



HAL
open science

Modélisation et simulation multi-agent de la formation et de la dynamique d'attitudes basées sur les croyances

Kei-Léo Brousmiche

► **To cite this version:**

Kei-Léo Brousmiche. Modélisation et simulation multi-agent de la formation et de la dynamique d'attitudes basées sur les croyances. Système multi-agents [cs.MA]. Université Pierre et Marie Curie - Paris VI, 2015. Français. NNT : 2015PA066623 . tel-01342445

HAL Id: tel-01342445

<https://theses.hal.science/tel-01342445v1>

Submitted on 6 Jul 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.



**THÈSE DE DOCTORAT DE
L'UNIVERSITÉ PIERRE ET MARIE CURIE**

Spécialité

Informatique

École doctorale Informatique, Télécommunications et Électronique (Paris)

Présentée par

Kei-Léo Brousmiche

Pour obtenir le grade de

DOCTEUR de L'UNIVERSITÉ PIERRE ET MARIE CURIE

Titre de la thèse :

**Modélisation et simulation multi-agent de la formation et de la
dynamique d'attitudes basées sur les croyances**

soutenue le 16 décembre 2015

devant le jury composé de :

M. Frédéric AMBLARD	Examineur
M. Rémy COURDIER	Rapporteur
M. Guillaume DEFFUANT	Rapporteur
M. Jean-Daniel KANT	Encadrant de thèse
M. Nicolas MAUDET	Examineur
M. Nicolas SABOURET	Directeur de thèse

RÉSUMÉ

Nous étudions dans cette thèse la formation et la dynamique des attitudes sociales à l'aide de la simulation multi agent. L'attitude peut se définir comme une évaluation globale d'un objet social sur la base d'informations cognitives ou affectives. Nos travaux s'inscrivent dans le cadre de la simulation sociale qui tente de reproduire informatiquement des phénomènes sociaux complexes à une échelle macroscopique, sur la base de la représentation des individus et de leurs interactions au niveau microscopique. Tandis que les approches existantes dans cette discipline font généralement abstraction des travaux en sciences humaines sur le sujet de l'attitude, nous proposons de suivre une approche psychomimétique en micro-fondant le modèle cognitif de nos agents sur des théories issues de psychologie sociale et des sciences cognitives. Ainsi, nous proposons un modèle de dynamique d'attitude combinant des théories issues de travaux en sciences humaines et sociales de la perception des individus, la communication inter-personnelle et médiatique, la révision de croyances, la réponse émotionnelle ainsi que le sentiment de surprise. Ce modèle a pour objectif de reproduire au niveau microscopique la dynamique des attitudes vis-à-vis d'acteurs qui effectuent des actions observées par la population. Nous avons procédé à une analyse fonctionnelle des différents composants du modèle sur des scénarii abstraits afin d'étudier les capacités de notre modèle, en particulier les phénomènes descriptibles tels que la diffusion de l'information, la résistance à la désinformation ou le processus de conformité. Le modèle a été appliqué dans le contexte des opérations militaires françaises de stabilisation en Afghanistan. L'objectif de cette expérience consiste à reproduire les sondages d'opinions vis-à-vis des Forces en présence, récoltés durant l'intervention, à partir d'un scénario militaire qui a été reconstitué en partenariat avec les officiers en charge des opérations de 2011 à 2012. Les résultats de simulations qui suivent un processus de calibration du modèle affichent une erreur inférieure à 3 points d'écart par rapport aux données réelles. Enfin, nous proposons une analyse microscopique des résultats en appliquant des techniques de classifications automatiques sur les individus afin d'expliquer les différentes tendances d'attitudes au sein de la population.

ABSTRACT

We study in this thesis the problem of social attitude formation and dynamics using multi agent simulation. The concept of attitude could be defined as a global evaluation of a social object, based on cognitive or affective information. Our works belongs to the field of social simulation which aims to reproduce in a virtual environment complex social phenomenon at a macroscopic level based on microscopic representations of individuals and their interactions. While existing approaches in this field rarely consider the results of studies in human sciences on the topic of attitude, we propose to follow a psychomimtic approach by micro-founding the cognitive model of our agents on human and social sciences' theories

on individual's perception, inter-personal and media communication, belief revision, affective responses and the sentiment of unexpectedness. This model aims to reproduce at a microscopic level attitude dynamics toward actors who perpetuate actions witnessed by the individuals. We have proceeded to a functional analysis of the model's various components based on abstract scenarios in order to study the capabilities of our model, and more precisely the describable phenomenon such as information diffusion, resistance to disinformation or the conformity process. The model has been applied in the context of French military operations of stabilisation in Afghanistan. The goal of this experience consists in reproducing opinion polls results of the locals toward the present Forces, collected during the intervention, based on a military scenario which has been recreated in partnership with officers who were in charge of operations between 2011 and 2012. Simulation results that follow a model calibration process show an error below 3 points of disparity compared to the real data. Finally, we propose a microscopic analysis of the results by applying automatic classification techniques on the simulated individuals in order to explain the multiple attitudes tendencies in the population.

PUBLICATIONS

Quelques idées et figures de ce manuscrit sont déjà parues dans les publications suivantes :

Brousmiche, K. L., Kant, J.-D., and Sabouret, N. (2014a). Modélisation de l'impact des croyances et de la communication sur la formation et la dynamique des attitudes : une approche multi-agents. In *JFSMA*.

Brousmiche, K.-L., Kant, J.-D., Sabouret, N., Fournier, S., and Prenot-Guinard, F. (2014b). Modelling the impact of beliefs and communication on attitude dynamics : a cognitive agent-based approach. In *ESSA*.

Brousmiche, K.-L., Kant, J.-D., Sabouret, N., and Prenot-Guinard, F. (2015a). From field data to attitude dynamics. In *ESSA*.

Brousmiche, K.-L., Kant, J.-D., Sabouret, N., Prenot-Guinard, F., and Fournier, S. (2014c). The role of emotions on communication and attitude dynamics : an agent-based approach. In *WCSS*.

Brousmiche, K.-L., Kant, J.-D., Sabouret, N., Prenot-Guinard, F., and Fournier, S. (2015b). Modeling the impact of combat and influence actions on population attitudes towards forces : An agent-based approach. NATO Science and Technology Organization.

REMERCIEMENTS

Je voudrais tout d'abord remercier les membres du jury, Rémy Courdier et Guillaume Deffuant, rapporteurs de cette thèse, pour la pertinence de leurs remarques; Frédéric Amblard et Nicolas Maudet pour l'intérêt qu'ils ont porté à mes travaux. Je remercie également Amal El Fallah Segrouchni pour m'avoir accueilli au sein de l'équipe des Systèmes Multi-Agents du Laboratoire Informatique de Paris 6.

Je tiens à remercier tout particulièrement mon directeur de thèse, Nicolas Sabouret ainsi que Jean-Daniel Kant, encadrant de thèse, qui se sont toujours montrés disponibles tout au long de ce projet. Leurs encouragements et leurs précieux conseils m'ont été d'un très grand profit pour mener à bien cette thèse.

Je souhaite également remercier vivement François Prenot-Guinard et Stéphane Fournier pour leur implication tout au long de ma thèse. Leurs connaissances portant sur les thèmes militaires ainsi que sur l'ingénierie informatique ont beaucoup apportées tant sur les phases de définition de la thèse que pour son déroulement.

J'exprime toute ma gratitude à Jean-Louis Dessalles qui a grandement participé à la définition d'une partie non négligeable du modèle proposé dans cette thèse à travers des séances de travail en commun.

Mes remerciements vont ensuite aux doctorants et anciens collègues de l'équipe, Edouard Amouroux, Costin Caval, Kévin Chapuis, Olivier Goudet, Emmanuel Hadoux, Thomas Hureau et Hazaël Jones, tant pour les moments de procrastinations que pour les débats existentiels.

Enfin, un grand merci à mes parents, mes amis ainsi qu'à ma conjointe Djaïa Dubreuil qui m'ont apporté tout leur soutien.

ACRONYMES

AFD	Agence Française de Développement
BCM	Bounded Confidence Model
CIAE	Centre Interarmé des Actions sur l'Environnement
CICDE	Centre Interarmé de Concepts, de Doctrines et d'Expérimentations
CMA-ES	Covariance Matrix Adaptation Evolution Strategy
COP	Combat OutPost
COBAN	COmmunication of Beliefs using Associative Networks
CSA	Institut d'étude et de conseils
CSO	Collaboration Support Office de l'OTAN
ELM	Elaboration Likelyhood Model
GTIA	Groupement Tactique Inter-Armé
HSM	Heuristic Systematic Model
IED	Improvised Explosive Device
ISAF	International Security Assistance Force
OI	Opération d'Information
OMI	Opération Militaire d'Influence
ONG	Organisation Non Gouvernementale
ONU	Organisations des Nations Unies
OTAN	Organisation du Traité de l'Atlantique Nord
OWA	Ordered Weighted Agregator
SMA	Système Multi Agent
SOM	Self Organising Map
TFLF	Task Force La Fayette
TTP	Tactics, Techniques and Procedures
USAID	Agence des Etats-Unis pour le développement international

TABLE DES MATIÈRES

1	INTRODUCTION	1
2	ATTITUDE : DE LA PSYCHOLOGIE SOCIALE VERS LA SIMULATION SOCIALE	5
2.1	Attitude en psychologie sociale	5
2.1.1	Différentes définitions	5
2.1.2	Fonctions de l'attitude	8
2.1.3	Mesures de l'attitude	9
2.1.4	Composantes de l'attitude	11
2.1.5	Influences cognitives sur l'attitude	13
2.1.6	Influences émotionnelles	20
2.1.7	Attitude en psychologie sociale : bilan	22
2.2	Attitudes en simulation sociale	22
2.2.1	Premiers modèles	23
2.2.2	Des attitudes privées aux opinions publiques	23
2.2.3	Attitudes multi dimensionnelles et exposition sélectionnée	26
2.2.4	Simulation des attitudes politiques et le comportement de vote	28
2.2.5	Théorie de la Simplicité	31
2.3	Synthèse : vers un modèle de la dynamique d'attitudes	33
3	LE MODÈLE POLIAS	37
3.1	Structure de données	37
3.1.1	Protagonistes	37
3.1.2	Objets sociaux	38
3.1.3	Croyances sur des actions	38
3.1.4	Attitudes	40
3.1.5	Messages	40
3.2	Acquisition de croyance	41
3.3	Révision de croyance	41
3.4	Intérêt d'une action	43
3.4.1	L'intensité émotionnelle	44
3.4.2	Surprise	44
3.5	Communication Intra-Population	47
3.5.1	Topologie du Réseau	47
3.5.2	Algorithme de communication	49
3.6	Construction d'une attitude	51
3.6.1	Evaluation du bénéfice d'une action	52
3.6.2	Accessibilité	53
3.6.3	Attitude de Fazio	53
3.6.4	Coresponsabilité	54
3.7	Paramètres de simulation	55
3.8	Conclusions	56
4	ANALYSES FONCTIONNELLES DU MODELE	59
4.1	Impact de l'émotion et de la surprise : étude fonctionnelle	59
4.1.1	Premier protocole	60
4.1.2	Second protocole	62

4.2	Impact des groupes sociaux sur la dynamique des attitudes	64
4.3	Etude fonctionnelle du mécanisme de communication	66
4.3.1	Impact du payoff et de la date de l'information sur sa diffusion	67
4.3.2	Le rôle du seuil de communication T_{com}	72
4.4	Révision de croyance et résistance à la désinformation	74
4.4.1	Première expérience	74
4.4.2	Seconde expérience	77
4.5	Bilan de l'analyse fonctionnelle du modèle	78
5	APPLICATION A UN CAS REEL	81
5.1	Contexte militaire : Opérations de stabilisation	82
5.2	Un cas d'étude : l'intervention française en Afghanistan	83
5.2.1	Reconstitution du scénario militaire	84
5.2.2	Sondages d'opinions	92
5.3	Calibration du modèle et résultats macroscopiques	97
5.3.1	Algorithme de calibration : CMA-ES	97
5.3.2	Protocole de calibration	98
5.3.3	Résultats macroscopiques de calibration	100
5.4	Analyses microscopiques	102
5.5	Application du modèle : Bilan	108
6	CONCLUSION ET PERSPECTIVES	111
6.1	Bilan	111
6.2	Ethique	112
6.3	Perspectives	113
	BIBLIOGRAPHIE	115

