les œuvres de fiction (littérature, cinéma, séries télévisées...) auxquels les utilisateurs ont pu être exposés ou l'image qu'eux-mêmes se font de la technologie sont susceptibles d'influencer la façon dont ils conçoivent la possibilité (ou l'impossibilité) d'établir une relation avec un système informatique.

Pour continuer à explorer l'apport du regulatory focus en tant que composante de la personnalité d'un agent virtuel, il serait intéressant de mettre en place des études cross-situationnelles, où les utilisateurs interagissent avec le même agent mais dans différents contextes (une interaction ludique et une pédagogique par exemple). Par ailleurs, en complément des questionnaires remplis en hétéro-attribution, d'autres mesures de la personnalité perçue pourraient être mises en place. Des tests plus "projectifs" pourraient être utilisés. Par exemple, dans le cadre de notre interaction ludique, une situation de jeu (pour laquelle la décision de l'agent promotion serait différente de celle de l'agent prévention) serait présentée à l'utilisateur, qui de-

vrait alors "prédire" quel serait le choix de l'agent avec lequel il a joué.

7.2 réflexions générales et perspectives à long-terme

7.2.1 *Personnalité et informatique*

Dans la lignée de notre perspective socio-cognitive, nous pensons que l'informatique affective pourrait s'inspirer des frameworks intégratifs de la personnalité comme celui proposé par Mayer (2005). Dans ce framework, la personnalité peut être vue comme une forme de propriété émergente du système complexe formé par l'ensemble des sous-systèmes impliqués. Mais si la personnalité est une propriété émergente, un modèle de personnalité artificielle n'est-il pas qu'un pis-aller en attendant une modélisation computationnelle des sous-systèmes sus-cités ?

D'après Sun and Wilson (2014), toute architecture cognitive devrait pouvoir créer une personnalité sans ajout ou modification de ses mécanismes et de ses processus internes. Toujours d'après Sun and Wilson (2014), l'architecture CLARION permet de développer différentes personnalités simplement en manipulant les motivations du sous-système motivationnel. Des simulations montrent que statistiquement, CLARION peut simuler des interactions entre agents comparables à des interactions entre humains.

Mais dans ce cas, plutôt que de porter l'effort sur des modèles de personnalité, faudrait-il se centrer sur la modélisation fine de processus cognitifs en lien avec la personnalité ? L'étude computationnelle de la personnalité serait-elle alors uniquement la définition de l'ensemble des paramètres nécessaires à l'émergence de la personnalité dans ces processus ?

Cependant, il est à rappeler que dans cette optique intégrative, la personnalité intègre nombre d'adaptations caractéristiques ainsi que l'internalisation d'une biographie personnelle (McAdams and Pals, 2006). La personnalité est donc un phénomène complexe, qui se joue aux niveaux intra- comme interpersonnel et dans une interaction avec la sphère socio-culturelle. L'aspect "bac à sable" des architectures informatiques peut apporter une vue simplifiée, sans être simpliste, de ce phénomène en permettant de rejouer des situations identiques avec la modification d'un seul paramètre ou d'ajouter de façon incrémentale des processus supplémentaires pour observer les interactions.

Plus que de créer des machines avec une personnalité "reconnaissable", la place de l'informatique affective au sein d'une science de la personnalité est peut-être de fournir l'occasion d'observer l'émer-

gence de ces motifs émotionnels, motivationnels ou comportementaux récurrents que nous nommerons personnalité.

Concernant l'apport d'une personnalité artificielle à l'interaction homme-machine, un but recherché est souvent de proposer la personnalité "la plus adaptée" pour l'utilisateur. Généralement, l'atteinte de cet objectif est mesuré en termes d'appréciation de la machine par l'utilisateur, comme cela a été fait dans ce travail.

Dans un premier temps, l'idée qu'il existe une personnalité "plus adaptée" peut être discutée, principalement en regard du rôle de la machine. Des agents ou des robots destinés à occuper des fonctions précises, pour des interactions ponctuelles, peuvent effectivement avoir une personnalité particulière car les utilisateurs pourraient attendre certaines caractéristiques. Ces caractéristiques sont liées aux stéréotypes associés à ladite fonction : e. g. une hôtesse d'accueil est agréable et enjouée, un agent d'assurance est prudent et consciencieux... Par contre, pour un compagnon artificiel destiné à être en interaction quotidiennement avec son utilisateur et qui pourra tenir plusieurs rôles sociaux, une personnalité plus nuancée et surtout donnant l'impression que le compagnon est un individu à part entière serait peut-être plus adaptée ?

Dans un deuxième temps, est-il toujours souhaitable d'optimiser l'appréciation qu'une personne va reporter à propos de l'interaction qu'elle a pu avoir avec une machine ? Il est certes important que la machine n'apparaisse pas comme détestable. Mais le fait que la machine soit appréciée sur les courtes interactions proposées lors d'expérimentations assure-t-il que les utilisateurs soient prêts à interagir de façon plus durable ? Si la bonne réalisation de tâches communes peut être un élément d'évaluation pour des systèmes orientés tâche, quelles sont les variables pertinentes pour évaluer des systèmes sociaux destinés à être des compagnons artificiels ?

La découverte de ces variables pertinentes permettant de mesurer au mieux l'apport de systèmes affectifs dans l'interaction homme-machine est un objet de recherche important et ce n'est pas le seul défi méthodologique auquel soit confrontée l'informatique affective, particulièrement dans le domaine de la personnalité.

7.2.2 *Problématiques méthodologiques*

Dans le domaine particulier de la personnalité artificielle, la problématique méthodologique principale est probablement de savoir comment dépasser la formation d'impression pour aller vers un jugement de personnalité. Dans le modèle du continuum de la formation d'impression (Fiske and Neuberg, 1990), les individus se forment en premier lieu une impression de façon rapide et automatique, en s'appuyant sur de la catégorisation sociale et des stéréotypes. À l'autre

bout du continuum, la personne jugée n'est plus seulement considérée au travers de ce prisme catégoriel mais est évaluée de façon individuelle, comme un être unique. La question est donc de savoir comment encourager le positionnement des participants vers cette extrémité du continuum dans leur jugement d'une entité artificielle.

Plusieurs obstacles rendent cet objectif difficile à atteindre. Le premier est d'ordre pratique. Pour obtenir un véritable jugement de personnalité, des études à long-terme sont nécessaires. De telles études demandent des protocoles lourds, où les participants interagissent avec la machine de nombreuses fois, de façon régulière et idéalement dans différentes situations. Ce qui amène une nouvelle difficulté, d'ordre technologique : une telle interaction ne nécessite-t-elle pas un système global, avec de nombreuses capacités de perception, d'action, de mémorisation, un système capable de s'adapter et de prendre en compte différents contextes ? Or un tel système n'est pas disponible pour le moment, mais l'informatique affective ne doit pas s'interdire pour autant l'étude de phénomènes globaux comme la personnalité. De plus, la démarche scientifique encourage l'étude indépendante de chaque brique du système avant de tester le système complet. Des compromis expérimentaux doivent donc être faits, qui ne sont un problème que si les résultats desdites expérimentations ne sont pas discutés en regard de ces compromis.

Une autre problématique méthodologique, qui n'est pas propre à la question de la personnalité, tient à l'utilisation des questionnaires lors de l'évaluation de systèmes affectifs. En effet, l'outil de mesure le plus utilisé est probablement le questionnaire, que celui-ci soit une adaptation d'un questionnaire validé pour les humains ou une création directe des auteurs. Cependant, la psychométrie est une science et les construits ainsi mesurés sont complexes à mettre en évidence, comme discuté au chapitre 4 avec la création du RFQ-PF. La question de ce qui *est* mesuré est encore une fois un point important à discuter.

Le questionnement du construit n'est certes pas l'enjeu principal de la recherche en informatique affective, et les chercheurs de ce domaine sont plus des utilisateurs, que des créateurs, de questionnaires. Mais cette utilisation innovante de mesures issues de la psychologie éclaire une nouvelle facette du construit étudié. En plus d'être questionnée par l'informaticien, ces mesures devraient permettre d'initier un dialogue avec les psychologues qui s'intéressent justement à la définition du construit sous-jacent. Dans cette optique, des informations complémentaires, utiles à tous, pourraient être obtenues en mettant plus en avant les analyses qualitatives, comme par exemple les entretiens d'explicitation, dans l'évaluation des interactions homme-machine.

De façon plus générale encore, l'informatique affective doit également s'approprier la méthodologie de la psychologie expérimentale

pour l'étude de l'interaction homme-machine. Nous parlons d'appropriation au sens où l'application de cette méthodologie dans un cadre informatique soulève de nouvelles problématiques propres au domaine. Dans la présentation de nos expérimentations, nous avons par exemple souligné le problème particulier du choix des conditions "contrôle". Il est important de prendre en compte ce à quoi les systèmes évalués sont comparés par rapport au concept étudié. Dans le cas de la personnalité, ce souci est exacerbé par le fait qu'une impression est rapidement formée par les participants et que si la question est posée, il est fort probable qu'une personnalité particulière soit reportée.

Cet exemple est assez représentatif des challenges liés à l'articulation des deux domaines que sont la psychologie et l'informatique. Comme nous l'avions présenté en introduction, les expérimentations utilisateurs ne peuvent être considérées comme des "validations" des choix d'implémentation. S'il parait raisonnable de penser que s'appuyer sur un modèle psychologique permet d'obtenir une implémentation plus pertinente d'un concept affectif, il ne faut pas oublier que ces études, qu'elles soient perceptives ou en interaction, n'évaluent pas l'implémentation en elle-même mais seulement les sorties qu'elle produit. Ces études répondent donc à un enjeu de recherche différent bien que d'égale importance. Il est possible que de ce point de vue, l'implémentation soit tout autant "invalidable" par essence, comme peuvent l'être les modèles psychologiques théoriques, car ils réifient une internalité non-observable. Comme pour les modèles théoriques, la validation de choix d'implémentation se situe peut-être plutôt dans le fait que ces choix permettent de représenter de façon plus complète que d'autres les comportements humains simulés.

Pour résumer, la validation d'un modèle informatique, sa perception par un utilisateur et son utilité dans une interaction homme-machine sont des questions différentes. D'un point de vue méthodologique, pour utiliser les outils les mieux adaptés et en tirer les conclusions les plus pertinentes, il est donc important de bien spécifier en amont les enjeux de recherche visés.

7.2.3 *Articuler psychologie et informatique*

Dans l'articulation entre informatique et psychologie pour l'informatique affective, nous pouvons distinguer différents enjeux : améliorer l'interaction homme-machine, approfondir la connaissance de l'humain dans son lien avec la machine, étudier l'humain en simulant informatiquement certaines de ses caractéristiques. Néanmoins, on peut se demander s'il est toujours judicieux de vouloir concilier ces différents objectifs. Chacun de ces buts amène des implications méthodologiques et applicatives différentes.

La résolution de ces contraintes donne généralement lieu à des compromis, techniques ou expérimentaux, qui n'existeraient pas forcément dans le cadre d'études visant un seul de ces objectifs. Cependant, malgré ces difficultés supplémentaires, les travaux de recherche cherchant à mêler ces problématiques peuvent apporter de nouvelles vues et réflexions intéressantes sur des questions déjà abordées avec le prisme d'un seul de ces buts.

Par ailleurs, en s'inspirant de la psychologie cognitive, l'informatique affective n'est-elle pas en train de rejouer ses débats ? Globalement, ne retrouve-t-on pas l'esprit de l'opposition entre approche behavioriste et cognitiviste dans l'opposition entre implémentation connexionniste ou symbolique ? Des questions similaires concernant la pertinence d'ouvrir la "boite noire" (qu'il s'agisse de l'individu pour les uns ou de réseaux auto-organisés pour les autres) et les façons d'appréhender l'étude des comportements observables (de l'homme ou de la machine) semblent émerger. De façon plus globale, il semblerait que le nœud du problème se trouve dans l'articulation entre description et explication. Car s'il est possible d'avoir pour objectif de recherche l'une ou l'autre, les deux concepts s'avèrent complémentaires : les approches descriptives ne devraient pas se montrer réductionnistes par rapport aux objets qu'elles étudient comme les approches explicatives ne devraient pas se priver de mots et d'outils de mesure des phénomènes qu'elles cherchent à expliquer.

D'ailleurs, à l'instar de Tiberghien (1999), nous pensons que les deux ne s'excluent pas et qu'il est possible de refuser de prendre des positions (pratiques ou épistémologiques) extrêmes. L'avenir de la personnalité artificielle (et peut-être même de l'informatique affective) se trouve sans doute dans l'utilisation d'approches mixtes à tous les points de vue : de la méthodologie, en combinant approche theory-driven et data-driven, à l'évaluation, en associant mesures qualitatives et quantitatives, en passant par l'implémentation, à la fois symbolique et connexionniste.

Cinquième partie

ANNEXES

A

ÉLÉMENTS RELATIFS À LA CONSTRUCTION ET À LA VALIDATION DU QUESTIONNAIRE RFQ-PF

A.1 PROVERBES DE LA PRÉ-SÉLECTION ET REPRÉSENTATIVITÉ PERÇUE

Tableau 15 – Proverbes prévention pré-sélectionnés sur leur pertinence pour l'échantillon tout-venant et avis des experts

Proverbes	Pertinence	Expert 1	Expert 2
Dans le doute abstiens-toi	86%	4	3
Prudence est mère de sûreté	86%	3	3
Quand on est bien, il faut s'y tenir	84%	1	4
Il ne faut pas quitter le certain pour l'incertain	80%	3	3
On sait ce que on quitte mais on ne sait pas ce qu'on trouve	80%	4	4
Il faut de la modération en toute chose	78%	1	3
À chacun son métier, les vaches seront bien gardées	77%	2	1
Mieux vaut prévenir que guérir	75%	3	3
Il faut garder une poire pour la soif	75%	3	3
Il vaut mieux tenir que courir	75%	3	4
Qui veut voyager loin ménage sa monture	72%	3	3
Un homme averti en vaut deux	72%	3	2
De deux maux il faut choisir le moindre	71%	4	2
Il ne faut pas mettre tous ses œufs dans le même panier	71%	4	3
Il ne faut pas aller plus vite que la musique	71%	1	3
Chat échaudé craint l'eau froide	70%	3	1
Qui va piano va sano	67%	3	1
Plus on se hâte, moins on avance	64%	1	2
Un tiens vaut mieux que deux tu l'auras	62%	4	4
Hâtez-vous lentement	58%	2	2

Avis experts : 1- Non pertinent ; 2- Plutôt non pertinent ; 3- Plutôt pertinent ; 4- Pertinent

En italique, les 10 proverbes de la version initiale du RFQ-PF

Tableau 16 – Proverbes promotion pré-sélectionnés sur leur pertinence pour l'échantillon tout-venant et avis des experts

Proverbes	Pertinence	Expert 1	Expert 2
La fortune sourit aux audacieux	100%	4	3
Qui ne tente rien n'a rien	88%	4	4
La vie est faite pour être vécue	87%	2	3
Il faut prendre la balle au bond	84%	3	3
Quand on veut, on peut	84%	4	3
Il faut battre le fer quand il est chaud	83%	3	2
Premier arrivé, premier servi	81%	1	2
On ne fait pas d'omelette sans casser des œufs	80%	1	4
À cœur vaillant, rien d'impossible	78%	4	3
Impossible n'est pas français	74%	1	2
Celui qui n'est pas mort peut tout faire	74%	4	2
Le monde appartient à ceux qui se lèvent tôt	72%	2	1
Il faut faire tourner le moulin lorsque le vent souffle	72%	3	3
Qui veut la fin veut les moyens	72%	1	1
Qui craint de se mouiller ne prendra pas de truites	67%	3	3
Trop de précautions nuit	65%	3	3
Celui qui n'ose pas ne doit pas se plaindre de sa malchance	64%	3	4
Tant qu'il y a de la vie, il y a de l'espoir	64%	1	2
Trop de prudence n'atteint pas son but	64%	3	3
A vivre, on apprend toujours quelque chose	61%	3	1

Avis experts : 1- Non pertinent ; 2- Plutôt non pertinent ; 3- Plutôt pertinent ; 4- Pertinent

En italique, les 10 proverbes de la version initiale du RFQ-PF

A.2 RFQ-PF ; version finale ; 18 items

Pour chaque proverbe, indiquez dans quelle mesure celui-ci peut s'appliquer à votre façon de voir la vie en général (en entourant le chiffre correspondant)

	Pas du tout						Tout à fait
01- Qui ne tente rien n'a rien	1	2	3	4	5	6	7
02- Il ne faut pas quitter le certain pour l'incertain	1	2	3	4	5	6	7
03- Qui veut voyager loin ménage sa monture	1	2	3	4	5	6	7
04- Il vaut mieux tenir que courir	1	2	3	4	5	6	7
05- Il faut prendre la balle au bond	1	2	3	4	5	6	7
06- Dans le doute abstiens-toi	1	2	3	4	5	6	7
07- Trop de prudence n'atteint pas son but	1	2	3	4	5	6	7
08- Un tiens vaut mieux que deux tu l'auras	1	2	3	4	5	6	7
09- On sait ce qu'on quitte mais on ne sait pas ce qu'on trouve	1	2	3	4	5	6	7
10- À cœur vaillant, rien d'impossible	1	2	3	4	5	6	7
11- Il faut garder une poire pour la soif	1	2	3	4	5	6	7
12- Il faut faire tourner le moulin lorsque le vent souffle	1	2	3	4	5	6	7
13- Mieux vaut prévenir que guérir	1	2	3	4	5	6	7
14- La fortune sourit aux audacieux	1	2	3	4	5	6	7
15- Trop de précautions nuit	1	2	3	4	5	6	7
16- Prudence est mère de sûreté	1	2	3	4	5	6	7
17- Il ne faut pas mettre tous ses œufs dans le même panier	1	2	3	4	5	6	7
18- Quand on veut, on peut	1	2	3	4	5	6	7

Scoring :

Pro = (01 + 05 + 07 + 10 + 12 + 14 + 15 + 18) / 8

Pre = (02 + 03 + 04 + 06 + 08 + 09 + 11 + 13 + 16 + 17) / 10

A.3 TRADUCTION DU RFQ UTILISÉE ; ANGEL AND STEINER (2010)

	Cet ensemble de questions vous demande d'indiquer à quelle fréquence ces évènements arrivent ou sont arrivés dans votre vie.	**Jamais 1**	**Rarement 2**	**Parfois 3**	**Souvent 4**	**Très Souvent 5**
Pro	1. Par rapport à la plupart des gens, obtenez-vous généralement ce que vous voulez dans la vie ?	☐	☐	☐	☐	☐
Prev	2. Vous arrive-t-il de « dépasser les bornes » en faisant des choses non-autorisées (par vos parents, la société, etc.) ?	☐	☐	☐	☐	☐
Pro	3. Vous arrive-t-il de faire des choses qui vous donnent l'envie d'aller plus loin, de travailler encore plus, de vous dépasser?	☐	☐	☐	☐	☐
Prev	4. Avez-vous souvent mis vos parents sur les nerfs ?	☐	☐	☐	☐	☐
Prev	5. Obéissez-vous souvent aux règles établies (par vos parents, la société, etc.) ?	☐	☐	☐	☐	☐
Prev	6. Avez-vous déjà fait des choses considérées comme répréhensibles (par vos parents, la société, etc.) ?	☐	☐	☐	☐	☐
Pro	7. Réussissez-vous souvent les différentes choses que vous tentez ?	☐	☐	☐	☐	☐
Prev	8. Faites-vous souvent preuve de prudence ?	☐	☐	☐	☐	☐
Pro	9. En général, réussissez-vous, comme vous le souhaitez, les choses importantes pour vous ?	☐	☐	☐	☐	☐
Pro	10. Globalement, avez-vous souvent le sentiment de progresser dans la vie et d'aller dans le sens du succès ?	☐	☐	☐	☐	☐
Pro	11. Pour le moment, avez-vous trouvé des activités qui vous intéressent vraiment et qui vous motivent à fournir des efforts ?	☐	☐	☐	☐	☐

B

RÈGLES DU JEU CAN'T STOP

B.1 MATÉRIEL DE JEU

Le matériel de jeu se compose d'un plateau de jeu comprenant 11 colonnes numérotées de 2 à 12, de 9 pions "camps de base" pour chaque joueur (ici 2 joueurs), de 3 pions "grimpeurs" communs aux joueurs et de 4 dés à 6 faces (Figure 40).