TABLE DES FIGURES

FIGURE 1	Exemple d'une représentation de l'attitude chez un fumeur selon le modèle de Fazio	19
FIGURE 2	Exemple d'évolution de l'attitude après l'acquisition d'une nouvelle information évaluative selon le modèle Méta-Cognitif tiré de Petty et collab. (2007)	20
FIGURE 3	Exemple d'une configuration aléatoire d'attitudes (à gauche) et l'équilibre atteint au bout de 18 pas de temps (à droite). (Chaque cellule représente un individu. Les I représentent l'opinion majoritaire; les - indiquent le point de vue opposé) Cet exemple est tiré de Nowak et collab. (1990)	25
FIGURE 4	Evolution de la distribution des attitudes selon différents scénarii. La colonne de gauche représente les cas où les impressions initiales sont distribuées selon une loi normale, en conséquence les attitudes suivent une distribution triangulaire. La colonne de droite représente les cas où les attitudes sont distribuées uniformément. Cette illustration est tirée de Urbig et Malitz (2007)	27
FIGURE 5	Volatilité de 100 électeurs sur 365 jours. Cette illustration est extraite de Kottonau et Pahl-Wostl (2004) - L'encadré 1 correspond à une période d'effort du parti A et 2 celle du parti B	29
FIGURE 6	Exemple de différence entre l'inattendu brut et personnel	32
FIGURE 7	Exemple de différence entre la complexité de génération et de description	33
FIGURE 8	Exemple de configuration d'attitudes inter-groupes sociaux	40
FIGURE 9	Exemples de réseaux sociaux employés dans nos simulations	48
FIGURE 10	Fonction de probabilité de communication	50
FIGURE 11	Exemple d'une attitude dans Fazio, 2007.	51
FIGURE 12	Attitude moyenne de la population en fonction de α et ξ	61
FIGURE 13	Evolution de la moyenne des attitudes dans le scénario perturbé	62
FIGURE 14	Δ en fonction de α et ξ	62
FIGURE 15	Δ en fonction de α avec $\xi = 0.01$	63
FIGURE 16	Evolution de la moyenne d'attitude des groupes sociaux dans le temps selon les trois profils cognitifs	65
FIGURE 17	Awarenness en fonction du temps	68
FIGURE 18	Awarenness en fonction de la surprise temporelle et le payoff de l'action selon trois profils cognitifs	69
FIGURE 19	Evolution temporelle de l'awarenness selon un profil cognitif <i>équilibré</i>	71
FIGURE 20	Awarenness en fonction du seuil de communication T_{com} et le payoff de l'action	73
FIGURE 21	Awarenness en fonction des probabilités de révision selon trois réseaux sociaux	75

FIGURE 22	Vues alternatives des figure 21. Adoption en fonction de π_0 et π_1 avec $k = \{4, 6, 8\}$	77
FIGURE 23	Carte de la Kapisa, une province de l'Afghanistan peuplée d'environ 360 000 habitants selon une densité de 195 habitants/km ² principalement situés dans des villes et villages sur l'axe de Vermont et dans les vallées (voir la figure 24)	85
FIGURE 24	Schématisation de l'évolution des axes d'efforts de la Task Force La Fayette (TFLF) sur la période étudiée au fil des tours d'opérations - Les points bleus représentent les principales villes où des COP/FOB sont installés, les zones bleues désignent les axes d'effort et les lignes marron caractérisent les principales vallées.	89
FIGURE 25	Scénario militaire (actions des Forces Bleu et Rouge en Kapisa entre 2010 et 2011)	90
FIGURE 26	Scénario militaire (actions des Forces Bleu et Rouge en Kapisa entre 2010 et 2011)	95
FIGURE 27	Scénario militaire (actions des Forces Bleu et Rouge en Kapisa entre 2010 et 2011)	96
FIGURE 28	Scénario d'initialisation	99
FIGURE 29	Résultats de calibration	100
FIGURE 30	Carte de Kohonen avec les poids des vecteurs du modèle de classification	103
FIGURE 31	Carte de Kohonen avec un aperçu des moyennes d'attitudes vis-à-vis des Forces	104
FIGURE 32	Mesures sur la classification de Kohonen	104
FIGURE 33	Moyennes et écart-types d'attitudes des classes 2 et 3	105
FIGURE 34	Moyennes et écart-types d'attitudes des classes 8 et 9	107
FIGURE 35	Moyennes et écart-types d'attitudes des classes 2 et 3	108

LISTE DES TABLEAUX

TABLE 1	Structure d'une table de révision de croyances générique.	43
TABLE 2	Table de révision de croyances employée dans nos travaux.	43
TABLE 3	Table de t_s de la figure 19a	72
TABLE 4	Table de révision de croyance employée dans nos travaux.	74
TABLE 5	% de la population ayant répondu favorablement aux questions à différentes dates	94

INTRODUCTION

Les chercheurs en psychologie sociale avancent que l'attitude, qui peut se traduire par le degré de favorabilité d'un individu vis-à-vis d'un objet social, jouerait un rôle majeur dans les mécanismes de la structuration des représentations cognitives, celui du bien-être interne ainsi que du comportement. L'influence sur le comportement, en particulier, a suscité de nombreux travaux sur le concept d'attitude : comprendre la dynamique de l'attitude permettrait d'élucider celle des comportements humains.

L'étude de ce phénomène peut se faire à deux niveaux. Au niveau individuel, où se situent les psychologues investiguant la structure et le fonctionnement du psychisme, nous nous intéressons à décrire, évaluer et expliquer les processus mentaux, en prenant en compte les manifestations de la subjectivité. Au niveau des populations, nous étudions les comportements sociologiques à l'échelle des groupes d'individus. Nous nous inspirons pour cela des théories sociologiques sur les phénomènes macroscopiques résultant notamment des interactions entre les individus. Une des difficultés dans la modélisation des activités humaines consiste à combiner ces deux niveaux de recherches : en effet, l'étude de l'activité humaine, de l'impact de l'attitude, met en œuvre à la fois des mécanismes psychologiques et des phénomènes sociologiques.

Une nouvelle discipline est apparue dans cette perspective : la simulation sociale. Elle vise à reproduire informatiquement des phénomènes sociaux et macroscopiques répertoriés par les sociologues, à l'aide d'une description microscopique des mécanismes et des interactions à un niveau individuel en s'inspirant de la psychologie. Pour ce faire, la simulation sociale utilise des systèmes multi-agents qui permettent de simuler des entités individuelles dotées d'attributs et de mécanismes personnels ainsi que les interactions entre ces entités.

Cependant, parmi les travaux en simulation sociale s'intéressant à la dynamique des attitudes, très peu se basent sur les théories issues de la psychologie. Ceci est lié à la difficulté de modéliser la dynamique des théories psychologiques sur l'attitude qui sont généralement établies de manière statique. Or, une telle approche permettrait d'une part d'accroître l'applicabilité des modèles à des problèmes concrets et, d'autre part, d'encourager un dialogue inter-disciplinaire entre les psychologues sociaux, les sociologues et les informaticiens. Par ailleurs, malgré les nombreuses données de sondages récoltées par les sociologues, très peu de travaux en simulation sociale les mettent à profit : la majorité des simulations porte sur des scénarii et des phénomènes abstraits.

Dans ce contexte, nous proposons au travers de cette thèse un modèle de dynamique d'attitudes conçu selon une nouvelle approche méthodologique. L'enjeu de nos travaux consiste à faire le lien entre les théories issues de la psychologie portant sur les attitudes à un niveau individuel et des phénomènes à l'échelle d'une population observée par les sociologues. Dans ce cadre, nous traiterons deux problématiques majeures :

- Appliquer une approche psychomimétique au problème de la simulation des dynamiques d'attitudes. Cette approche consiste à se référer aux théories

issues de la psychologie, afin d'établir le modèle cognitif des individus simulés. Ainsi, notre modèle d'attitude sera micro-fondé sur les travaux majeurs de cette discipline.

- Valider le modèle sur un cas d'application concret et mesuré. Nous appliquerons notre modèle au contexte des interventions militaires françaises en Afghanistan, suivant un processus de calibration sur des données récoltées sur le terrain.

Nos travaux ont abouti à :

1. Un modèle multi-agent fondé sur des théories psychologiques au niveau individuel telles que l'attitude en tant qu'ensemble de relations entre évaluations et objets sociaux, la révision de croyance, l'évaluation personnelle de l'information, l'impact affectif du stimuli, le sentiment de surprise ou encore la communication inter-personnelle.
2. Une analyse fonctionnelle du modèle qui a permis de reproduire des phénomènes sociaux observés et étudiés par les sociologues tels qu'une certaine tendance de la diffusion de l'information, l'effet des groupes sociaux sur la dynamique d'attitude, la résistance à la désinformation ou encore le phénomène de conformité.
3. Une méthodologie permettant d'appliquer le modèle à des problèmes concrets. Cette méthode consiste à calibrer les paramètres du modèle afin que les résultats de simulation reproduisent des résultats de sondages d'opinion. Deux types de données sont nécessaires : des résultats de sondages ainsi qu'une séquence d'événements correspondants.
4. Une méthode d'analyse microscopique des dynamiques d'attitudes s'appuyant sur une classification des individus à l'aide des cartes de Kohonen dans le but d'expliquer les différentes tendances émergentes dans une simulation.

Ce travail de recherche s'inscrit dans le cadre d'une thèse CIFRE entre l'Université Pierre et Marie Curie (UPMC) et Airbus Defence & Space. Dans ce contexte, nous avons appliqué notre modèle au contexte des opérations militaires françaises en Kapisa entre 2011 et 2012. A partir d'un scénario relatant les opérations militaires, le modèle a reproduit des dynamiques d'attitudes avec un écart inférieur à 3 points par rapport aux résultats enregistrés par des sondages d'opinion sur le terrain.

ORGANISATION DU DOCUMENT Ce manuscrit de thèse s'articule en six chapitres incluant ce chapitre d'introduction. Le second chapitre porte sur les travaux autour du concept de l'attitude. La première partie présente les recherches dans la discipline à l'origine de ce concept, c'est-à-dire la psychologie sociale. La seconde partie décrit en détails quelques travaux marquant en simulation sociale abordant le problème de la dynamique des attitudes. La synthèse de ce chapitre dégagera les principales problématiques que nous traiterons dans le cadre de cette thèse.

Le troisième chapitre présente notre modèle multi-agent sur la formation et la dynamique des attitudes dans une population. Les individus perçoivent les actions dans l'environnement qui constituent leurs bases de connaissances. En fonction de l'interprétation personnelle de leurs croyances, les individus modifient leurs attitudes selon un processus comprenant à la fois une composante rationnelle et

affective. De plus, les informations ainsi acquises peuvent être communiquées à travers un réseau social liant la population.

Le quatrième chapitre expose une analyse fonctionnelle du modèle sur des scénarii abstraits, permettant d'étudier le comportement du modèle, sa sensibilité aux paramètres ainsi que les divers phénomènes reproductibles. Divers phénomènes émergents et les propriétés du modèle y sont discutés.

Le cinquième chapitre porte sur l'application du modèle à un cas concret. La première partie présente le contexte applicatif des menaces asymétriques dans les interventions militaires de stabilisation. La seconde partie décrit les données terrains que nous avons récoltées sur un cas d'étude précis dans ce contexte. La troisième partie expose notre protocole de calibration du modèle sur ces données ainsi que les résultats de simulation.

Enfin, le sixième et dernier chapitre présente le bilan et les perspectives de nos travaux.

ATTITUDE : DE LA PSYCHOLOGIE SOCIALE VERS LA SIMULATION SOCIALE

En suivant l'approche psychomimétique, la première étape dans l'élaboration d'un modèle de dynamique d'attitude consiste à étudier les travaux à l'origine de ce concept afin d'en saisir le sens et d'identifier les avancées qui le fondent. Ainsi nous proposons dans la première partie de ce chapitre (section 2.1) un tour d'horizon des recherches portant sur le concept de l'attitude dans sa discipline mère : la psychologie sociale. Dans un second temps, nous présenterons en détail quelques travaux marquants en simulation sociale qui abordent le sujet de la dynamique des attitudes en appliquant des techniques d'intelligence artificielle 2.2. Enfin, sur la base des constats issus de ces deux études, nous élaborerons une synthèse sur les problématiques de recherches qui guideront la conception de notre modèle multi-agent 2.3.

2.1 ATTITUDE EN PSYCHOLOGIE SOCIALE

Cette section propose un tour d'horizon des recherches en sciences humaines et sociales et plus particulièrement en psychologie sociale relatives à la formation et à la dynamique des attitudes. Dans un premier temps, nous tenterons de définir le concept de l'attitude sur la base des aspects qui ont été mis en évidence dans les différents travaux du domaine (section 2.1.1). Nous étudierons ensuite les fonctions de l'attitude, leurs rôles dans les mécanismes cognitifs et sociaux (section 2.1.2). Nous verrons notamment la manière dont les psychologues mesurent l'attitude (section 2.1.3). Nous montrerons que ces travaux portant sur l'estimation ont permis d'identifier une caractéristique importante de l'attitude : l'accessibilité en mémoire. Enfin, nous terminerons cette première partie du chapitre en décrivant les modèles cognitivistes d'une part et émotionnels d'autre part, sur lesquels nous nous fonderons (sections 2.1.5 et 2.1.6).