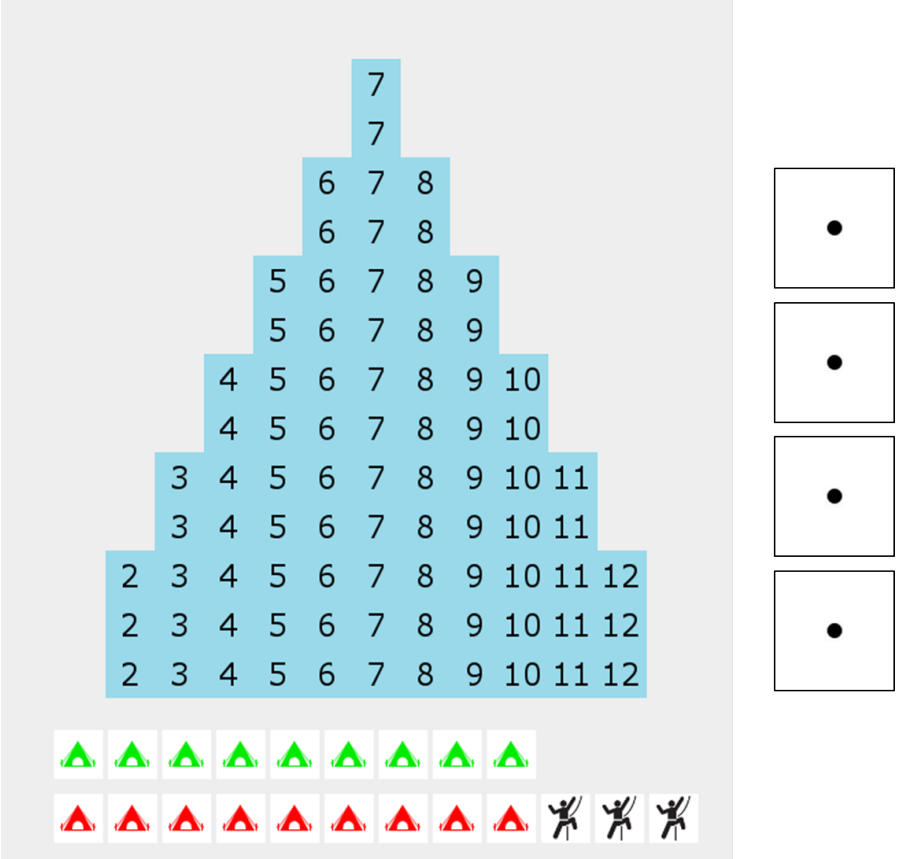


FIGURE 40 – Matériel du jeu Can't stop

B.2 BUT DU JEU

La partie se termine lorsqu'un des joueurs a réussi à poser des camps de base en haut de 3 colonnes différentes (quelle que soit leur hauteur).

B.3 Déroulement de la partie

Une partie se déroule en tours de jeu successifs. A son tour, un joueur réalise plusieurs phases dont les étapes sont décrites ci-après. Lorsque son tour se termine, il passe la main au joueur suivant et ainsi de suite jusqu'à la fin de la partie.

Un tour est composé de 1 ou plusieurs phases de jeu, chaque phase de jeu comprenant 4 étapes.

Etape 1 : lancer des dés Le joueur commence par lancer les 4 dés. Sur la Figure 41a, le résultat des dés est 2, 3, 4, 5.

Etape 2 : calcul des mouvements Un mouvement est décrit par une paire de chiffres calculée en additionnant les dés 2 à 2. Ces chiffres correspondent aux colonnes sur lesquelles vont pouvoir se positionner les grimpeurs. Après le premier lancer, il y a donc 3 mouvements possibles : (5 et 9), (6 et 8), (7 et 7) ()Figure 41b).

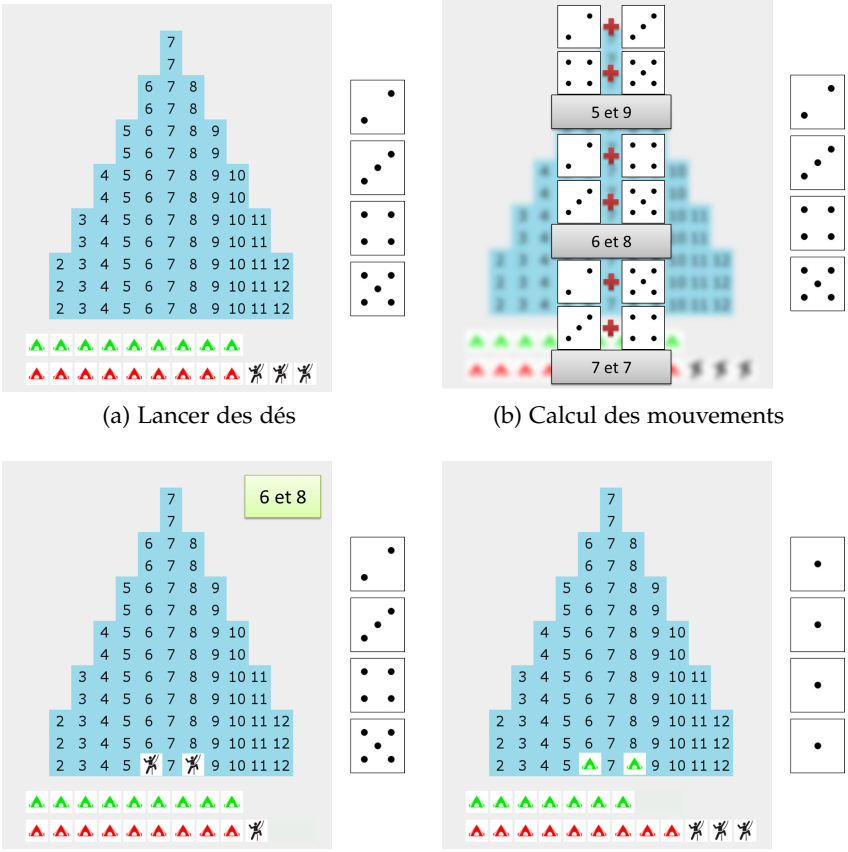
Etape 3 : choix du mouvement Le joueur en cours choisit alors les colonnes sur lesquelles il va faire avancer ses grimpeurs. Lorsque les grimpeurs sont à l'extérieur du plateau (début de tour), il est possible de les faire rentrer sur n'importe quelle colonne définie dans le mouvement. Sur la Figure 41c, le joueur choisit le mouvement (6 et 8). Un grimpeur est donc posé sur la première case de la colonne 6 et un autre sur la première case de la colonne 8.

Une fois les grimpeurs sur le plateau, ils ne peuvent pas changer de colonne pour ce tour. Les mouvements obtenus doivent donc correspondre aux colonnes déjà occupées par des grimpeurs pour pouvoir être jouables. Si après un lancer de dés, aucun des mouvements obtenus n'est possible, on dit qu'il y a "décrochage". Le tour du joueur courant s'arrête immédiatement, les grimpeurs sont retirés du plateau et c'est le tour du joueur suivant.

Etape 4 : décision stop ou encore Si le joueur a pu déplacer au moins 1 grimpeur, alors il peut choisir d'arrêter son tour ou de rejouer. S'il choisit de rejouer, une nouvelle phase de jeu commence et il reprend à la phase 1. S'il choisit de s'arrêter, les grimpeurs sont remplacés par ses camps de base (Figure 41d). Son avancée est alors sauvegardée. Lors de son prochain tour, si le joueur choisit de faire avancer un grimpeur sur une colonne où il a déjà un camp, le grimpeur entre sur le plateau sur la case directement supérieur au camp plutôt qu'en bas de la colonne.

Si un grimpeur est positionné sur la dernière case d'une colonne et que le joueur s'arrête, alors la colonne est fermée. Plus personne (ni ce

joueur, ni ses adversaires) ne peut utiliser cette colonne dans la suite du jeu. Lorsqu'un joueur a fermé trois colonnes, la partie s'arrête.



(a) Lancer des dés

(b) Calcul des mouvements

(c) Avancée des grimpeurs

(d) Pose des camps de base

FIGURE 41 – Illustration des 4 étapes composant une phase de jeu

En conclusion, le déroulement d'une partie à deux joueurs est décrit par le diagramme présenté en Figure 42.

Un tutoriel vidéo présentant les règles du jeu Can't Stop est disponible à l'adresse suivante :

https ://www.youtube.com/watch ?v=kQs6JBXotxw

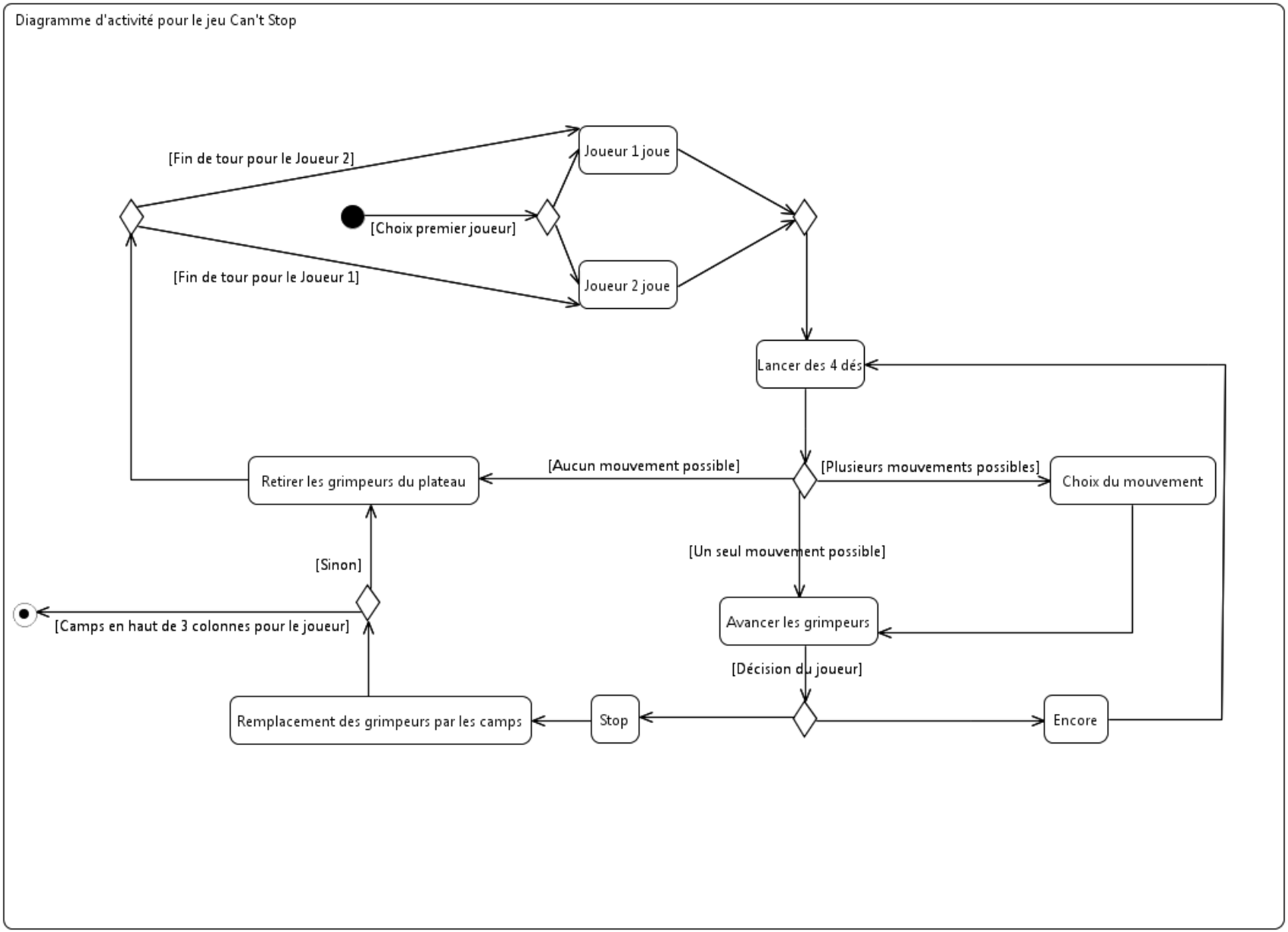


FIGURE 42 – Diagramme d'activité représentant une partie de Can't Stop

C

STATISTIQUES DESCRIPTIVES POUR LA SECONDE EXPÉRIMENTATION AVEC INTERACTIONS RÉPÉTÉES

C.1 PERCEPTION DU CHANGEMENT

Tableau 17 – Fréquences concernant la perception du changement

	Perception du changement			
	Ok		Erreur	
Avec MARC	Fréquence	%	Fréquence	%
RF-Pre	10	11%	6	7%
RF-Pro	13	15%	7	8%
Switch	9	10%	7	8%
Sans MARC				
RF-Pre	6	7%	5	6%
RF-Pro	6	7%	4	5%
Switch	9	10%	5	6%
Total	53	61%	34	39%