2.1.1 *Différentes définitions*

Le terme attitude est communément utilisé dans son sens usuel : une manière de tenir son corps ou une manière d'être qui manifeste un certain sentiment. Dans cette première définition, le terme désigne un comportement observable, une posture exprimant un état d'esprit. En psychologie sociale, une attitude désigne l'état d'esprit lui-même, caché par essence : c'est une variable latente qui va influencer les manifestations observables (le comportement). Comme pour la plupart des constructions théoriques en psychologie, le concept d'attitude a été défini sous différentes formes. Parmi les définitions les plus influentes, nous pouvons citer les suivantes :

- Allport (1935) : « Un état mental de préparation à l'action organisé à travers l'expérience, exerçant une influence directive et dynamique sur le comportement ». La principale notion proposée dans cette première définition est la

tendance à l'action : l'attitude est un des facteurs déterminant le comportement d'un individu. Cette attitude, ou « état mental », serait construite au cours du temps, en fonction du vécu du sujet, à l'opposé de l'instinct (l'attitude n'est pas innée). Enfin, l'influence de cette variable mentale aurait une direction (positive ou négative).

- **Cacioppo et Petty (1982)** : « Une tendance générale et durable, un sentiment positif ou négatif vis-à-vis d'un objet, d'une personne ou d'un sujet ». Cette seconde définition introduit une nouvelle notion importante : un individu a une attitude vis-à-vis d'un objet. La nature de cet objet paraît être relativement distendue, elle peut être une entité quelconque, y compris une personne. Enfin, d'après ces chercheurs, l'attitude serait durable dans le temps.
- **Zanna et Rempel (1988)** : « La catégorisation d'un objet stimulus selon une dimension évaluative ». Nous retrouvons dans cette définition la notion d'objet d'attitude, cette fois caractérisée comme un stimulus qui va induire une réaction chez le sujet. L'idée essentielle proposée ici est que l'attitude intégrerait un mécanisme d'évaluation de l'objet de l'attitude. De plus, ces évaluations serviraient à classer les objets en différentes catégories.
- **Eagly et Chaiken (1993)** : « Une tendance psychologique qui s'exprime par l'évaluation d'une entité particulière avec un degré de favorabilité ou de défavorabilité. » Cette définition paraît être très proche de la précédente. Pourtant, elle diffère sur un point crucial en se libérant du caractère comportemental de l'attitude. En effet, l'attitude est réduite à l'évaluation d'un objet.
- **Fazio (1995)** : « Une association mémoire entre un objet donné et un résumé d'évaluation de l'objet ». Cette dernière définition ajoute un caractère schématique : l'attitude possède une représentation mémoire. Elle serait caractérisée par des associations mémoire entre la représentation de l'objet que se fait le sujet et ses évaluations (*ex.* : « voiture-pollution », « écologique-coûteux »).

En dépit des différences subtiles entre ces définitions, elles affirment toutes que la formation d'une attitude implique un jugement évaluatif de l'objet en question. En d'autres termes, l'attitude dérive d'une décision entre aimer ou ne pas aimer, favoriser ou défavoriser un certain objet, une personne ou une idée.

Dans nos travaux, nous définissons une attitude comme une évaluation globale d'un objet (physique, humain ou conceptuel) sur la base d'informations cognitives ou affectives.

Dans les paragraphes suivants, nous détaillerons les attributs de l'attitude que nous venons d'évoquer : l'objet de l'attitude, sa direction ainsi que sa valeur. Nous définirons également deux concepts majeurs liés à l'attitude : la notion de croyance et l'opinion.

LA NATURE DE L'OBJET DE L'ATTITUDE Nous avons illustré le concept d'attitude en faisant référence à divers objets, sujets de l'attitude. Survient alors la question de la nature de l'objet : est-ce que tout peut être objet d'une attitude ? La réponse semble être relativement flexible (*ex.* : **Zanna et Rempel (1988)**; **Eagly et Chaiken (1993)**; **Fazio (1995)**). Tout concept pouvant être évalué selon une dimension de favorabilité est susceptible d'être la cible d'une attitude. En conséquence, ces concepts peuvent être abstraits (*ex.* : la religion) ou concrets

(*ex.* : un ordinateur). Le soi (*ex.* : estime de soi) ou d'autres personnes (*ex.* : une actrice) peuvent faire l'objet d'attitudes, tout comme une politique sociétale (*ex.* : la peine capitale) ou encore un groupe social (*ex.* : les chercheurs académiques). Dans notre cas d'étude, les objets d'attitudes portent essentiellement sur les acteurs (dans le sens d'effecteurs à l'origine des actions) et les groupes sociaux qui divisent le terrain humain (*ex.* : des ethnies).

DEUX MESURES MAJEURES Sur la base de ces définitions, nous pouvons identifier au moins deux dimensions de l'attitude. Premièrement, les attitudes ont une certaine valence (*ex.* : plaisant, déplaisant) ou plus généralement une direction (positive ou négative). Par exemple, nous pouvons supposer que vous, lecteur de ce manuscrit, manifestez une attitude positive envers la recherche et négative dans l'hypothèse d'un rejet de sa contribution. Deuxièmement, les attitudes ont une certaine intensité, une amplitude. Par exemple, l'un peut détester le foie poêlé, mais un autre peut ne pas l'aimer dans une moindre mesure, au point de se forcer à en manger lors d'un dîner de gala. Nous verrons dans les prochaines sections que ces deux mesures jouent un rôle important sur l'influence de l'attitude dans le traitement de l'information et le comportement.

CROYANCES Une autre notion sous-jacente à ces définitions est le concept de croyance. Il ne s'agit pas dans ce contexte d'une croyance religieuse ou spirituelle. En psychologie, une croyance fait référence à l'acceptation d'une information factuelle (Stevenson (2010)), par exemple « Je crois que la Terre tourne autour du soleil. » ou « Je crois que la Force étrangère a fait une patrouille hier. ». Ces croyances constituent la connaissance (ou « l'expérience » en reprenant la définition d'Allport (1935)) des individus qui peuvent porter sur toutes informations interprétables par le sujet. Notamment, les évaluations d'un objet auxquelles font référence les définitions de l'attitude peuvent également être représentées sous forme de croyances (*ex.* : « Je crois que la Force étrangère m'est bénéfique. »). Bien entendu, ces évaluations peuvent elles-mêmes se construire sur la base des croyances factuelles liées à la Force (*ex.* : « Car elle fait des patrouilles »).

Ainsi, l'attitude est fondée en grande partie sur les croyances de l'individu (Fishbein et Ajzen (1975); Fazio (1995)). Nous reviendrons sur cette vision cognitive de l'attitude dans la section 2.1.5.

OPINION ET ATTITUDE Le caractère implicite de l'attitude implique au moins deux niveaux : le niveau génotypique sur lequel elle se forme et se transforme ; le niveau phénotypique où elle s'exprime de manière observable (Beauvois et Deschamps (1996)). L'opinion se situe au second niveau et permet d'inférer les attitudes qui lui sont sous-jacentes : elles représentent une des formes d'expression de l'attitude. En d'autres termes, la quantification de l'opinion ne concerne épistémologiquement que la partie verbale de l'expression du sujet et n'est en aucun cas la réflexion transparente de l'attitude. Dans nos recherches, nous nous intéressons au niveau inférieur, aux attitudes. Cependant, malgré cette distinction claire entre ces deux concepts, nous sommes contraints de confronter nos sorties de simulation à des résultats de sondages d'opinion lors de la phase d'évaluation (voir section 2.1.3).

2.1.2 Fonctions de l'attitude

Nous avons vu que les attitudes peuvent se manifester par des opinions ou des comportements. Dans cette section, nous présentons les fonctions concrètes de l'attitude proposées principalement par [Katz \(1960\)](#) et [Smith \(1956\)](#).

FONCTION HEURISTIQUE Elle correspond au rôle de l'attitude lors de décisions simples, que ce soit pour un comportement ou l'évaluation d'une information. L'attitude assigne un résumé évaluatif du pour et du contre relatif à un objet social. Ce résumé va guider et faciliter la prise de décision. Par exemple, l'attitude peut servir à se rapprocher des éléments bénéfiques et se prémunir de risques potentiels (comme l'ont montré [Maio et Haddock \(2004\)](#)). Selon [McGuire \(1968\)](#), elles sont « un guide simplifié et pratique du comportement approprié à l'égard d'objets spécifiques ». Autre exemple, chez [Fishbein et Ajzen \(1975\)](#), l'attitude fait partie des critères de décision préparant l'action.

AJUSTEMENT SOCIAL L'attitude sert également à se rapprocher des personnes que nous aimons et s'éloigner de celles que nous n'apprécions pas (*ex.* : [Pratkanis et Greenwald \(1989\)](#)). La notion de distance ne se limite pas à la proximité géographique mais elle est aussi cognitive : par exemple une personne peut être amenée à boire un soda vanté commercialement par son chanteur préféré. Simultanément, cet acte va peut-être le rapprocher de certaines autres personnes susceptibles de partager ses préférences. L'attitude va jouer alors un rôle de sociabilisation.

FONCTION SCHÉMATIQUE L'attitude structure la mémoire selon un schéma cognitif. Selon [Pratkanis et Greenwald \(1989\)](#), « Un schéma consiste en une organisation d'un sous-ensemble de savoirs relevant d'un domaine limité. Il s'agit à la fois de contenus (organisation de l'information) et de procédures (usage de l'information dans la connaissance). Par contraste avec une heuristique, qui est une simple règle, un schéma est une organisation d'un ensemble de procédures et de données particulières à un domaine ». En effet, de nombreux chercheurs (*ex.* : [Pratkanis et Greenwald \(1989\)](#); [Fazio \(1995\)](#); [Petty et Cacioppo \(1986\)](#)) voient dans l'attitude une organisation schématique des informations qui lui sont associées. Or, le schéma facilite l'encodage et l'accès à l'information en mémoire : il aide à la compréhension des messages, la reconstruction lors du rappel et l'apprentissage. Par conséquent, nous pensons qu'une représentation computationnelle de l'attitude ne doit pas se limiter à une (ou plusieurs) variable(s) mais doit présenter un caractère schématique, au moins dans son processus de formation.

EXTERNALISATION OU MAINTENANCE DE SOI L'attitude peut servir à résoudre des conflits internes allant à l'encontre de l'estime de soi par le processus d'externalisation. Par exemple, une personne maladroit au billard aura tendance à baisser son attitude envers ce jeu afin de préserver son estime de soi de sa mauvaise performance (*ex.* : [Bem \(1973\)](#); [Snyder \(1974\)](#); [Steele \(1988\)](#)).

Dans nos travaux, les attitudes accomplissent à la fois un rôle heuristique lors de l'estimation des événements ainsi qu'une fonction schématique pour l'organisation de la connaissance. De plus, notre objectif *in fine* consiste à reproduire

l'évolution des attitudes au fil des événements tel que nous pouvons observer dans la réalité. A cette fin, nous devons être en mesure de mesurer les attitudes, aussi bien réelles que simulées.

2.1.3 Mesures de l'attitude

Comme nous l'avons évoqué, une attitude n'est pas directement observable de part sa nature latente, comme pour la plupart des constructions théoriques en psychologie. Cependant, il est possible de l'inférer à partir de ses manifestations (Olson et Fazio (2003)). Les psychologues sociaux ont ainsi élaboré de nombreuses techniques permettant une estimation des attitudes. Il est important de présenter le caractère mesurable de l'attitude. D'une part pour mieux comprendre les recherches sur cette thématique et, d'autre part, afin de permettre une évaluation des résultats de nos simulations. En effet, ces évaluations constitueront un des apports majeurs de notre contribution.

Dans cette section, nous présentons succinctement les deux grandes catégories de mesure d'attitude : explicite (directe) et implicite (indirecte). Cette dichotomie a fait l'objet d'une longue histoire en psychologie. Elle oppose les processus qui requièrent une attention consciente du sujet, à l'opposé de ceux qui traitent un état inconscient (ex. : apprentissage, mémoire, anticipation d'un comportement). Appliquée à la mesure de l'attitude, elle distingue simplement les cas où l'on demande directement l'attitude au sujet et les processus qui n'attendent pas de réponses verbales ou écrites.

2.1.3.1 Mesures explicites

Les mesures explicites se résument souvent à demander directement aux participants leurs opinions vis-à-vis d'un objet social ou d'un sujet (ex. : « Quelle est votre opinion sur l'usage de robots pour les courses de dromadaires ? »). Se pose alors la question de l'échelle de mesure de l'attitude. Les techniques de construction d'échelle d'attitude que nous présentons ci-après sont employées encore aujourd'hui, bien qu'elles aient été élaborées entre 1929 et 1944. Très peu d'alternatives ont été proposées depuis, probablement du fait que le problème est aujourd'hui moins central, l'intérêt s'étant déplacé vers les méthodes multidimensionnelles. Nous présentons ici les modèles qui ont fondé les méthodes actuellement employées.