C.2 regulatory focus

Tableau 18 – Statistiques descriptives pour le score promotion

		Induction		
Avec MARC		Neutre	Pre	Pro
RF-Pre	moyenne	4,54	4,70	4,58
	ecart-type	1,07	1,15	1,02
	CQV	17%	16%	16%
	Finn coef.	0,72	0,67	0,74
RF-Pro	moyenne	3,71	4,08	4,68
	ecart-type	1,05	0,95	1,20
	CQV	21%	19%	16%
	Finn coef.	0,72	0,78	0,64
Switch	moyenne	3,99	4,40	4,77
	ecart-type	1,51	1,37	1,30
	CQV	28%	15%	10%
	Finn coef.	0,43	0,53	0,58
Sans MARC		Neutre	Pre	Pro
RF-Pre	moyenne	4,35	4,38	4,23
	ecart-type	1,27	1,43	1,23
	CQV	12%	24%	14%
	Finn coef.	0,60	0,47	0,62
RF-Pro	moyenne	4,49	4,45	4,28
	ecart-type	1,69	1,56	1,50
	CQV	35%	13%	26%
	Finn coef.	0,29	0,40	0,44
Switch	moyenne	4,22	4,07	4,36
	ecart-type	1,24	1,42	0,84
	CQV	15%	20%	10%
	Finn coef.	0,61	0,50	0,83

Tableau 19 – Statistiques descriptives pour le score prévention

		Induction		
		Neutre	Pre	Pro
Avec MARC				
RF-Pre	moyenne	3,71	3,87	4,05
	ecart-type	1,44	1,28	0,99
	CQV	10%	28%	14%
	Finn coef.	0,71	0,59	0,75
RF-Pro	moyenne	4,07	3,92	4,13
	ecart-type	1,27	0,89	1,20
	CQV	24%	15%	13%
	Finn coef.	0,60	0,80	0,64
Switch	moyenne	4,03	4,04	3,87
	ecart-type	1,38	1,33	1,15
	CQV	17%	25%	22%
	Finn coef.	0,53	0,56	0,67
Sans MARC		Neutre	Pre	Pro
RF-Pre	moyenne	3,89	3,81	3,72
	ecart-type	1,03	1,24	1,03
	CQV	18%	19%	10%
	Finn coef.	0,74	0,61	0,73
RF-Pro	moyenne	4,73	4,41	4,51
	ecart-type	1,37	1,39	1,16
	CQV	21%	19%	12%
	Finn coef.	0,53	0,52	0,67
Switch	moyenne	4,42	4,29	4,28
	ecart-type	1,16	1,49	1,30
	CQV	13%	12%	14%
	Finn coef.	0,66	0,45	0,58

C.3 crédibilité

Tableau 20 – Statistiques descriptives pour l'appréciation

		Induction		
Avec MARC		Neutre	Pre	Pro
RF-Pre	moyenne	2,81	2,62	2,27
	ecart-type	1,04	0,99	1,18
RF-Pro	moyenne	2,87	2,58	2,71
	ecart-type	0,93	1,18	1,03
Switch	moyenne	3,09	3,00	2,97
	ecart-type	0,96	0,89	0,85
Sans MARC		Neutre	Pre	Pro
RF-Pre	moyenne	2,33	2,29	2,36
	ecart-type	1,04	1,18	0,98
RF-Pro	moyenne	2,93	2,64	2,82
	ecart-type	1,05	1,12	1,12
Switch	moyenne	3,03	3,11	3,27
	ecart-type	0,89	1,07	1,16

Tableau 21 – Statistiques descriptives pour l'intelligence perçue

		Induction		
		Neutre	Pre	Pro
Avec MARC				
RF-Pre	moyenne	3,31	2,83	2,98
	ecart-type	0,93	1,07	1,26
RF-Pro	moyenne	3,12	2,88	3,27
	ecart-type	1,14	1,00	1,00
Switch	moyenne	3,32	3,53	3,41
	ecart-type	1,16	1,18	0,87
		Neutre	Pre	Pro
Sans MARC				
RF-Pre	moyenne	2,743	2,73	2,87
	ecart-type	1,16	0,76	1,05
RF-Pro	moyenne	3,39	3,27	3,42
	ecart-type	1,01	1,27	1,08
Switch	moyenne	3,36	3,56	3,29
	ecart-type	0,96	1,00	0,92

C.4 perception de l'interaction

Tableau 22 – Statistiques descriptives pour le Flow - Intérêt

Avec MARC		Session 1	Session 2	Session 3
RF-Pre	moyenne	3,60	3,12	3,24
	ecart-type	1,06	1,23	1,19
RF-Pro	moyenne	3,88	2,67	3,58
	ecart-type	0,85	1,23	1,13
Switch	moyenne	3,36	3,61	3,36
	ecart-type	0,85	0,87	0,64
Sans MARC		**Session 1**	**Session 2**	**Session 3**
RF-Pre	moyenne	3,50	3,25	3,04
	ecart-type	1,06	0,92	0,99
RF-Pro	moyenne	3,44	3,65	3,56
	ecart-type	0,88	1,00	1,09
Switch	moyenne	3,70	3,57	3,57
	ecart-type	0,70	0,90	0,73

Tableau 23 – Statistiques descriptives pour le Flow - Distraction

Avec MARC		Session 1	Session 2	Session 3
RF-Pre	moyenne	1,90	2,52	2,81
	ecart-type	1,34	1,60	1,60
	CQV	33%	50%	60%
RF-Pro	moyenne	2,46	2,58	2,12
	ecart-type	1,41	1,41	1,26
	CQV	60%	60%	50%
Switch	moyenne	1,94	2,33	2,67
	ecart-type	1,06	1,03	1,33
	CQV	47%	20%	30%
Sans MARC		**Session 1**	**Session 2**	**Session 3**
RF-Pre	moyenne	2,14	2,86	3,07
	ecart-type	1,46	1,56	1,49
	CQV	50%	52%	41%
RF-Pro	moyenne	2,65	2,00	2,47
	ecart-type	1,46	1,37	1,28
	CQV	60%	33%	50%
Switch	moyenne	2,27	1,93	2,00
	ecart-type	1,46	1,37	1,28
	CQV	50%	43%	50%

Tableau 24 – Statistiques descriptives pour la coprésence

		Session 2	Session 3
Avec MARC			
RF-Pre	moyenne	2,28	2,81
	ecart-type	1,02	0,98
RF-Pro	moyenne	2,61	2,81
	ecart-type	0,97	0,97
Switch	moyenne	3,10	2,97
	ecart-type	0,92	0,86
		Session 2	Session 3
Sans MARC			
RF-Pre	moyenne	2,73	2,47
	ecart-type	1,19	1,10
RF-Pro	moyenne	2,96	2,77
	ecart-type	0,87	0,89
Switch	moyenne	2,90	2,98
	ecart-type	0,93	1,09

D

FORMULAIRES DE CONSENTEMENT ET AUTORISATIONS

D.1 PREMIÈRE EXPÉRIMENTATION (ADULTES)

Original à conserver par le laboratoire

Consentement pour la participation à une expérience

Je soussigné(e) : **Né(e) le :**

Demeurant :

certifie donner librement mon accord pour participer gracieusement à l'expérience décrite ci-dessous :

Description de l'expérience :
Ce travail de recherche s'inscrit dans le cadre du projet « MoCA ; Mon petit monde de compagnons » visant à examiner les interactions entre humains et compagnons artificiels (agents virtuels, robots). La présente étude concerne seulement l'évaluation des stratégies de l'humain au cours d'un jeu de plateau sur ordinateur.

Objectifs de l'expérience :
Le but de cette étude est d'enregistrer l'interaction entre deux humains ou l'interaction entre un humain et un agent virtuel au cours d'un jeu et d'évaluer différentes dimensions psychologiques afin de pouvoir créer un agent virtuel présentant des caractéristiques de personnalité.

L'expérience se déroule de la façon suivante :
Lors de cette expérimentation, un expérimentateur de l'équipe Cognition, Perception et Usages (CPU) du LIMSI-CNRS, UPR 3251 (Bâtiment 508, Rue John von Neumann, Campus Université Paris-Sud, 91403 Orsay; contact mail : caroline.faur@limsi.fr, sous la direction du Pr Jean-Claude Martin, martin@limsi.fr) vous expliquera les règles du jeu de plateau auquel vous allez jouer ainsi que le fonctionnement de l'interface. Vous jouerez ensuite une partie avec un adversaire (humain ou agent virtuel). L'ensemble de la partie sera enregistrée via une webcam. Les expérimentateurs vous demanderont de répondre ensuite à un questionnaire concernant vos caractéristiques personnelles et psychologiques.

Durée de l'expérience : environ 20min

Je déclare avoir été expressément informé(e) du projet du CNRS de recueillir les données que je lui fournirai puis de les étudier et analyser dans le cadre des recherches mentionnées ci-dessus.

Je comprends que ma participation n'est pas obligatoire et que je peux stopper ma participation à tout moment sans avoir à me justifier ni encourir aucune responsabilité. Mon consentement ne décharge pas les organisateurs de la recherche de leurs responsabilités et je conserve tous mes droits garantis par la loi[1].

A cet effet, j'autorise expressément le LIMSI, d'une part, à me filmer et, d'autre part, à collecter des informations fournies via des questionnaires. Je cède tous mes droits sur ces enregistrements et données pour une durée illimitée à compter de la signature de la présente autorisation. Je comprends que les données filmées et enregistrées ainsi que les informations recueillies sont strictement confidentielles et à usage exclusif des investigateurs concernés, pour des fins de recherche.

J'ai été informé(e) que mon identité n'apparaîtra dans aucun rapport ou publication et que toute information me concernant sera traitée de façon confidentielle. J'accepte que les images et données enregistrées et fournies à l'occasion de cette étude puissent être conservées dans une base de données et faire l'objet d'un traitement informatisé non nominatif par le LIMSI, unité de recherche du CNRS.

En conséquence de quoi, je garantis le CNRS contre tout recours et/ou action que pourraient former les personnes physiques ou morales estimant avoir des droits quelconques à faire valoir sur l'utilisation des données.

Je reconnais et accepte que le CNRS conserve toute liberté pour exploiter ou ne pas exploiter intégralement ou par extraits les données.

Je déclare que la présente autorisation est accordée pour le monde entier et pour une durée de 5 (cinq) an (s) à compter de la signature de la présente, renouvelable par tacite reconduction, sauf dénonciation de la part du signataire.

Fait à ……. le …../…../…….
Signature précédée de la mention « Lu et approuvé »

[1] Conformément à la loi « informatique et libertés » du 6 janvier 1978 modifiée en 2004, vous pouvez obtenir communication et le cas échéant, rectification ou suppression des informations vous concernant, en vous adressant au CNRS, 3 rue Michel-Ange, 75794 PARIS CEDEX 16.

D.2 Seconde expérimentation (adolescents)

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

LABORATOIRE D'INFORMATIQUE POUR LA MECANIQUE ET LES SCIENCES DE L'INGENIEUR

Céline Clavel
Maître de Conférences en Psychologie
Université Paris Sud & LIMSI CNRS
celine.clavel@limsi.fr
01 69 85 81 05
06 72 79 20 82

Orsay, le 05 janvier 2016

A l'attention des parents d'élèves

Objet : Participation à un atelier « jeu vidéo »

Notre équipe de recherche réalise actuellement une étude auprès de tous les élèves de 4ème du collège Alain Fournier.

Cette étude s'inscrit dans le cadre d'un projet national en collaboration avec différents laboratoires de recherche sur le développement des compagnons artificiels (personnages virtuels et robots personnels) capables d'établir et d'entretenir une relation sur le long terme avec un usager.

Nous souhaiterions proposer à votre enfant de participer à cet atelier, à raison de quatre séances d'une heure par semaine, entre le 11/01/2016 et le 19/02/2016, afin de lui faire découvrir notre plateforme d'agents virtuels et de lui permettre au cours de ces séances d'interagir avec eux sous une forme ludique. Notre objectif est d'évaluer dans quelles mesures différentes caractéristiques des compagnons artificiels contribuent à l'intérêt, l'engagement et l'implication de l'élève dans l'interaction, favorisent la compréhension de la "psychologie" du compagnon artificiel et soutiennent le développement d'une relation sur le long terme.

Cet atelier sera également l'occasion d'un échange avec les élèves sur la recherche et le développement informatique, et de leur présenter l'université Paris Saclay.

Monsieur Arieu, le principal du collège Alain Fournier, a donné son accord pour collaborer à cette recherche et a accepté de nous accueillir au sein de son établissement pour que nous mettions en place cet atelier durant quelques séances.

Nous souhaiterions ainsi pouvoir proposer aux élèves volontaires d'utiliser notre jeu et de répondre à des questionnaires sur leur expérience de jeu. Nous avons besoin de votre accord pour que votre enfant puisse effectivement participer à cette étude.

Cette recherche est tout à fait anonyme et n'a aucune relation avec les évaluations scolaires des élèves. Les productions des enfants sont strictement confidentielles et aucun résultat individuel ne sera communiqué. Seuls les chercheurs de l'équipe auront accès à ces données. Les résultats de cette étude seront exprimés sous forme de statistiques d'ensemble et vous seront communiqués si vous le souhaitez.

Nous vous remercions de contribuer au bon déroulement de cette étude en autorisant votre enfant à y participer. A cette fin, nous vous demandons de remplir et signer le document ci-joint puis de le remettre à votre enfant.

Nous vous prions d'agréer, Madame, Monsieur, l'expression de nos très cordiales salutations.

Céline Clavel & les membres de l'équipe du projet MOCA

Autorisation de fixation, de reproduction et d'exploitation de l'image et de la voix d'un mineur

Je soussigné(e) : Nom, Prénom :

Père Mère Tuteur *(rayer les mentions inutiles)*

Adresse :

Certifie

- donner librement mon consentement à la présente autorisation et
- donner mon accord pour que l'enfant dont je suis le responsable légal (ci-après l'Enfant) :

Nom, Prénoms :

Né(e) le :

Sexe : Masculin Féminin *(rayer les mentions inutiles)*

participe gracieusement à l'expérience décrite en annexe (ci-après le Projet) et aux enregistrements et photographies associées (ci-après les « Enregistrements »).

Déclare avoir été expressément informé(e) (i) du Projet du CNRS (Centre National de la Recherche Scientifique) de filmer l'Enfant puis éventuellement d'en faire un montage aux fins de travaux de recherche ou de diffusions ultérieurs et (ii) plus précisément du fait que la réalisation des Enregistrements lié au Projet peut nécessiter la présence d'une équipe de tournage afin d'enregistrement d'images fixes et/ou animées et/ou de la voix.

En conséquence de quoi et conformément aux dispositions relatives au droit à l'image, à la voix et aux droits de la personnalité, j'autorise expressément le CNRS dans le cadre du Projet, à titre gratuit et non exclusif, à :

- fixer et reproduire l'image fixe et/ou animée et/ou la voix de l'Enfant, en tout ou partie, ensemble et/ou séparément, à titre gracieux, par tous procédés, sur tous supports, en tous formats (tels que notamment photographies papiers ou numériques, films vidéos ou numériques) et ce en vue de leur communication au public.
- exploiter ou faire exploiter les Enregistrements, avec le logo et/ou la mention du CNRS et/ou du LIMSI, intégralement ou par extraits, à des fins de recherche et éventuellement à des fins de valorisation économique de la recherche (intégration dans des bases de données, …), sous toutes formes et tous supports par tous modes d'exploitation connus ou inconnus à ce jour, par tous réseaux de transmission ainsi que par tous réseaux de communications électroniques (par téléchargement et/ou en mode streaming) et en tous formats.

Je garantis que l'Enfant n'est pas lié par un contrat exclusif relatif à son image et/ ou à sa voix.

En conséquence de quoi, je garantis le CNRS contre tout recours et/ou action que pourraient former les personnes physiques ou morales estimant avoir des droits quelconques à faire valoir sur l'utilisation des Enregistrements.

Je reconnais et accepte que le CNRS conserve toute liberté pour exploiter ou ne pas exploiter intégralement ou par extraits les Enregistrements.

Je déclare que la présente autorisation est accordée (i) pour le monde entier et (ii) pour une durée de 5 an (s) à compter de la signature des présentes, renouvelable par tacite reconduction, sauf dénonciation de la part du signataire.

Fait à le

Signature précédée de la mention « Lu et approuvé »

Conformément à la loi « informatique et libertés » du 6 janvier 1978 modifiée en 2004, vous pouvez obtenir communication et le cas échéant, rectification ou suppression des informations vous concernant, en vous adressant au CNRS, 3 rue Michel-Ange, 75794 PARIS CEDEX 16.

E

PUBLICATIONS

CONFÉRENCES INTERNATIONALES

Faur, C., Caillou, P., Martin J.-C. and Clavel C. (2015) A Socio-cognitive Approach to Personality : Machine-learned Game Strategies as Cues of Regulatory Focus. In *Proceedings of 6th biannual Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII 2015)*, pp 581-587, September 21-24, Xi'An, China.

Faur, C., Martin, J.-C. and Clavel, C. (2015) Matching artificial agents' and users' personalities : designing agents with regulatory-focus and testing the regulatory fit effect. In *Proceedings of 37th annual meeting of the Cognitive Science Society (CogSci 2015)*, pp 662-667, July 23-25, Pasadena, USA.

Faur, C., Clavel, C. and Martin, J.-C. (2013) PERSEED : a Self-based Model of Personality for Virtual Agents Inspired by Socio-cognitive Theories. In *Proceedings of 5th biannual Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII 2013)*, pp 467-472, September 2-5, Geneva, Switzerland.

Clavel, C., **Faur, C.**, Martin, J.-C., Pesty, S. and Duhaut D. (2013) Artificial Companions with Personality and Social Role. In *Proceedings of the 2013 IEEE Symposium on Computational Intelligence for Creativity and Affective Computing (CICAC 2013)*, pp. 87-95, April 16-19, Singapore.

CONFÉRENCES NATIONALES

Pécune, F., **Faur, C**., Ochs, M., Clavel, C., Pelachaud, C. and Martin J.-C. (2014) *Modélisation de l'influence de la personnalité d'un compagnon artificiel sur ses attitudes sociales*. Workshop Affects, Compagnons Artificiels et Interaction (WACAI 2014), June 30 - July 1, Rouen, France.

Faur, C., Clavel, C. and Martin, J.-C. (2013) *La personnalité comme catalyseur de l'intercompréhension entre un utilisateur et un agent virtuel : le modèle PERSEED*. Conférence III - Intercompréhension de l'intraspécifique à l'interspécifique (III 2013), September 30 - October 1, Lorient, France.

CHAPITRE D'OUVRAGE

Faur, C., Clavel, C. and Martin, J.-C. (2015) La personnalité comme catalyseur de l'intercompréhension entre un utilisateur et un agent

virtuel : le modèle PERSEED. Dans M. Grandgeorge, F. Pugnière-Saavedra, B. Le Pevedic et C. Jost (éds.), *Vers une communication homme-animal-machine ? Contribution interdisciplinaire*, pp. 17-42, EME Editions.

communications orales

Faur, C., Clavel, C. and Martin J.-C. (2014) *Endowing Artificial Companions with Personality Using Regulatory-focus*. 28th the International Congress of Applied Psychology (ICAP 2014), July 08-13, Paris, France, 2014.

Faur, C., Clavel, C. and Martin, J.-C. (2014) *Affective Interactions with Virtual Agents : Theories and Models for Emotions and Personality*. 28th International Congress of Applied Psychology (ICAP 2014), July 08-13, Paris, France.

en cours de soumission

Faur, C., Martin, J.-C. and Clavel, C. (2016) Measuring Chronic Regulatory Focus with Proverbs : Development and Psychometrics Properties of a French Scale.

Allport, G. (1937). *Personality : A psychological interpretation*. Holt, New York.

André, E., Klesen, M., Gebhard, P., Allen, S., and Rist, T. (2000). Integrating Models of Personality and Emotions into Lifelike Characters. In *Affective interactions*, pages 150–165. Springer, Berlin Heidelberg.

Angel, V. and Steiner, D. D. (2010). Questionnaire des Focus Régulateurs Chroniques (QFRc) - version ajustée. Non publié, Université de Nice-Sophia Antipolis.