THURSTONE Thurstone et Chave (1929) ont proposé de transposer aux attitudes les méthodes de mesure qu'ils avaient élaborées en psychométrie, puis en psychophysique. Le principe consiste à répéter les questions et à mesurer la fréquence des réponses en fonction d'au moins deux stimuli A et B, sachant que les réponses à un stimulus ne sont pas constantes, car imprécises. Par exemple, si la réponse au stimulus A est positive dans 80% des réponses et dans 60% des cas pour le stimulus B, on peut en déduire que l'attitude envers A est supérieure à B. Maintenant, en ajoutant un certain « degré de favorabilité » sur les réponses, il est désormais possible de mesurer une distance entre les stimuli : si la probabilité $A > B$ est de 0,7 avec un degré moyen de 5 et $A > C$ de 0,8 avec un degré moyen de 4, on peut en conclure que $0.7 * 5 > 0.8 * 4$ et $B > C$. Cependant, cette méthode postule l'homogénéité totale des juges (les participants) : d'une part, à

chaque question le participant ne doit pas se souvenir de sa réponse précédente, exercice difficile dans le cas où il est conscient qu'on lui demande son opinion sur un sujet précis (variation intra-personnelle); d'autre part ils doivent tous avoir interprété le stimulus de la même manière (variations inter-personnelles). Or, chacun possède sa propre représentation d'un objet social.

LIKERT Likert (1932) propose à l'inverse une méthode mettant la cohérence au centre du problème de mesure. Son approche se base sur le principe d'une étude de corrélation par une analyse factorielle. Si l'on considère un énoncé modérément favorable à l'objet social en question, plus le participant sera défavorable, plus la probabilité qu'il soit en désaccord augmentera. Inversement, plus il sera favorable, plus il sera probable qu'il soit en accord. Le principe consiste à présenter une série d'énoncés qui seront plus ou moins favorables et d'étudier la courbe d'accord / de désaccord des participants selon une échelle de 5, 7 ou 9 valeurs qualitatives bipolaires. Si cette courbe est monotone, elle sera qualifiée de « caractéristique opératoire ». Si cette courbe ne montre aucune corrélation avec l'attitude, les réponses ne pourront pas être considérées comme une bonne expression de l'attitude, en conséquence elles ne pourront pas être évaluées. Dans le cas où les réponses sont mesurables, des poids sont attribués à chacune d'entre elles. Ces poids sont proportionnels au nombre d'occurrences de la réponse dans l'échantillon total. Plus une réponse est rare, plus son poids est important. La somme des réponses pondérées donne la valeur d'attitude inférée.

LIMITES Bien que ces mesures directes en recherches empiriques sur la psychologie de l'attitude aient été largement adoptées, ces techniques présentent néanmoins certaines limites. Par exemple, les participants ne sont pas toujours conscients de leur propre attitude sous-jacente vis-à-vis de certains objets (*ex.* : le racisme, les armes à feu), comme l'a montré Greenwald (1977). D'autres chercheurs ont en outre constaté qu'une subtile différence dans la formulation d'un énoncé pouvait entraîner un impact direct sur l'attitude mesurée Schwartz et Baer (1991). L'une des plus importantes critiques de la mesure directe porte sur la *gestion de l'impression* : le participant a tendance à adapter ses réponses dans l'intention de se présenter sous le meilleur aspect auprès de l'expérimentateur, comme l'a montré Paulhus (1998).

2.1.3.2 Mesures implicites

Afin de contourner les problèmes que présentent les mesures explicites, les psychologues ont développé de nombreuses stratégies d'estimation implicites ou indirectes.

EVALUATION DES ADJECTIFS Comme nous l'avons présenté au début de ce chapitre, Fazio (1995) définit l'attitude comme étant un ensemble d'associations mémoire entre l'objet social et ses évaluations. De plus, chacune de ces évaluations est caractérisée par une valeur d'accessibilité déterminant le degré d'accès du sujet à l'information. Le processus d'évaluation des adjectifs (« evaluative priming » en anglais) consiste à mesurer le temps de latence entre la perception du stimuli, qui n'est pas directement l'objet en question mais un de ses traits ou « adjectifs », et la réponse du participant. Par exemple, « blouse

blanche », « hôpital », « ordonnance » *etc.* pour les médecins. Afin de mesurer cet écart temporel, l'expérience est effectuée sur un ordinateur : l'écran affiche les stimuli (les adjectifs), le participant doit répondre en appuyant soit sur le bouton « favorable » ou « défavorable ». L'idée sous-jacente est qu'une information très accessible met moins de temps à être remémorée et vice-versa (Fazio et collab. (1995)). Ce test permet alors de mesurer la différence du signe et de la valence de deux attitudes : le signe est donné par la touche actionnée et la valence par la latence.

Cette procédure peut être affinée. Au lieu de répondre si l'on est favorable ou pas au mot affiché, le participant est dans un premier temps stimulé par l'objet de l'attitude, puis doit dans un second temps classer des adjectifs, selon qu'ils soient « bon » ou « mauvais ». Le participant ayant activé la représentation de l'objet rencontrera des difficultés pour les classer sans commettre d'erreurs. Par exemple, après avoir activé l'objet « famine », le sujet mettra beaucoup plus de temps à classer l'adjectif « bonheur » dans la catégorie « bon », au pire des cas il peut aussi se tromper en le classant dans la catégorie « mauvais ». Le taux d'erreurs pondéré par les temps de latence donnera alors le score de l'attitude.

TEST D'ASSOCIATION IMPLICITE Greenwald et collab. (1998) ont proposé une méthode de mesure implicite s'inspirant de l'évaluation des adjectifs. A l'instar de la méthode précédente, le participant doit classer non seulement des adjectifs mais également des objets d'attitudes. Ce test a la particularité de mesurer différentes attitudes sur plusieurs séquences d'adjectifs. En comparant les résultats sur chacun des objets, il permet de mesurer l'attitude de manière sensiblement plus précise que dans la méthode exposée précédemment.

2.1.3.3 *Les autres types de mesures*

D'autres types de mesure, s'appuyant sur d'autres leviers, ont été développés. Dans les années 60 - 70, les chercheurs ont expérimenté de nombreuses mesures psychophysiques telles que la conductance cutanée, la dilatation de la pupille ou encore l'électromyographie faciale. Cependant, le rapport entre la précision et le coût de mise en œuvre a freiné leur popularité. Plus récemment, les recherches sur les techniques de mesure de l'activité électrique cérébrale semblent ouvrir de nouveaux horizons. Quoiqu'il en soit, toutes ces techniques physiologiques ne permettent pas de mesures s'appliquant à de larges échantillons.

Dans cette section, nous avons montré que l'obtention des mesures d'attitudes représente une tâche particulièrement complexe. Nos travaux ne portent pas sur la mesure de l'attitude en soi. Cependant, cette notion reste importante pour nous, dans le sens où l'évaluation du modèle et de ses sorties se basera sur des mesures d'opinions ainsi que ses composantes.

2.1.4 *Composantes de l'attitude*

Rosenberg et Hovland (1960) proposent que l'attitude soit constituée de trois composantes en relation les unes aux autres : une composante affective, comportementale et cognitive. La composante affective exprime la réponse émotionnelle, le sentiment du sujet lorsqu'il est confronté physiquement à l'objet ou lors de

l'activation de sa représentation (l'objet est évoqué). La composante comportementale fait référence aux actions et intentions du sujet vis-à-vis de l'objet. Enfin, la composante cognitive regroupe toutes les croyances relatives à l'objet. Par exemple, Mehdi assiste à une projection d'un film en plein air, organisée par l'OTAN. Le film ainsi que l'ambiance conviviale lui plaisent (composante affective). Il a notamment pu croiser ses amis du quartier avec qui il s'est réuni sur la place principale, autour d'un thé. « La prochaine fois, je devrais inviter Asma, ça devrait également lui plaire », se dit-il (composante comportementale). Par ailleurs, il est conscient que ces actes sociaux ont et auront un impact positif sur la dynamique économique du village (composante cognitive). Ce modèle ne contraint aucun ordre chronologique ou hiérarchique selon lequel se structure l'attitude.

Breckler (1984) a prouvé l'existence et la distinction entre ces trois composantes, sans que cela implique qu'elles soient indépendantes entre elles. Dans son expérience, les sujets devaient répondre à un questionnaire décliné en trois parties portant sur l'affectif, le comportement et la cognition vis-à-vis des serpents. (i) les serpents sont-ils gentils ou cruels (cognition), (ii) les serpents vous rendent-ils anxieux ou content, (iii) voulez-vous les toucher (comportement). L'étude conclut à un faible taux de corrélation entre les scores de chacune des composantes. Par conséquent, les trois composantes sont empiriquement distinctes.

COMPOSANTE COMPORTEMENTALE EXTERNALISÉE Bien que le modèle de Rosenberg et Hovland ait été largement accepté par la communauté scientifique, certains ont soulevé l'ambiguïté de la composante comportementale. En effet, l'observation d'un comportement en adéquation avec l'attitude n'est pas systématique. De ce fait, l'aspect comportemental constitue souvent un concept externe à l'attitude et non pas une composante interne qui la constituerait (*ex.* : Fishbein et Ajzen (1975); Fazio (1995)). Les recherches sur le lien entre attitude et comportement qui perdurent depuis plus d'un demi siècle confirment ce constat (*ex.* : Simon (1955); Wicker (1969); Ajzen et Fishbein (1977); Deci et Ryan (1980); Ajzen (1991); Kraus (1995); Finlay et collab. (1997); Fishbein et Ajzen (2005)). Il existe cependant une ambiguïté car la formulation originelle de Rosenberg de la composante affective incluait aussi bien les comportements futurs (intentions) que les comportements antérieurs (*ex.* : « Les dernières fois que je suis allé au cinéma en plein air, ça c'est mal passé »). Le fait est que dans la majorité des travaux sur l'attitude, cette composante n'est pas prise en considération.

RELATION ENTRE LES COMPOSANTES ET L'ATTITUDE GLOBALE Les composantes permettent-elles de prédire l'attitude globale et, si c'est le cas, dans quelle mesure? Cette question, légitime et nécessaire pour la suite de nos travaux, a été étudiée par de nombreux chercheurs en psychologie. Notons tout d'abord que, d'une part du fait de la popularité des mesures des composantes affectives et cognitives, et d'autre part pour les raisons que nous avons évoquées précédemment, la majorité de ces études ne prennent pas en compte la composante comportementale.

L'un des premiers travaux d'enquête concernant l'importance de la cognition et de l'affectif sur l'attitude a été mené par Abelson et collab. (1982) en étudiant la prédictibilité des attitudes vis-à-vis des élections présidentielles américaines. Dans cette étude, les sujets étaient amenés à décrire les traits personnels de

chaque candidat (évaluations cognitives), leurs sentiments (affectifs) envers eux et, dans un second temps, leurs attitudes. Abelson et ses collègues ont pu identifier une importante corrélation entre la favorabilité des réponses affectives et l'attitude finale. Une corrélation entre les croyances et l'attitude apparaissait également, mais dans une moindre mesure. D'autres travaux (*ex.* : Eagly et collab. (1994); Haddock et Zanna (1997)) ont par la suite abouti à des résultats similaires, confirmant l'hypothèse que l'information affective et cognitive permet de prévoir l'attitude dans le domaine politique.

Dans le cas des attitudes inter-groupes, Esses et collab. (1993) ont dirigé une série d'études sur l'importance des informations cognitives et affectives sur les attitudes discriminatoires. La méthode employée est proche des travaux d'Abelson : les sujets décrivent des groupes sociaux à l'aide d'informations cognitives et affectives avant de donner leurs attitudes envers le groupe. A l'instar des résultats précédents, les deux composantes permettaient de prédire l'attitude. En outre, ces études ont permis de constater que la relativité de la contribution de chacune des composantes à l'attitude (leurs poids dans la construction de l'attitude) dépend du signe de l'attitude : dans le cas où l'attitude envers le groupe est négative, la composante cognitive est plus prédictive que la composante affective ; dans le cas contraire (attitude positive) la composante affective prédomine.

D'après ces études il semble nécessaire, dans nos travaux, que l'attitude prenne en compte à la fois (1) des informations cognitives et (2) des informations affectives.

Ainsi, (1) nous amène à étudier les influences cognitives sur l'attitude (dans la prochaine section) et (2) les influences affectives (dans la section 2.1.6).