Ayduk, z. and Gyurak, A. (2008). Applying the Cognitive-Affective Processing Systems Approach to Conceptualizing Rejection Sensitivity. *Social and Personality Psychology Compass*, 2(5) :2016–2033.

Bailenson, J. N., Yee, N., Merget, D., and Schroeder, R. (2006). The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. *Presence : Teleoperators and Virtual Environments*, 15(4) :359–372.

Bandura, A. (1986). *Social Foundations of Thought and Action : A Social Cognitive Theory*. Prentice-Hall, Englewood Cliffs, New Jersey.

Baron-Cohen, S., Leslie, A., and Frith, U. (1985). Does the autistic child have a "theory of mind" ? *Cognition*, 21(1) :37–46.

Bartle, R. (1996). Hearts, clubs, diamonds, spades : Players who suit MUDs. *Journal of MUD research*, 1(1).

Bartneck, C., Kulić, D., Croft, E., and Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, 1(1) :71–81.

Becker, C., Prendinger, H., Ishizuka, M., and Wachsmuth, I. (2005). Evaluating affective feedback of the 3d agent max in a competitive cards game. In *Affective computing and intelligent interaction*, pages 466–473. Springer.

Benfield, J. A., Szlemko, W. J., and Bell, P. A. (2007). Driver personality and anthropomorphic attributions of vehicle personality relate to reported aggressive driving tendencies. *Personality and Individual Differences*, 42(2) :247–258.

Benyon, D. and Mival, O. (2010). From human-computer interactions to human-companion relationships. In *Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia (IITM 2010)*, Allahabad, India.

Bevacqua, E., De Sevin, E., Hyniewska, S. J., and Pelachaud, C. (2012). A listener model : introducing personality traits. *Journal on Multimodal User Interfaces*, 6(1-2) :27–38.

Bevacqua, E., de Sevin, E., Pelachaud, C., McRorie, M., and Sneddon, I. (2010). Building credible agents : Behaviour influenced by personality and emotional traits. In *International Conference on Kansei Engineering and Emotion Research (KEER 2010)*, Paris, France.

Bianco, A. T., Higgins, E. T., and Klem, A. (2003). How "Fun/Importance" Fit Affects Performance : Relating Implicit Theories to Instructions. *Personality and Social Psychology Bulletin*, 29(9) :1091–1103.

Block, J. (1993). Studying personality the long way. In Funder, D., Parke, R., Tomlinson-Keasey, C., and Block, J., editors, *Studying lives through time : Personality and development*, pages 9–41. American Psychological Association, Washington, DC.

Bonett, D. G. (2006). Confidence interval for a coefficient of quartile variation. *Computational Statistics & Data Analysis*, 50(11) :2953–2957.

Bosse, T. and Zwanenburg, E. (2014). Do Prospect-Based Emotions Enhance Believability of Game Characters ? A Case Study in the Context of a Dice Game. *Affective Computing, IEEE Transactions on*, 5(1) :17–31.

Branje, S. J., Van Aken, M. A., Van Lieshout, C. F., and Mathijssen, J. J. (2003). Personality judgments in adolescents' families : The perceiver, the target, their relationship, and the family. *Journal of Personality*, 71(1) :49–81.

Burgoon, J. K., Bonito, J. A., Bengtsson, B., Cederberg, C., Lundeberg, M., and Allspach, L. (2000). Interactivity in human–computer interaction : A study of credibility, understanding, and influence. *Computers in Human Behavior*, 16(6) :553–574.

Caci, H., Deschaux, O., and Baylé, F. J. (2007). Psychometric properties of the French versions of the BIS/BAS scales and the SPSRQ. *Personality and Individual Differences*, 42(6) :987–998.

Camacho, C. J., Higgins, E. T., and Luger, L. (2003). Moral value transfer from regulatory fit : What feels right is right and what feels wrong is wrong. *Journal of Personality and Social Psychology*, 84(3) :498–510.

Cattell, R. (1943). The description of personality : Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, 38(4) :476–506.

Cervone, D. (2005). Personality Architecture : Within-Person Structures and Processes. *Annual Review of Psychology*, 56(1) :423–452.

Cervone, D. (2006). Systèmes de personnalité au niveau de l'individu : vers une évaluation de l'architecture sociocognitive de la personnalité. *Psychologie française*, 51(3) :357–376.

Costa, P. T. and McCrae, R. R. (1992a). Four ways five factors are basic. *Personality and Individual Differences*, 13(6) :653–665.

Costa, P. T. and McCrae, R. R. (1992b). *Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO FFI) : Professional manual*. Psychological Assessment Resources, Odessa, FL, USA.

Costa, P. T. and McCrae, R. R. (1995). Domains and facets : Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of personality assessment*, 64(1) :21–50.

Courgeon, M., Celine, C., and Martin, J.-C. (2014). Modeling facial signs of appraisal during interaction : impact on users' perception and behavior. In *Proceedings of the 13th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2014)*, pages 765–772, Paris, France.

Courtney, M. G. R. and Gordon, M. (2013). Determining the number of factors to retain in EFA : Using the SPSS R-Menu v2.0 to make more judicious estimations. *Practical Assessment, Research & Evaluation*, 18(8) :1–14.

Csikszentmihalyi, M. (1997). *Finding flow : The psychology of engagement with everyday life*. Basic Books.

Damian, R. I. and Robins, R. W. (2012). Investigations into the Human Self : A Naturalist Perspective. *Social Cognition*, 30(4) :431–448.

De Bruin, L., Van Elk, M., and Newen, A. (2012). Reconceptualizing second-person interaction. *Frontiers in human neuroscience*, 6.

Debanne, T., Angel, V., and Fontayne, P. (2014). Decision-making during games by professional handball coaches using regulatory focus theory. *Journal of Applied Sport Psychology*, 26(1) :111–124.

Egges, A., Kshirsagar, S., and Magnenat-Thalmann, N. (2004). Generic Personality and Emotion Simulation for Conversational Agents. *Computer Animation and Virtual Worlds*, 15(1) :1–13.

Elliott, C. (1994). Research problems in the use of as Shallow Artificial Intelligence model of personality and emotion. In *Proceedings of the 12th National Conference on Artificial Intelligence (AAAI 1994)*, pages 9–15, Seattle, WA, USA.

Epley, N., Waytz, A., and Cacioppo, J. (2007). On Seeing Human : A Three-Factor Theory of Anthropomorphism. *Psychological Review*, 114(4) :868–886.

Eysenck, H. (1970). *The structure of human personality*. Methuen, London, UK, 3rd edition.

Fein, E. C. and Klein, H. J. (2011). Personality Predictors of Behavioral Self-Regulation : Linking behavioral self-regulation to five-factor model factors, facets, and a compound trait. *International Journal of Selection and Assessment*, 19(2) :132–144.

Finn, R. H. (1970). A note on estimating the reliability of categorical data. *Educational and Psychological Measurement*, 30(1) :71–76.

Fiske, S. T. and Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes : Influences of information and motivation on attention and interpretation. *Advances in experimental social psychology*, 23 :1–74.

Fodor, J. A. (1983). *The modularity of mind : An essay on faculty psychology*. MIT press.

Friedman, R. S. and Förster, J. (2001). The effects of promotion and prevention cues on creativity. *Journal of personality and social psychology*, 81(6) :1001–1013.

Gaudreau, P., Sanchez, X., and Blondin, J.-P. (2006). Positive and negative affective states in a performance-related setting : Testing the factorial structure of the PANAS across two samples of French-Canadian participants. *European Journal of Psychological Assessment*, 22(4) :240–249.

Gebhard, P. (2005). ALMA : a layered model of affect. In *Proceedings of the 4th international joint conference on Autonomous Agents and Multiagent Systems (AAMAS 2005)*, pages 29–36, Utrecht, Netherlands.

Gmytrasiewicz, P. J. and Lisetti, C. L. (2002). Emotions and personality in agent design and modeling. In *Game theory and decision theory in agent-based systems*, pages 81–95. Springer.

Goldberg, L. R. (1982). From Ace to Zombie : Some explorations in the Language of Personality. In Spielberger, C. and Butcher, J., editors, *Advances in Personality Assessment*, volume 1, pages 203–234. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Grant, H. and Higgins, E. T. (2003). Optimism, promotion pride, and prevention pride as predictors of quality of life. *Personality and Social Psychology Bulletin*, 29(12) :1521–1532.

Gratch, J., Marsella, S., Wang, N., and Stankovic, B. (2009). Assessing the validity of appraisal-based models of emotion. In *Proceedings of the 3rd International Conference on Affective Computing and Intelligent Interaction (ACII 2009)*, pages 147–154, Amsterdam, Netherlands.

Gray, J. A. (1990). Brain systems that mediate both emotion and cognition. *Cognition & Emotion*, 4(3) :269–288.

Hamstra, M. R., Bolderdijk, J. W., and Veldstra, J. L. (2011). Everyday risk taking as a function of regulatory focus. *Journal of Research in Personality*, 45(1) :134–137.

Harlow, R. E., Friedman, R. S., and Higgins, E. T. (1997). The Regulatory Focus Questionnaire. Unpublished, Columbia University.

Hayes, P. and Ford, K. (1995). Turing test considered harmful. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI 1995)*, pages 972–977, Montréal, Québec, Canada.

Higgins, E. T. (1987). Self-Discrepancy : A Theory Relating Self and Affect. *Psychological Review*, 94(3) :319–340.

Higgins, E. T. (1997). Beyond pleasure and pain. *American psychologist*, 52(12) :1280–1300.

Higgins, E. T. (2005). Value from regulatory fit. *Current Directions in Psychological Science*, 14(4) :209–213.

Higgins, E. T., Friedman, R. S., Harlow, R. E., Idson, L. C., Ayduk, O. N., and Taylor, A. (2001). Achievement orientations from subjective histories of success : Promotion pride versus prevention pride. *European Journal of Social Psychology*, 31(1) :3–23.

Higgins, T. and Tykocinski, O. (1992). Self-discrepancies and biographical memory : Personality and cognition at the level of psychological situation. *Personality and Social Psychology Bulletin*, 18(5) :527–535.

Holmes, G., Pfahringer, B., Kirkby, R., Frank, E., and Hall, M. (2002). Multiclass alternating decision trees. In *Proceedings of the 13th European Conference on Machine Learning (ECML 2002)*, pages 161–172, Helsinki, Finland.

Hooper, D., Coughlan, J., and Mullen, M. R. (2008). Structural Equation Modelling : Guidelines for Determining Model Fit. *Electronic Journal of Business Research Methods*, 6(1) :53–60.

Hoyle, R. H. (2010). Personality and self-regulation. In Hoyle, R. H., editor, *Handbook of personality and self-regulation*, pages 1–18. Wiley-Blackwell, Oxford, UK.