2.1.5 Influences cognitives sur l'attitude

Le processus de la formation et de la transformation de l'attitude est un processus complexe qui fait appel à divers mécanismes et à des informations de natures variées (*ex.* : Hovland et collab. (1953); Kelman (1961); Chaiken (1980)), comme nous l'illustrons dans l'exemple suivant.

Vous rencontrez une personne pour la première fois, par exemple lors d'une réunion de travail. Dès les premières secondes, vous vous forgez une impression de cette personne. Une impression qui peut se baser sur son apparence physique ou vestimentaire, sa manière de se tenir, de s'exprimer et/ou le fait qu'il ait été introduit par un collègue que vous appréciez, etc. Au cours de la réunion, l'acquisition de plus amples informations sur cette personne via ses prises de paroles vont probablement modifier cette première impression. Ces intuitions et informations peuvent ainsi se combiner et former une attitude relativement stable dans le temps vis-à-vis de cette personne. Cette attitude vous reviendra en mémoire lorsque vous croiserez de nouveau ce personnage.

Nous avons déjà mentionné que les trois (ou deux) composantes de l'attitude ont une influence sur l'attitude. Dans cette section, nous présentons les principaux travaux en psychologie sociale décrivant les processus d'influences cognitives. La section suivante décrira les influences affectives sur l'attitude (cf. section 2.1.6).

De très nombreuses recherches ont été menées sur l'effet des informations cognitives sur l'attitude. Cette approche cognitiviste a été si importante que les modèles majeurs sur le changement de l'attitude en général se sont focalisés sur le rôle de la cognition dans la persuasion. La persuasion et sa thématique (processus par lequel l'attitude du sujet est modifiée intentionnellement) existent depuis les temps les plus anciens : les rhéteurs et sophistes du V^e siècle avant J-C. avaient déjà développé des techniques oratoires de persuasion. Pourtant, il aura fallu attendre la Seconde Guerre mondiale pour que la persuasion fasse l'objet d'investigations scientifiques. Les modèles scientifiques de persuasion peuvent être scindés en deux générations qui se sont historiquement succédées : l'étude des types d'informations ayant une influence sur l'attitude et les modèles décrivant le processus d'influence.

2.1.5.1 *Les types d'informations*

MODÈLE DE YALE Les premières recherches scientifiques sur la persuasion furent initiées par Carl Hovland (*ex.* : Hovland et collab. (1949, 1953)) dans l'armée américaine, l'objectif étant de maintenir le moral des soldats. Ces recherches se poursuivirent dans l'Université de Yale en diversifiant les applications pratiques, comme par exemple dans le secteur de la publicité. La majorité de ces recherches portaient sur l'efficacité d'un ensemble d'attributs caractéristiques de la situation sur la persuasion : « *Qui dit quoi sur quel canal qui avec quel effet ?* » (Lasswell (1948)). Le « *qui* » concerne les attributs de la source de la communication influençant la persuasion. Notre cas d'étude s'intéresse au processus de la dynamique des attitudes notamment en fonction des communications des acteurs vers la population et des échanges de messages inter-personnels. La question du *qui* joue alors un rôle majeur dans nos travaux : elle concerne les attributs des acteurs et des individus qui émettent un message. Le « *quoi* » porte sur la structuration, la forme du message. Mise en rapport avec notre cas d'étude cette notion relève d'une granularité trop ténue (nous ne pouvons pas représenter différentes structures de messages plus ou moins argumentatives), nous ne dévierons pas sur les détails de ces recherches. La question de « *qui* » porte principalement sur la mise en garde des sujets, c-à-d. le mécanisme de résistance contre le processus de persuasion lorsque les individus ont conscience de l'intention de persuasion de leur interlocuteur (*ex.* : Greenwald (1968)). Ces mécanismes étant difficilement observables à l'échelle d'une population, nous n'en ferons pas un objet de nos travaux. En conséquence, au cours des prochains paragraphes, nous nous bornerons à décrire les caractéristiques de la source.

NATURE ET CARACTÉRISTIQUES DE LA SOURCE La source d'un message peut être un expert du thème relaté (*ex.* : un dentiste pour une publicité de dentifrice), une personne physiquement attractive (*ex.* : une pin-up), une personnalité digne de confiance (*ex.* : un maire, un chef de tribu), une personne exerçant un haut statut social (*ex.* : un patron). Mais cette source peut tout aussi bien être dénuée de l'une ou de toutes ces qualités. La source ne se limite pas forcément à une seule personne, elle peut provenir d'un groupe (*ex.* : des manifestants), une agence (*ex.* : l'Agence Nationale de la Sécurité), une organisation (*ex.* : ONG) ou même provenant d'inconnus. On peut identifier principalement deux carac-

téristiques de la source susceptible d'exercer un effet sur son efficacité dans la persuasion : la crédibilité et l'attractivité.

LA CRÉDIBILITÉ fait référence à un ensemble de jugements complexes portés sur la source du message, concernant son niveau de confiance, sa capacité apparente et son objectivité (Hovland et collab. (1953)). Diverses expériences ont démontré l'impact de la crédibilité sur la persuasion. Un exemple, Hewgill et Miller (1965) ont manipulé le degré de crédibilité perçu en faisant varier le statut annoncé de l'auteur d'une communication sur l'utilité des abris atomiques pendant la guerre froide. L'auteur supposé était soit un scientifique (source très crédible), soit un étudiant (peu crédible). Les conclusions de ces expérimentations convergent toutes : une source est d'autant plus crédible que le niveau d'éducation, « d'intelligence », du statut social et de la compétence professionnelle qui lui étaient attribués augmentait (Hass (1981)). En outre, Chaiken (1980) et d'autres chercheurs ont montré que selon le taux d'implication du sujet, la communication n'était traitée qu'en se basant sur la crédibilité, sans tenir compte du contenu argumentatif.

Les attributs statiques de la source ne sont pas les seuls facteurs qui influencent le degré de crédibilité : des études ont démontré que la position défendue par la source était notamment un critère important dans le processus de la persuasion (ex. : Kelley (1973)). En général, on s'attend à ce qu'un auteur défende une opinion allant dans le sens de ses propres intérêts. Or, si elle s'avère en sa défaveur, il est possible que nous jugions la source comme étant plus crédible ou sincère (à moins que ses facultés cognitives ne le défavorisent abusivement). Les théories d'attributions causales ont mis en évidence que la perception de la crédibilité de la source pouvait être le résultat d'inférences subjectives complexes (Kelley (1973)). D'autres chercheurs ont également rejoint la qualification de ces inférences causales comme étant déterminantes dans l'estimation de la confiance et de la crédibilité des sources (ex. : Eagly et collab. (1981)).

Dans notre cas d'étude, nous pouvons par exemple imaginer qu'un acteur diffusant un message en sa faveur soit moins crédible qu'un membre de ma famille qui me relate un fait dont il a été témoin. Il nous paraît important de pouvoir décrire de tels phénomènes.

L'ATTRACTIVITÉ de la source détient un rôle complémentaire à la crédibilité. Selon Kelman (1961), l'attitude peut être modifiée en fonction de deux processus : la source est sympathique, l'attitude est influencée par un processus d'*identification*. Lorsque la source est crédible, elle est influencée par un processus d'*internalisation*. Le changement qui résulte de l'*identification* n'est pas fortement intégré dans le système de croyances et de valeurs de l'individu : les arguments cognitifs occupent un rôle secondaire. En contraste, dans le processus d'*internalisation*, la persuasion dépend de la pertinence des informations et de l'argumentation. Plus une source est experte (crédible), plus les informations de la communication sont évaluées. Plus la source est sympathique, moins les informations sont traitées par le sujet. La persuasion dépend alors de la durée du lien affectif reliant la source au sujet (Norman (1975)).

L'attractivité de la source semble donc être un attribut important. Cependant, dans notre contexte applicatif, il est difficile de récolter des données sur cette notion (ex. : attractivité d'une Force armée).

2.1.5.2 Réponses cognitives

Nous avons compris que la persuasion peut dépendre d'un processus cognitif analysant les informations relatives au sujet de l'attitude. Cependant, nous ne disposons d'aucune précision sur la manière dont l'acceptation et l'évaluation de ces informations vont modeler l'attitude.

Sur ce point, Greenwald (1968) a émis l'hypothèse que l'étude de la persuasion doit intégrer les réponses cognitives du sujet (les pensées remémorées) succédant à la réception du message. De nombreuses expérimentations ont été effectuées afin de vérifier cette hypothèse. Le principe consistait à récolter les réponses cognitives du sujet après lui avoir exposé un message persuasif (*ex.* : « Votez untel. » ; réponses cognitives : « c'est un bon orateur », « mais sa politique est trop libérale », etc.). Elles aboutirent toutes à un même résultat : plus les réponses cognitives sont positives, plus les sujets sont favorables au texte (*ex.* : Brock (1967); Petty et collab. (1976)).

Sur la base de ce constat, Fishbein et Ajzen (1981) ont proposé le premier modèle majeur d'impact des réponses cognitives sur la formation de l'attitude. Leur modèle se fonde sur l'hypothèse que les croyances (au sens défini dans le §2.1.1) de l'individu sont à la base des attitudes. C'est à dire qu'un message persuasif modifie l'attitude lorsqu'il modifie les croyances sous-jacentes de l'attitude et/ou leurs évaluations. Néanmoins, Fishbein et Ajzen notent que toutes les croyances ne se valent pas : certaines sont plus importantes que d'autres dans la détermination de l'attitude. Selon eux, les croyances *modalement saillantes* ou *primaires* auraient le plus grand impact. En d'autres termes, ce sont celles qui ont le plus de chance de venir à l'esprit. Ce modèle est communément connu sous la dénomination « framework d'anticipation-valeur ».

Deux principaux canaux sont prévus dans ce modèle : soit l'attitude est modifiée par le changement de l'évaluation d'une croyance, soit par l'ajout d'une croyance. Par exemple, un client souhaite se procurer une voiture chez un concessionnaire. Son attention est attirée par un véhicule monospace mais son prix lui paraît excessif. Il est difficile de modifier cette évaluation du prix en soi. Mais le commercial parvient à le convaincre que la 4x4 familiale qui l'intéressait initialement était bien plus onéreuse, relativisant ainsi le prix du monospace. Le vendeur surenchérit, il énumère la liste des caractéristiques avantageuses de la voiture (sa climatisation en série, son système de sonorisation, etc.). La relativisation du prix et l'addition des nouvelles croyances sur le véhicule auront engendré une nouvelle attitude chez le client qui finira par l'acheter.

D'après ce modèle, il nous paraît important que toutes représentations d'attitude fondées sur des croyances soient munies d'un mécanisme de révision de croyance. Ce mécanisme doit permettre à minima (1) l'ajout de nouvelles croyances et (2) modifier les caractéristiques des croyances existantes (*ex.* : le prix d'un véhicule).

2.1.5.3 Dynamiques d'influences de l'information sur l'attitude

Dans les précédentes sections, nous avons évoqué qu'il existe généralement deux types d'informations influençant la formation de l'attitude : les caractéristiques de la source et les réponses cognitives. Nous avons présenté dans le dernier paragraphe un des premiers modèles décrivant la dynamique de l'influence des réponses cognitives sur l'attitude. Des chercheurs ont tenté de compléter ce modèle

en y greffant un mécanisme d'influence des caractéristiques de la source. Nous verrons cependant que cette vision modulaire a été abandonnée au profit de modèles proposant un processus unifié pour les deux types d'informations. Dans cette section, nous présentons la génération de modèles dual ainsi que la transition vers les modèles unifiés. Enfin, nous exposerons les formulations récentes de l'attitude sur lesquelles notre modèle est fondé.

DES MODÈLES DICHOTOMIQUES AUX MODÈLES UNIFIÉS La Théorie du Jugement Social (Sherif et Hovland) propose un mécanisme d'adaptation de l'attitude en fonction de l'acceptation d'une information qui intègre les impacts des différentes caractéristiques de la source discutés dans le paragraphe 2.1.5.1. D'après cette théorie, le sujet aurait un intervalle d'attitude centré sur sa position initiale de référence. Cet intervalle serait de plus composé de trois compartiments : le compartiment d'acceptation, celui du rejet et le compartiment neutre. La taille de ces compartiments sont affectées par le degré d'implication du sujet par rapport au message. Selon le compartiment dans lequel tombe le message persuasif, le sujet accepte, rejette ou ignore l'information. L'attitude de référence peut alors converger ou diverger de la position avancée par le message. Les auteurs proposent notamment que l'évaluation du message est généralement sujet à des biais d'interprétations. Bien que cette théorie ait largement inspiré les modèles qui le succéderont, cette première tentative de modélisation présente deux limites majeurs. Sur un plan théorique, le modèle d'attitude n'est pas explicite et ne tient pas compte des recherches effectués sur la représentation de l'attitude (Siero et Doosje). Sur un plan méthodologique, le mécanisme des compartiments couplé à la dynamique de l'attitude est difficilement validable car non mesurable (Fraser et Stacey).