Ijaz, K., Bogdanovych, A., and Simoff, S. (2011). Enhancing the believability of embodied conversational agents through environment-, self- and interaction-awareness. In *Proceedings of the 34th Australasian Computer Science Conference (ACSC 2011)*, pages 107–116, Perth, Australia.

In'nami, Y. and Koizumi, R. (2013). Structural Equation Modeling in Educational Research. In Khine, M. S., editor, *Application of Structural Equation Modeling in Educational Research and Practice*, pages 23–51. SensePublishers, Rotterdam, Netherlands.

Isbister, K. and Doyle, P. (2002). Design and evaluation of embodied conversational agents : A proposed taxonomy. In *The First International Joint Conference on Autonomous Agents & Multi-Agent Systems (AAMAS 2002)*, Bologna, Italy.

Isbister, K. and Nass, C. (2000). Consistency of personality in interactive characters verbal cues, non-verbal cues, and user characteristics. *International Journal of Human-Computer Studies*, 53 :251–267.

James, W. (1890). *Principles of psychology*, volume 1. Henry Holt and Company, New York.

Johansson, M. and Verhagen, H. (2014). Social believability in games - the early years. In *Proceedings of Social Believability in Games Workshop, Colocated with the 9th International Conference on the Foundations of Digital Games (FDG 2014)*, Fort Lauderdale, FL, USA.

Johnson, D. and Gardner, J. (2010). Personality, motivation and video games. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction (OZCHI 2010)*, pages 276–281, Brisbane, Australia.

Joosse, M., Lohse, M., Perez, J. G., and Evers, V. (2013). What you do is who you are : The role of task context in perceived social robot personality. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2013)*, pages 2134–2139, Karlsruhe, Germany.

Karwowski, M. (2011). Teacher personality as predictor of perceived climate for creativity. *International Journal of Creativity and Problem Solving*, 21(1) :37–52.

Keller, J. (2008). On the development of regulatory focus : the role of parenting styles. *European Journal of Social Psychology*, 38(2) :354–364.

Klein, S. B. (2012). "What is the Self ?" : Approaches to a Very Elusive Question. *Social Cognition*, 30(4) :363–366.

Klimstra, T. A., Hale III, W. W., Raaijmakers, Q. A., Branje, S. J., and Meeus, W. H. (2009). Maturation of personality in adolescence. *Journal of personality and social psychology*, 96(4) :898–912.

Kshirsagar, S. and Magnenat-Thalmann, N. (2002). A multilayer personality model. In *Proceedings of the 2nd international symposium on Smart graphics (SG 2002)*, pages 107–115, Hawthorne, NY, USA.

Lardi, C., Billieux, J., d'Acremont, M., and Linden, M. V. d. (2008). A French adaptation of a short version of the Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ). *Personality and Individual Differences*, 45(8) :722–725.

Lee, A. Y. and Aaker, J. L. (2004). Bringing the Frame Into Focus : The Influence of Regulatory Fit on Processing Fluency and Persuasion. *Journal of Personality and Social Psychology*, 86(2) :205–218.

Lee, K. M., Peng, W., Jin, S.-A., and Yan, C. (2006). Can robots manifest personality ? : An empirical test of personality recognition, social responses, and social presence in human–robot interaction. *Journal of communication*, 56(4) :754–772.

Leite, I., Mascarenhas, S., Pereira, A., Martinho, C., Prada, R., and Paiva, A. (2010). "Why Can't We Be Friends ?" An Empathic Game Companion for Long-Term Interaction. In *Proceedings of the 10th International Conference on Intelligent Virtual Agents (IVA 2010)*, pages 315–321, Philadelphia, PA, USA.

Leonardelli, G. J., Lakin, J. L., and Arkin, R. M. (2007). A regulatory focus model of self-evaluation. *Journal of Experimental Social Psychology*, 43(6) :1002–1009.

Leys, C. and Schumann, S. (2010). A nonparametric method to analyze interactions : The adjusted rank transform test. *Journal of Experimental Social Psychology*, 46(4) :684–688.

Lim, M. Y., Dias, J., Aylett, R., and Paiva, A. (2012). Creating adaptive affective autonomous NPCs. *Autonomous Agents and Multi-Agent Systems*, 24(2) :1–25.

Lisetti, C. L. (2002). Personality, Affect and Emotion Taxonomy for Socially Intelligent Agents. In *Proceedings of the 15th International Florida Artificial Intelligence Research Society Conference (FLAIRS 2002)*, pages 397–401, Pensacola, FL, USA.

Liu, K., Tolins, J., Fox Tree, J., Neff, M., and Walker, M. (2015). Two techniques for assessing virtual agent personality. *IEEE Transactions on Affective Computing*, 7(1) :94–105.

Lockwood, P., Jordan, C. H., and Kunda, Z. (2002). Motivation by positive or negative role models : regulatory focus determines who will best inspire us. *Journal of personality and social psychology*, 83(4) :854–864.

Ludolph, R. and Schulz, P. J. (2015). Does regulatory fit lead to more effective health communication ? A systematic review. *Social Science & Medicine*, 128 :142–150.

Mayer, J. (2005). A Tale of Two Visions : Can a New View of Personality Help Integrate Psychology. *American Psychologist*, 60(4) :294–307.

McAdams, D. P. and Pals, J. L. (2006). A new Big Five : fundamental principles for an integrative science of personality. *American Psychologist*, 61(3) :204–217.

McCrae, R. R. and Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of personality and social psychology*, 52(1) :81–90.

Mehrabian, A. (1996). Analysis of the Big-five Personality Factors in Terms of the PAD Temperament Model. *Australian Journal of Psychology*, 48(2) :86–92.

Mendelson, M. and Aboud, F. (1997). Measuring friendship quality in late adolescents and young adults : McGill Friendship Questionnaires. Unpublished manuscript, retrieved from www.midss.ie.

Mendelson, M. J. and Aboud, F. E. (1999). Measuring friendship quality in late adolescents and young adults : McGill Friendship Questionnaires. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 31(2) :130–132.

Mischel, W. and Shoda, Y. (1995). A Cognitive-Affective System Theory of Personality : Reconceptualizing Situations, Dispositions, Dynamics and Invariance in Personality Structure. *Psychological Review*, 102(2) :246–268.

Mischel, W., Shoda, Y., and Smith, R. (2004). *Introduction to Personality : Towards an Integration*. John Wiley & Sons, 7th edition.

Moffat, D. (1997). Personality parameters and programs. In Trappl, R. and Petta, P., editors, *Creating personalities for synthetic actors*, number 1195 in Lecture Notes In Artificial Intelligence, pages 120–165. Springer Verlag.

Morf, C. C. (2006). Personality Reflected in a Coherent Idiosyncratic Interplay of Intra-and Interpersonal Self-Regulatory Processes. *Journal of Personality*, 74(6) :1527–1556.

Nass, C. and Lee, K. M. (2001). Does computer-synthesized speech manifest personality ? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology : Applied*, 7(3) :171–181.

Nass, C., Moon, Y., Fogg, B., Reeves, B., and Dryer, D. C. (1995). Can computer personalities be human personalities ? *International Journal of Human-Computer Studies*, 43(2) :223–239.

Nass, C., Steuer, J., and Tauber, E. R. (1994). Computers are social actors. In *Proceedings of the ACM Special Interest Group on Computer-Human Interaction conference on Human factors in computing systems (SIGCHI 1994)*, pages 72–78, Boston, MA, USA.

Neff, M., Toothman, N., Bowmani, R., Fox Tree, J., and Walker, M. (2011). Don't Scratch ! Self-adaptors Reflect Emotional Stability. In *Proceedings of the 11th International Conference on Intelligent Virtual Agents (IVA 2011)*, pages 398–411, Reykjavik, Iceland.

Niewiadomski, R., Demeure, V., and Pelachaud, C. (2010). Warmth, competence, believability and virtual agents. In *Proceedings of the 10th International Conference on Intelligent Virtual Agents (IVA 2010)*, pages 272–285, Philadelphia, PA, USA.

Ortony, A., Clore, G. L., and Collins, A. (1990). *The cognitive structure of emotions*. Cambridge university press.

Ouschan, L., Boldero, J. M., Kashima, Y., Wakimoto, R., and Kashima, E. S. (2007). Regulatory Focus Strategies Scale : A measure of individual differences in the endorsement of regulatory strategies. *Asian Journal Of Social Psychology*, 10(4) :243–257.

Paiva, A. (2011). Empathy in Social Agents. *International Journal of Virtual Reality*, 10(1) :1–4.

Park, E. S., Hinsz, V. B., and Nickell, G. S. (2015). Regulatory fit theory at work : Prevention focus' primacy in safe food production. *Journal of Applied Social Psychology*, 45(7) :363–373.

Pereira, A., Prada, R., and Paiva, A. (2012). Socially Present Board Game Opponents. In *Proceedings of the 9th International Conference on Advances in Computer Entertainment (ACE 2012)*, pages 101–116, Kathmandu, Nepal.

Petisca, S., Dias, J., and Paiva, A. (2015). More social and emotional behaviour may lead to poorer perceptions of a social robot. In *Proceedings of the 7th International Conference on Social Robotics (ICSR 2015)*, pages 522–531, Paris, France.

Picard, R. W. (1995). Affective computing. Technical Report 321, MIT Media Lab, Perceptual Computing Group.

Picard, R. W. (1999). Affective computing for hci. In *Proceedings of the 8th International Conference on Human-Computer Interaction (HCI 1999)*, pages 829–833, Munich, Germany.

Picard, R. W. and Klein, J. (2002). Computers that recognise and respond to user emotion : theoretical and practical implications. *Interacting with computers*, 14(2) :141–169.

Plaisant, O., Courtois, R., Reveillere, C., Mendelsohn, G., and John, O. (2010). Validation par analyse factorielle du Big Five Inventory français (BFI-Fr). Analyse convergente avec le NEO-PI-R. *Annales Medico-Psychologiques*, 168(2) :97–106.

Poznanski, M. and Thagard, P. (2005). Changing personalities : towards realistic virtual characters. *Journal of Experimental & Theoretical Artificial Intelligence*, 17(3) :221–241.

Rao, A. S. and Georgeff, M. P. (1995). BDI agents : From theory to practice. In *Proceedings of the First International Conference on Multi-agent Systems (ICMAS 1995)*, pages 312–319, San Fransisco, CA, USA.

Read, S. J., Monroe, B., Brownstein, A., Yang, Y., Chopra, G., and Miller, L. C. (2010). A Neural Network Model of the Structure and Dynamics of Human Personality. *Psychological Review*, 117(1) :61–92.

Revelle, W. (1995). Personality processes. *Annual Review of Psychology*, 46 :295–328.

Revelle, W. and Scherer, K. R. (2009). Personality and emotion. In Sander, D. and Scherer, K. R., editors, *Oxford companion to emotion and the affective sciences*, pages 304–306. Oxford University Press.