Deux modèles, le « Elaboration Likelyhood Model » (ELM) et le « Heuristic-Systematic Model » (HSM), respectivement proposés par Petty et Cacioppo (1986); Chaiken et Eagly (1989), reprennent par la suite les travaux de Fishbein et Ajzen et ceux de Sherif et Hovland, en suggérant que le rôle des réponses cognitives à un message varie en fonction des personnes et des situations. Selon ces modèles, il existerait deux modes de traitement des stimuli :

- Le mode délibératif (ou cognitif) : le sujet se sent suffisamment concerné par le stimulus et capable de porter un jugement élaboré. Il va alors analyser les informations et l'argumentaire du message en se servant de ses réponses cognitives.
- Le mode systématique (ou réactif) : le sujet n'est pas impliqué par le thème, sa motivation et/ou son aptitude à décoder le message n'est pas suffisante. Dans ce cas, la construction de son attitude va se baser sur des caractéristiques de la source (crédibilité et attractivité) ou sur les attributs du message tels que sa forme ou son esthétique.

Ces modèles diffèrent dans leurs définitions de la motivation (l'HSM définit différents types de motivations), la considération du niveau d'ancrage de l'attitude (selon l'ELM, une attitude délibérée devrait être plus ancrée et provoquer un plus grand impact sur le comportement qu'une attitude faible) et l'atteignabilité de l'attitude désirée (dans le cas où le sujet entre dans le processus délibératif, il n'est pas garanti que l'attitude résultante soit tranchée en faveur ou en défaveur, d'après HSM).

Cependant, Kruglanski et Thompson (1999) se sont interrogés sur cette dualité entre les processus réactifs, rapides et légers d'un côté et de l'autre cognitifs, élaborés et plus profonds. Selon eux, cette dichotomie n'a pas lieu d'être modélisée. Ils proposent que l'influence soit modélisée selon un unique processus. L'idée consiste à considérer toutes connaissances relatives à l'objet de l'attitude comme étant des prémisses de l'attitude sans faire la distinction entre les informations simples (*ex.* : caractéristiques de la source) et les informations plus complexes telles que l'argumentaire, les conséquences potentielles, etc. En effet, la source de l'information est tout aussi pertinente que les arguments, car il n'existe aucun autre moyen d'estimer leurs véracités. De ce fait, Kruglanski affirme qu'il n'existe pas de différences fondamentales entre ces deux types d'information.

Cette proposition nous amène à penser que nos travaux devront se baser sur un modèle adoptant une vision mono-procédurale de l'élaboration de l'attitude. C'est à dire un modèle dans lequel le rôle des caractéristiques de la source des messages ainsi que leurs contenus s'intègrent dans un processus unique de formation d'attitude. Dans les prochains paragraphes, nous présenterons les modèles récents de la dynamique d'attitude qui reprennent cette vision unifiée.

ATTITUDE DE FAZIO Fazio (2007) propose un modèle dans lequel l'attitude est considérée comme étant un ensemble d'associations mémoires entre l'objet de l'attitude et des évaluations de connaissances qui lui sont relatives. Chacune de ces associations présente une accessibilité en mémoire, plus ou moins forte. Dit autrement, la probabilité que le sujet se souvienne d'une information est plus ou moins importante selon son accessibilité. Cette notion d'accessibilité est issue des précédents travaux de Fazio et de ses collègues, portant sur la mesure de l'attitude indirecte que nous avons présentée précédemment (cf. §2.1.3.2).

En résumé : les attitudes sont définies selon la synthèse (ou agrégation) des évaluations. Cela n'implique aucunement que les attitudes soient nécessairement froides, en ne se basant que sur des évaluations rationnelles de croyances. Le terme « évaluation » n'est pas limité à des estimations analytiques comme chez Fishbein et Ajzen, mais englobe également des réactions affectives ou même des comportements antérieurs ou futurs (intentions) comme c'est le cas chez Zanna et Rempel (1988). Cette formulation de l'attitude est ainsi compatible avec le framework de type anticipation-valeur, les théories affectives et comportementales.

Une des particularités de ce modèle réside dans le fait qu'il se démarque de la majorité des autres travaux du domaine dans lesquels l'attitude est une variable hypothétique, c-à-d. une boîte noire permettant de faire le lien entre l'état mental du sujet et ses comportements. En définissant l'attitude comme un ensemble d'associations en mémoire, l'attitude devient tangible : lorsque qu'un individu a formé de telles associations évaluatives, l'attitude est représentée dans sa mémoire. L'attitude correspond alors à une forme de connaissance (pour être plus précis : une connaissance-évaluation) qui est représentée en mémoire comme pour toutes les autres connaissances. Tout comme nous associons « pain » avec « beurre », « médecin » et « infirmière » ou « beurk » avec des cafards, le whisky ou le chocolat selon les goûts.

Autre aspect important, ce modèle permet d'obtenir des attitudes caractérisées par une intensité, une amplitude (le second attribut d'une attitude, cf. 2.1.1). Elle est donnée par l'accessibilité de l'attitude, elle-même fonction de la somme des accessibilités des évaluations qui la constitue. Par exemple, un fan de foot

présente des associations évaluatives très ancrées dans sa mémoire portant sur de nombreuses équipes, coachs et joueurs. Il formera certainement une attitude envers les nouveaux joueurs et aura des réactions affectives lors des victoires ou défaites de certaines équipes. Par contre, un autre individu n'ayant que très peu d'associations évaluatives dans ce domaine, ne formera probablement pas d'attitude très prononcée envers un nouveau joueur et ne réagira pas autant que le fan aux résultats des matchs. La présence et accessibilité de ces informations dépendent de *l'intérêt* que portent les sujets à leurs égards.

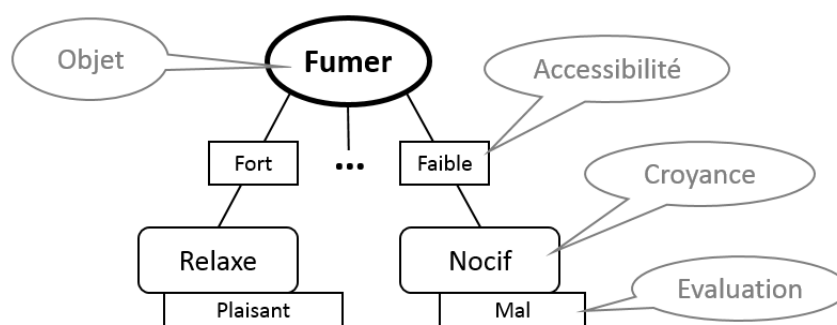


FIGURE 1 – Exemple d'une représentation de l'attitude chez un fumeur selon le modèle de Fazio

MODÈLE MÉTA-COGNITIF En se basant sur les travaux de Fazio, Petty et collab. (2007) ont proposé un nouveau modèle d'attitude l'étendant. Le principe consiste à reprendre la notion d'associations mémoire entre objet et évaluations en considérant que les évaluations peuvent être « acceptées » ou « rejetées ». Ces deux cas sont représentés par des *étiquettes* aussi appelées des *méta-cognitions* à valeur binaire sur les évaluations (méta-cognitions dans le sens où elles représentent des cognitions sur d'autres). Considérons par exemple un fumeur qui n'a jamais eu conscience du danger du tabac (voir la figure 2). Son évaluation vis-à-vis d'une cigarette est alors positive. Il apprend que fumer représente un danger pour sa santé, ce qui va lui ajouter une évaluation négative à la cigarette. Deux cas de figures se présentent : soit il rejette cette nouvelle information, auquel cas son attitude et sa représentation de la cigarette ne changent pas. Soit il accepte la nouvelle évaluation, son attitude sera dite *ambivalente* et elle en sera abaissée.

L'étiquetage permet aux individus d'avoir des croyances ambivalentes et, par conséquent, des attitudes ambivalentes. Ce mécanisme est néanmoins très proche de celui de la révision de croyances qui permet entre autre d'accepter ou de rejeter des informations. En effet, même en s'abstenant de l'étiquetage, la révision de croyance offre la possibilité aux individus de posséder différentes évaluations de signe opposées (*ex.* : « La cigarette me relaxe, mais je sais aussi qu'elle est nocive pour ma santé. ») : les attitudes peuvent être ambivalentes vis-à-vis de l'objet (cf. figure 1).

Nous venons de voir les influences cognitives sur l'attitude en relation avec la composante cognitive suivant la conclusion de la section 2.1.4. Dans la prochaine section, nous présentons les influences affectives.

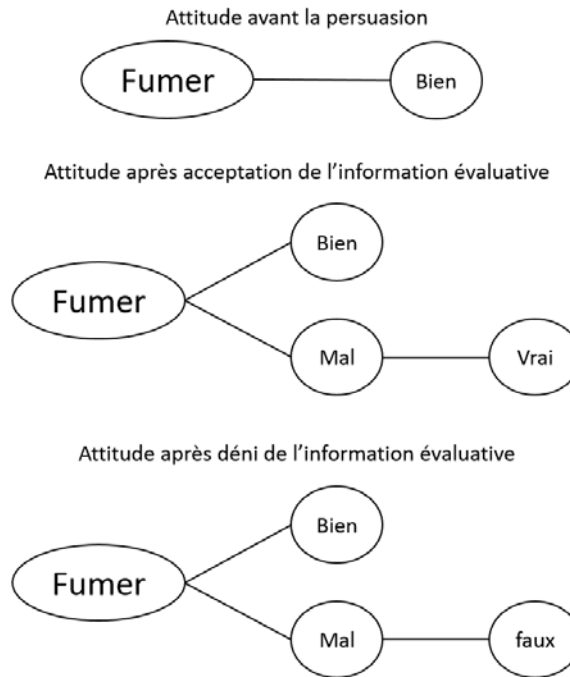


FIGURE 2 – Exemple d'évolution de l'attitude après l'acquisition d'une nouvelle information évaluative selon le modèle Méta-Cognitif tiré de Petty et collab. (2007)

2.1.6 Influences émotionnelles

Nous proposons d'étudier les travaux sur les influences affectives selon trois grandes catégories répertoriées par Maio et Haddock (2010) : l'effet de l'exposition du sujet à l'objet, l'apprentissage émotionnel et le rôle de l'humeur du sujet / ambiance de l'environnement. Cependant, dans notre cas d'étude, nous n'envisageons pas d'analyser l'humeur des individus à l'échelle d'une population. C'est pourquoi nous ne développerons que les deux premières catégories.

EXPOSITION Diverses études ont montré que la simple répétition d'expositions à un objet exerce un effet positif sur l'attitude. Ceci est valable pour n'importe quel type d'objet (*ex.* : un tableau, la nourriture, une personne, un pays) (Zajonc (1968)). Toutefois il ne s'agit pas d'un principe absolu, il est indispensable de prendre en considération l'attitude initiale vis-à-vis de l'objet. Lorsque l'attitude initiale est positive ou neutre, la fréquence d'exposition tend à la renforcer de manière significative. Dans le cas contraire, si l'attitude est initialement négative, l'effet n'est pas constaté. A l'extrême, l'attitude peut baisser dans certains cas Perlman et Oskamp (1971).

L'effet d'exposition dépend de plusieurs facteurs : l'habitude et l'ennui (Berlyne (1970)). Lorsque le sujet s'habitue à un stimulus négatif à la suite d'un certain nombre d'expositions, sa réponse émotionnelle négative va s'estomper. A l'opposé, lorsque le sujet s'ennuie de retrouver un stimulus positif, sa réponse émotionnelle positive va baisser. Bornstein (1989) proposa une extension à ce modèle dans lequel le sujet peut percevoir inconsciemment un stimulus. Au niveau inconscient, les individus ont plus de facilité à traiter une information déjà apprise par le biais d'un apprentissage subliminal. Cette *aisance de perception*

générée après l'exposition subliminale répétée, associe inconsciemment un sentiment agréable au stimulus, familiarise le stimulus et encourage un état d'esprit positif (*ex.* : [Winkielman et Cacioppo \(2001\)](#)).

APPRENTISSAGE ÉMOTIONNEL Il s'agit de l'association entre l'objet exposé et l'émotion ressentie lors de l'exposition. Nous avons identifié trois types d'apprentissages émotionnels : le conditionnement par l'exposition, le conditionnement par le comportement et le conditionnement par l'observation.