Rietzschel, E. F. (2011). Collective regulatory focus predicts specific aspects of team innovation. *Group Processes & Intergroup Relations*, 14(3) :337–345.

Robins, R. W., Caspi, A., and Moffitt, T. E. (2000). Two personalities, one relationship : both partners' personality traits shape the quality of their relationship. *Journal of personality and social psychology*, 79(2) :251.

Romano, D. M., Sheppard, G., Hall, J., Miller, A., and Ma, Z. (2005). Basic : A believable, adaptable socially intelligent character for social presence. In *Proceedings of the 8th Annual International Workshop on Presence (PRESENCE 2005)*, London, UK.

Rothbart, M. K. (2007). Temperament, Development, and Personality. *Current Directions in Psychological Science*, 16(4) :207–212.

Rothmund, T., Baumert, A., and Schmitt, M. (2012). Can Network Models Represent Personality Structure and Processes Better than Trait Models Do ? *European Journal of Personality*, 26(4) :444–445.

Sandercock, J., Padgham, L., and Zambetta, F. (2006). Creating adaptive and individual personalities in many characters without handcrafting behaviors. In *Proceedings of the 6th International Conference on Intelligent Virtual Agents (IVA 2006)*, pages 357–368, Marina Del Rey, CA, USA.

Schaffer, O. (2013). Crafting fun user experiences : A method to facilitate flow. White paper, Human Factors International.

Scherer, K. R. (2000). Psychological models of emotion. *The neuropsychology of emotion*, 137(3) :137–162.

Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. In *Appraisal processes in emotion : Theory, methods, research*, pages 92–120. Oxford University Press, New York and Oxford.

Scherer, K. R. and Sangsue, J. (2004). Le système mental en tant que composant de l'émotion. In *XXVe Journées d'études de l'Association de Psychologie Scientifique de Langue Française (APSLF)*, pages 11–36, Coimbra, Portugal.

Schmitt, T. A. (2011). Current Methodological Considerations in Exploratory and Confirmatory Factor Analysis. *Journal of Psychoeducational Assessment*, 29(4) :304–321.

Scholl, B. J. and Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, 4(8) :299–309.

Sears, D. O. (1983). The person-positivity bias. *Journal of Personality and Social Psychology*, 44(2) :233–250.

Shah, J., Higgins, T., and Friedman, R. S. (1998). Performance incentives and means : how regulatory focus influences goal attainment. *Journal of personality and social psychology*, 74(2) :285–293.

Shermer, M. (2012). *The Believing Brain*. Times Books, New York.

Signoretti, A., Feitosa, A., Campos, A. M., Canuto, A. M., and Fialho, S. V. (2010). Increasing the Efficiency of NPCs Using a Focus of Attention Based on Emotions and Personality. In *Proceedings of the 9th Brazilian Symposium on Games and Digital Entertainment (SBGAMES 2010)*, pages 171–181, Florianópolis, Santa Catarina, Brazil.

Skowronski, J. J. (2012). I, Me, Mine : Variations on the Theme of Selfness. *Social Cognition*, 30(4) :396–414.

Soto, C. J., John, O. P., Gosling, S. D., and Potter, J. (2008). The developmental psychometrics of big five self-reports : Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, 94(4) :718–737.

Sue-Chan, C., Wood, R. E., and Latham, G. P. (2012). Effect of a coach's regulatory focus and an individual's implicit person theory on individual performance. *Journal of Management*, 38(3) :809–835.

Sultan, S. and Bureau, B. (1999). Which optimism in health psychology ? *European review of applied psychology*, 49(1) :43–51.

Summerville, A. and Roese, N. J. (2008). Self-report measures of individual differences in regulatory focus : A cautionary note. *Journal of Research in Personality*, 42(1) :247–254.

Sun, R. (2006). The CLARION cognitive architecture : Extending cognitive modeling to social simulation. In Sun, R., editor, *Cognition and multi-agent interaction*, pages 79–99. Cambridge University Press.

Sun, R. and Wilson, N. (2014). A model of personality should be a cognitive architecture itself. *Cognitive Systems Research*, 29 :1–30.

Sviatlana, D., Busemann, S., and Schommer, C. (2012). Artificial Conversational Companions : A Requirements Analysis. In *Proceedings of the 4th International Conference on Agents and Artificial Intelligence (ICAART 2012)*, pages 282–289, Vilamoura, Portugal.

Tao, J. and Tan, T. (2005). Affective computing : A review. In *Affective computing and intelligent interaction*, pages 981–995. Springer.

Tapus, A., Tapus, C., and Matarić, M. J. (2008). User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics*, 1(2) :169–183.

Tay, B., Jung, Y., and Park, T. (2014). When stereotypes meet robots : The double-edge sword of robot gender and personality in human–robot interaction. *Computers in Human Behavior*, 38 :75–84.

Tiberghien, G. (1999). La psychologie cognitive survivra-t-elle aux sciences cognitives ? *Psychologie française*, 44 :265–283.

Tseng, S. and Fogg, B. (1999). Credibility and computing technology. *Communications of the ACM*, 42(5) :39–44.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236) :433–460.

Uskul, A. K., Sherman, D. K., and Fitzgibbon, J. (2009). The cultural congruency effect : Culture, regulatory focus, and the effectiveness of gain-vs. loss-framed health messages. *Journal of Experimental Social Psychology*, 45(3) :535–541.

Vallieres, E. F. and Vallerand, R. J. (1990). Traduction et validation canadienne-française de l'échelle de l'estime de soi de rosenberg. *International journal of psychology*, 25(2) :305–316.

Van den Bosch, K., Brandenburgh, A., Muller, T., and Heuvelink, A. (2012). Characters with Personality ! In *Proceedings of the 12th International Conference on Intelligent Virtual Agents (IVA 2012)*, pages 426–439, Santa Cruz, CA, USA.

Van-Dijk, D. and Kluger, A. N. (2004). Feedback sign effect on motivation : Is it moderated by regulatory focus ? *Applied Psychology*, 53(1) :113–135.

Van Stekelenburg, J. (2006). *Promoting or preventing social change : Instrumentality, identity, ideology and group-based anger as motives of protest participation.* PhD thesis, Vrije Universiteit, Amsterdam, Netherlands.

Vinciarelli, A. and Mohammadi, G. (2014a). More Personality in Personality Computing. *IEEE Transactions on Affective Computing*, 5(3) :291–300.

Vinciarelli, A. and Mohammadi, G. (2014b). A Survey of Personality Computing. *IEEE Transactions on Affective Computing*, 5(3) :273–291.

Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect : the panas scales. *Journal of personality and social psychology*, 54(6) :1063–1070.

Waytz, A., Morewedge, C., Epley, N., Monteleone, G., Gao, J.-H., and Cacioppo, J. (2010). Making Sense by Making Sentient : Effectance Motivation Increases Anthropomorphism. *Journal of Personality and Social Psychology*, 99(3) :410–435.

Zhao, G. and Pechmann, C. (2007). The impact of regulatory focus on adolescents' response to antismoking advertising campaigns. *Journal of Marketing Research*, 44(4) :671–687.

ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre : Approche computationnelle du regulatory focus pour des agents interactifs : un pas vers une personnalité artificielle

Mots clés : interaction homme-machine ; informatique affective ; personnalité ; regulatory focus ; agents artificiels ; agents animés

Résumé : L'essor de l'informatique affective ouvre aujourd'hui la porte à la création de dispositifs artificiels dotés d'une forme d'intelligence sociale et émotionnelle. L'étude de l'interaction homme-machine dans ce contexte offre de nombreuses pistes de recherche. Parmi celles-ci se trouve la question de la personnalité : comment modéliser certaines caractéristiques de la personnalité pour une entité artificielle ? Comment ces caractéristiques influencent-elles le déroulement de l'interaction entre l'homme et la machine ?

Cette question globale soulève plusieurs questions de recherche : Comment définir la personnalité ? Sur quels modèles et théories issus de la psychologie peut-on s'appuyer pour développer une personnalité artificielle ? Quelle méthodologie adopter pour aborder l'implémentation d'un concept psychologique complexe ? Qu'apporte le développement informatique d'une personnalité à l'interaction homme-machine ? Au domaine de la psychologie de la personnalité ? Comment évaluer expérimentalement ces apports ? Pour aborder ces questions, nos travaux se positionnent dans une optique pluridisciplinaire, au croisement de l'informatique et de la psychologie.

Au regard de sa pertinence pour une approche computationnelle, nous avons modélisé la régulation du Soi comme une composante de la personnalité. Ce concept est approché à partir de la théorie du regulatory focus. Sur cette base théorique, un cadre de travail conceptuel et un modèle computationnel sont proposés. Un questionnaire mesurant le regulatory focus a également été développé et validé.

Ces propositions théoriques sont mises en œuvre dans deux implémentations data-driven (dimensionnelle vs socio-cognitive) dotant des agents de regulatory focus en utilisant de l'apprentissage automatique. Deux études utilisateurs (interaction unique avec agent artificiel vs sessions répétées avec agent animé), présentées dans un cadre ludique, ont permis d'étudier la perception du regulatory focus chez un agent et son impact sur l'interaction.

Nos résultats encouragent l'utilisation du regulatory focus en informatique affective et ouvrent des perspectives sur les liens théoriques et méthodologiques entre informatique et psychologie.

Title : Computational approach of regulatory focus for interactive agents: towards an artificial personality

Keywords : affective computing; human-computer interaction; artificial agents; personality; regulatory focus ; animated agents

Abstract : The development of affective computing is leading to the design of artificial devices endowed with a form of social and emotional intelligence. The study of human-computer interaction in this context offers many research tracks. Among them is the question of personality : how to model some features of personality for a artificial device ? How these features will influence the course of interaction with users ?

This goal rises several research questions: how to define personality? On which models and theories from psychology should we rely to define some artificial personality? Which methodology will help to address the implementation of such a complex psychological concept? What could artificial personality bring to the field of human-computer interaction? And to the psychology of personality? How to experimentally evaluate these contributions? To address these issues, this thesis takes a multidisciplinary approach, at the crossing of computing science and psychology.

Given its relevance to a computational approach, we modeled self-regulation as a component of personality. This concept is approached from the regulatory focus theory. On this theoretical basis, a conceptual framework and a computational model are proposed.

Our theoretical proposals led to two data-driven implementations (dimensional vs. socio-cognitive) which endowed our artificial agents with regulatory focus by using machine-learning. A French questionnaire measuring regulatory focus was designed and validated. Two user studies (brief interaction with artificial agents vs. repeated sessions with animated agents), where the regulatory focus of agents is conveyed via game strategies, enabled the study of regulatory focus perception and its impact on the interaction.

Our results support the use of regulatory focus in affective computing and open perspectives on the theoretical and methodological links between computer science and psychology.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France