Le conditionnement par l'exposition consiste à présenter plusieurs fois l'objet de l'attitude accompagné d'une sensation affective. Ce processus ressemble fortement au conditionnement classique décrit par Ivan Pavlov, notamment célèbre pour son expérience du chien qui salive (réponse comportementale) lorsqu'il entend la sonnette (stimulus conditionné) associée à un aliment (stimulus inconditionné, qui déclenche une réponse sans la nécessité d'un apprentissage). Ces deux types de conditionnement ne diffèrent que par la nature de la réponse attendue. Le conditionnement par l'exposition attend une réponse affective interne. Les particularités notables de ce type d'apprentissage : les associations peuvent se faire de manière inconsciente (*ex.* : [Cacioppo et collab. \(1992\)](#); [Olson et Fazio \(2001\)](#)) ; l'apprentissage est résistant à l'extinction du stimulus inconditionné (une fois associée, la simple répétition du stimulus conditionné suffit à renforcer l'apprentissage) ; au bout d'un certain nombre d'expositions au couple de stimuli inconditionnel / conditionnel, l'association est accomplie ([De Houwer et collab. \(2001\)](#)).

Le conditionnement comportemental associe une réponse émotionnelle à un comportement accompli pour l'encourager, le renforcer. Par exemple, c'est par ce processus que les parents sourient à leurs enfants lorsqu'ils disent « merci » ou « s'il te plaît » afin de les encourager à être polis. Cet apprentissage peut avoir lieu aussi bien lorsque le sujet est conscient que lorsqu'il ne l'est pas. L'effet étant plus important dans le cas où le sujet n'a pas conscience de la manipulation ([Insko et Butzine \(1967\)](#)). Par ailleurs, une récompense trop importante peut avoir un effet inverse que l'on qualifie de *sur-justification* : le sujet va douter de la légitimité de son comportement ([Crano et Sivacek \(1984\)](#)). Par exemple, si on vous offre un voyage au Pérou pour avoir fini votre plat au restaurant, vous seriez en droit de vous demander si le plat était propre ou impropre à la consommation ou s'il s'agissait d'un canular.

Le conditionnement par l'observation intervient lorsqu'un individu est témoin d'une réponse émotionnelle chez une autre personne qui vient d'accomplir un comportement et ressent l'émotion par empathie (*ex.* : [Bandura \(1969\)](#); [Weiss \(1977\)](#)). C'est ce que l'on peut définir comme un conditionnement comportemental par procuration et correspond à l'une des manières dont les enfants apprennent à se comporter en « mimant » leurs parents ([Gerull et Rapee \(2002\)](#)).

Comme nous l'avons mentionné précédemment, dans nos travaux, nous nous focalisons sur la dynamique des attitudes. Le conditionnement comportemental et le conditionnement par l'observation nécessite une représentation explicite des comportements individuels. Néanmoins, il apparaît essentiel qu'un modèle d'attitude intègre *a minima* la notion d'exposition au stimuli comme mécanisme émotionnel.

2.1.7 *Attitude en psychologie sociale : bilan*

Dans cette section, nous avons vu que la notion d'attitude prend une place centrale en psychologie sociale. En effet, ce concept a un rôle majeur dans les mécanismes de prise de décision, de sociabilisation et d'organisation de la connaissance vis-à-vis d'un objet social. Cependant, malgré les abondantes recherches dans ce domaine, la mesure de l'attitude reste approximative en raison de son caractère latent. Néanmoins, elle présente au moins trois caractéristiques mesurables : la direction, l'intensité et l'accessibilité.

L'attitude se compose d'une dimension cognitive et affective qui influence sa formation ainsi que sa dynamique. Les informations cognitives (ou réponses cognitives) qui portent sur l'objet tels que les croyances sur ses impacts et leurs évaluations fondent les attitudes (Fazio (2007)). Par ailleurs, ces croyances peuvent notamment être acquises par différents canaux tels que la communication ou la perception directe. Ainsi la question de la crédibilité de la source devient importante dans l'estimation de la confiance accordée à l'information.

La dimension affective de l'attitude dépend principalement de deux types de mécanismes. Le premier est en rapport avec la fréquence d'exposition du sujet à l'objet d'attitude qui peut engendrer un phénomène d'habituation ou d'ennui. Le second porte sur le conditionnement dans lequel le sujet associe l'objet de l'attitude ou un comportement envers celui-ci à une sensation affective.

Dans la prochaine section, nous verrons comment ce concept de l'attitude est modélisé dans les simulations informatiques dont le but est de décrire les dynamiques d'attitudes et d'opinions dans des populations humaines.

2.2 ATTITUDES EN SIMULATION SOCIALE

Les sciences comportementales et sociales telles que la psychologie sociale et les sciences cognitives ont permis de déceler d'importants processus guidant la cognition, l'émotion, l'intention et les comportements humains. Parmi ces théories, nous pouvons par exemple mentionner l'apprentissage social, la motivation émotionnelle ou encore le comportement planifié (De Houwer et collab. (2001); Hogg et Abrams (1990); Ajzen (1991)). Malgré les décennies de recherches sur ces thématiques, la description temporelle des dynamiques résultantes des interactions individuelles au sein d'une population reste un problème difficile. La modélisation de tels problèmes résulte en des systèmes dits *complexes* car leurs dynamiques ne peuvent pas être exprimées par la somme des dynamiques de leurs composantes. Dans ce type de problème, la description des relations entre les dynamiques au niveau micro (dynamiques individuelles) et des phénomènes au niveau macro (dynamique du groupe, du système dans son ensemble) relève d'une extrême difficulté. Or, la compréhension de telles dynamiques complexes est nécessaire dans l'accomplissement de l'un des objectifs principaux des sciences sociales : éviter et gérer les problèmes sociétaux.

La simulation sociale combine la modélisation informatique basée sur l'approche agents aux sciences sociales. En effet, la modélisation orientée agents est fondée sur une modélisation microscopique des individus et de leurs interactions, permettant ainsi d'analyser les dynamiques individuelles et les manifestations émergentes au niveau du système (Drogoul et collab. (2003)). Dans ce sens, elle constitue une approche adéquate à l'étude des phénomènes sociaux complexes.

En conséquence, la modélisation informatique a participé à la compréhension et à l'identification de processus influençant la dynamique sociale. Par exemple, des recherches dans ce domaine ont permis d'éclaircir certains phénomènes sociaux réels tels que la ségrégation, les dynamiques d'opinion qui amènent à une polarisation de la population ou encore l'expansion des minorités (Sakoda (1971); Deffuant et collab. (2000); Hegselmann et Krause (2002); Nowak et collab. (1990)).

Le concept des attitudes a notamment fait l'objet de nombreuses recherches en simulation sociale. Dans cette section, nous présentons les principaux modèles de dynamiques d'attitudes qui ont été implémentés dans cette discipline.

2.2.1 Premiers modèles

Les premiers travaux portant sur la dynamique d'attitudes s'inspiraient des modèles issus de la physique statistique. De ce fait, les attitudes étaient souvent représentées sous la forme de valeurs binaires. Par exemple, un grand nombre de recherches a porté sur l'implémentation du modèle de l'électeur¹ dans lequel l'attitude d'un individu est directement associée à son comportement de vote (1 pour une voix favorable et 0 dans le cas contraire). Autre exemple de modèle statistique qui a souvent été repris, le modèle d'Ising dans lequel les phénomènes collectifs sont produits par des interactions locales entre des particules à deux états (Glauber (1963)). C'est par exemple de cas de Nowak et collab. (1990) que nous décrivons dans le paragraphe suivant.

2.2.2 Des attitudes privées aux opinions publiques

Nowak et collab. (1990) ont proposé de simuler informatiquement le processus de la révision d'attitude dans une population résultant de la théorie de l'impact social de Latane (1981). Cette méta théorie avance que les croyances, attitudes, comportements d'un individu sont influencés par ceux de ses voisins. Cette influence dépendrait de la *magnitude* de l'impact des sources sur l'individu en question qui est fonction de

- la *force* : elle peut représenter la crédibilité ou l'attractivité de la source, elle a été obtenue selon des données telles que l'âge ou le statut de l'individu au cours d'autres travaux en sciences sociales ;
- l'*immédiateté* : correspond à la proximité entre la source et l'individu, cette proximité n'est pas forcément limitée à une dimension spatiale mais aussi sociale, temporelle ou même cognitive (*ex.* : attitudes, croyances partagées).
- le nombre de sources d'influences.

Cette théorie s'inspire des équations statistiques décrivant les forces électromagnétiques en physique des particules. En effet, les particules représentent ici des individus et la force correspond à un vecteur d'attributs variés.

Les agents sont caractérisés par trois paramètres initialisés selon une distribution uniforme :

- L'attitude : une valeur binaire indiquant si l'individu est « pour » ou « contre », elle n'a pas d'objet spécifique ;

1. Processus stochastique d'interaction de particules modélisé à l'aide des mathématiques probabilistes (<https://goo.gl/t9x0Ui>)

- *La capacité de persuasion* : une valeur comprise dans $[0, 100]$, elle correspond à la *force* de la source ;
- *La capacité de résistance* : une valeur comprise dans $[0, 100]$.

Ces individus sont localisés sur une matrice carrée, l'immédiateté se calcule alors selon la distance euclidienne entre le sujet et la source. De plus, un agent est associé à un groupe défini par tous les individus limitrophes partageant la même attitude.

L'impact de persuasion est défini par :

$$i_p = N_o^{1/2} \left(\frac{\sum (p_i/d_i^2)}{N_o} \right)$$

où N_o le nombre de sources (taille du groupe influant), p_i la capacité de persuasion de l'individu i du groupe de persuasion et d_i la distance entre l'individu et le destinataire (l'individu qui est influencé). *L'impact de la résistance* est défini par :

$$i_r = N_r^{1/2} \left(\frac{\sum (r_i/d_i^2)}{N_r} \right)$$

où N_r le nombre d'individus du groupe de résistance (taille du groupe ciblé par l'influence), r_i la capacité de résistance de l'individu i du groupe de résistance. La règle de révision de l'attitude consiste à faire adopter à un sujet l'attitude de la source si l'impact de persuasion du groupe de la source est plus grand que la résistance du groupe du sujet. Suivant une révision d'attitude, l'individu se voit réassigner de nouvelles valeurs de capacités de persuasion et de résistance.

Les résultats de simulations montrent que les attitudes de la population ont une forte tendance à s'homogénéiser : toutes les simulations ont conduit à une attitude majoritaire adoptée par plus de 80% environ lorsque le système atteint un équilibre. De plus, [Nowak et collab. \(1990\)](#) ont remarqué que les groupes minoritaires avaient plus de chance de résister sur les bords de la matrice (les individus qui sont le moins connectés), de manière similaire aux groupes marginaux que l'on peut observer dans le monde réel (voir la figure 3). De manière générale, le modèle reproduit une dynamique de domination par laquelle l'attitude majoritaire en début de simulation est renforcée au cours de la simulation. Cependant, dans certains cas rares où un groupe minoritaire a une grande capacité de résistance, ce groupe pouvait croître au détriment de l'attitude majoritaire.

Bien que la représentation de l'attitude soit rudimentaire, ces travaux ont illustré la capacité des modèles informatiques à reproduire des phénomènes sociaux en s'inspirant de théories psycho-sociales. En effet, leur mécanisme d'interaction fait intervenir différentes notions issues des recherches sur l'attitude en psychologie que nous avons mentionnées dans la première partie de ce chapitre. Dans nos travaux, nous tenterons d'adopter une approche similaire et fonderons notre modèle d'attitude sur des théories issues de sa discipline originelle.

Par la suite, les modèles d'attitude se sont tournés vers des représentations continues. Parmi ces travaux, le *Bounded Confidence Model* (BCM) proposé par [Deffuant et collab. \(2000\)](#); [Hegselmann et Krause \(2002\)](#) a marqué un tournant en simulation sociale. En effet, ce modèle a suscité l'intérêt de nombreuses communautés scientifiques telles que la physique sociale ou les sciences complexes (*ex* : [Castellano et collab. \(2009\)](#)). Dans ce modèle, les agents adaptent graduellement leurs attitudes, représentées sous forme d'une variable continue, en fonction des



FIGURE 3 – Exemple d’une configuration aléatoire d’attitudes (à gauche) et l’équilibre atteint au bout de 18 pas de temps (à droite). (Chaque cellule représente un individu. Les I représentent l’opinion majoritaire ; les - indiquent le point de vue opposé) Cet exemple est tiré de [Nowak et collab. \(1990\)](#)

opinions des autres individus qui leur sont transmises par le biais de la communication inter-personnelle. L'idée proposée est que l'influence ne peut pas aboutir si l'écart entre les attitudes des interlocuteurs dépasse un certain seuil fixe. Dans le cas où cet écart est suffisamment faible, autrement dit si les attitudes sont proches, les individus vont rapprocher leurs attitudes selon un certain facteur constant. Avec cette règle locale de révision d'attitude, le modèle présente des résultats intéressants au niveau macro : on observe des phénomènes de polarisation avec la formation de différents groupes d'individus dont les attitudes convergent. Cependant, ce modèle présente deux limites vis-à-vis de notre objectif. D'une part, comme la majorité des autres travaux dans ce domaine, ce modèle traite le processus de la formation de l'attitude comme étant une *bo te noire*, en se focalisant sur l'impact des échanges d'opinions entre les individus. En effet, ces travaux ne modélisent que les conséquences de la formation de l'attitude sans traiter la question de la formation en soi (les causes). D'autre part, ce modèle ne s'appuie sur aucune théorie socio-psychologique, soulevant la question de sa plausibilité empirique. Sur ces points, [Urbig et Malitz \(2007\)](#) proposent un modèle, que nous présentons ci-après, traitant à la fois le problème de la formation et de l'impact des interactions qui s'appuie sur des théories issues des sciences humaines.

2.2.3 Attitudes multi dimensionnelles et exposition sélectionnée

Dans leur modèle, [Urbig et Malitz \(2007\)](#) proposent une représentation de l'attitude sous forme d'un vecteur *d impressions* vis-à-vis d'un certain objet social. Ces impressions représentent des évaluations de croyances sur des caractéristiques de l'objet, mais elles ne sont pas explicitement décrites dans le modèle. En conséquence, elles sont représentées par des valeurs continues comprises dans l'intervalle $[0, 1]$. En s'inspirant de la définition de l'attitude d'[Eagly et Chaiken \(1993\)](#), la valeur de l'attitude correspond à la moyenne de ces impressions. Au début des simulations, les attitudes (les impressions qui les composent, pour être plus précis) sont distribuées selon une loi uniforme.

Le mécanisme d'influence consiste à faire rencontrer aléatoirement des paires d'agents à chaque pas de temps afin qu'ils échangent (communiquent) leurs attitudes. Chacun des deux agents va comparer chacune de ses impressions à celles de son interlocuteur. Intervient alors le BCM, si les impressions sont suffisamment proches (selon un seuil ϵ_f constant) et si les deux attitudes globales sont suffisamment proches (selon un second paramètre fixe ϵ_a), les agents convergent leurs impressions de manière symétrique (selon un facteur de convergence fixé).

Les expérimentations portent principalement sur l'étude de la distribution des attitudes résultant des simulations. Ces résultats exposent différents types de convergences selon les valeurs de paramètres employées. Ainsi, leur modèle est capable de produire des formations de plusieurs groupes plus ou moins nombreux tels que dans le BCM (voir la figure 4) et une tendance générale de radicalisation par laquelle les attitudes convergent vers des valeurs extrêmes (proches des bornes).

Ces travaux présentent trois principales limites en regard de nos objectifs. Premièrement, bien que ce modèle aborde le problème de la formation de l'attitude en s'appuyant sur une théorie de l'attitude issue de la psychologie sociale, les évaluations de croyances (les impressions selon leur vocabulaire) sont données en amont des simulations et ne sont pas calculées par les agents. Le problème de

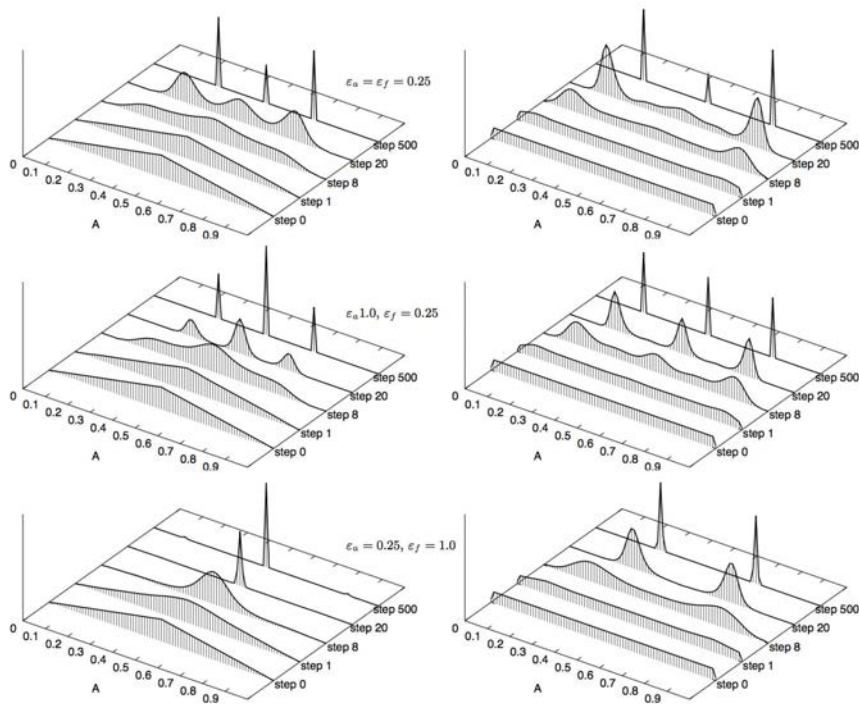


FIGURE 4 – Evolution de la distribution des attitudes selon différents scénarii. La colonne de gauche représente les cas où les impressions initiales sont distribuées selon une loi normale, en conséquence les attitudes suivent une distribution triangulaire. La colonne de droite représente les cas où les attitudes sont distribuées uniformément. Cette illustration est tirée de [Urbig et Malitz \(2007\)](#)

l'acquisition des croyances est écarté et le modèle se focalise sur les influences inter-personnelles.

Deuxièmement, ces influences se résument à l'application du BCM et la révision des croyances correspond à une moyenne pondérée des impressions ainsi obtenues. Pourtant, aucune théorie en sciences sociales ne supporte l'idée d'une moyenne de croyances pour décrire une révision de croyance. D'autre part, toutes les croyances ont une accessibilité égale en mémoire, contredisant ainsi les conclusions expérimentales en psychologie sociale et la notion de la rationalité limitée [Simon \(1955\)](#). Enfin, l'usage du BCM n'est pas compatible avec l'approche proposée dans ces travaux : bien que le modèle parte d'une définition psycho-sociale de l'attitude (celle d'[Eagly et Chaiken \(1993\)](#)), le BCM n'est pas défendable dans cette discipline particulière comme nous l'avons mentionné dans la section précédente.

Troisièmement, malgré les multiples expériences qui ont été conduites dans ces travaux, aucune ne fait appel à des données particulières. Il s'agit d'ailleurs de la principale critique de [Chattoe-Brown \(2014\)](#) sur les travaux en simulation sociale : ces études se concentrent principalement sur des analyses de sensibilité dans des contextes abstraits, les expériences ne sont pas validées sur des données réelles. Or, sans validation sur des cas réels, la plausibilité empirique ainsi que l'applicabilité des modèles restent en suspens. Dans nos travaux, nous verrons comment nous avons appliqué notre modèle à un cas d'étude concret, celui de l'intervention militaire française en Afghanistan.

Dans le prochain paragraphe, nous présentons un modèle d'attitude comblant certaines de ces lacunes.

2.2.4 *Simulation des attitudes politiques et le comportement de vote*

[Kottonau et Pahl-Wostl \(2004\)](#) proposent un modèle permettant de simuler la formation, la dynamique d'attitudes et le comportement de vote d'une population durant une campagne présidentielle. Trois types d'acteurs sont modélisés :

- Les candidats ou partis politiques : ils émettent des messages de persuasion en fonction de leurs stratégies de communication. Une stratégie de communication consiste à répartir les efforts de communications sur un intervalle de temps plus ou moins long qui précède le jour de l'élection. Il est à noter que le modèle ne prévoit que deux candidats antagonistes.
- Les médias : ces acteurs reprennent les communications des candidats afin de les amplifier. Cela se traduit dans le modèle par l'augmentation du nombre de messages persuasifs selon un facteur multiplicatif. En outre, les médias ne sont pas neutres, ils peuvent favoriser certains candidats qu'ils supportent (chaque média possède des couples de facteurs multiplicatifs pour chaque candidat).
- Les individus de la population : ils sont la cible des messages persuasifs ; leurs perceptions sont biaisées ; ils sont caractérisés par un degré d'intérêt (dynamique) aux élections déterminant l'impact des messages persuasifs ; ils conversent au sein d'un réseau social ; en fonction de leurs connaissances, ils se forment une attitude pour chacun des deux candidats.

La communication inter-personnelle est initiée selon une probabilité qui dépend de l'intérêt de l'individu ainsi que d'une constante, paramètre de simulation. Dans le cas où l'individu a une position ambivalente (ses deux attitudes envers les deux

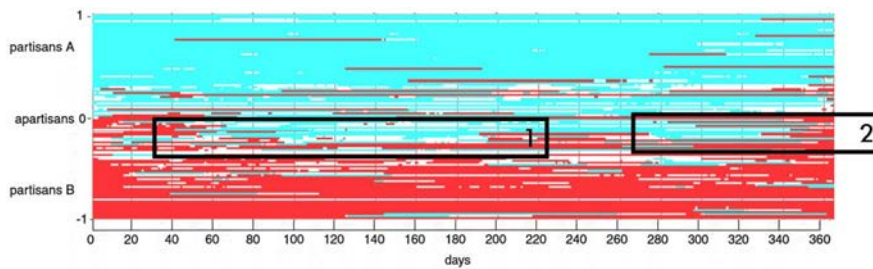


FIGURE 5 – Volatilité de 100 électeurs sur 365 jours. Cette illustration est extraite de [Kottonau et Pahl-Wostl \(2004\)](#) - L'encadré 1 correspond à une période d'effort du parti A et 2 celle du parti B

candidats sont similaires), il communique des arguments positifs et négatifs. Dans le cas contraire, si son attitude est suffisamment univalente, il communique la valence de son attitude. La crédibilité de cet individu est calculée en fonction de son niveau d'ambivalence par son interlocuteur. Par ailleurs, la crédibilité accordée aux individus est plus importante que celle des médias, elle-même supérieure à celle des candidats.

Le comportement de vote (le fait de se rendre aux urnes) est déterminé selon le niveau d'ambivalence : si un individu possède une attitude suffisamment forte (univalente) il ira voter.

Le niveau de persuasion des messages est déterminé selon deux facteurs : la *marque affective* du message qui est prédéfini par le modélisateur et la crédibilité de la source du message. Ces informations perçues par les individus constituent leurs connaissances sur les différents candidats. De plus, une fois intégrées, ces croyances sont pourvues d'une valeur d'accessibilité au sens de [Fazio \(1995\)](#).

L'accessibilité d'une croyance est fonction de la fréquence et de la récence des perceptions des messages persuasifs. Elle décroît en fonction du temps. Cette fonction s'inspire de [Anderson et Schooler \(1991\)](#) ainsi que de [Burnkrant et Sawyer \(1983\)](#).

Le biais de confirmation est modélisé par une hausse de la crédibilité d'une nouvelle information perçue, si celle-ci va dans le sens de l'attitude de l'individu en s'inspirant notamment de [Lord et collab. \(1979\)](#). Nous ne détaillerons pas les deux autres biais cognitifs car ils sont intégrés dans des formules relevant de mécanismes plus généraux.

Le niveau d'intérêt pour la politique est fonction de la fréquence de perception des messages persuasifs, en se référant à une définition de [Inkeles \(1969\)](#) (« interest in public affairs validated by keeping informed and [is] expressed through participation in civic action »).

L'intensité de la réponse (cognitive) du sujet à un message est une fonction de son accessibilité et de sa crédibilité.

Les intensités de réponses aux messages associés à un candidat sont transformées en une attitude envers cet objet social selon une somme normalisée.

Deux expérimentations ont été conduites dans le cadre de ces travaux. La première correspond à une analyse du modèle au niveau microscopique. L'objectif est de confirmer l'impact et le bon fonctionnement des différents mécanismes implémentés. Le protocole consiste à montrer les résultats d'une simulation afin d'observer les trajectoires d'attitudes des individus et les basculements d'atti-

tudes. Par exemple, dans la figure 5, nous pouvons observer que dans l'encadré (1) correspondant à une période d'effort du parti A (en bleu), de nombreux citoyens basculent en sa faveur. De la même manière, l'encadré (2) illustre la dernière rafale de communication du parti B (en rouge) grâce à laquelle il emporte la majorité.

La seconde expérience correspond à une analyse de sensibilité du modèle. Cette fois, le protocole appliqué consiste à simuler des scénarii (stratégies de communication) fictifs en modulant quelques paramètres majeurs. Les auteurs sont alors capables de comparer les stratégies selon les situations. Ces situations sont discriminées par (1) l'inhibition d'une partie du modèle et (2) une mémoire plus ou moins courte des agents.

Les auteurs concluent que cette approche permet de comparer différentes stratégies de communication en période électorale en les simulant au sein de ce modèle.

Ces travaux présentent de nombreux éléments de réponse à notre problème de formation et de dynamique d'attitudes d'une population faisant face à des acteurs sociaux. En effet, les candidats de ce modèle correspondent aux acteurs sociaux que nous cherchons à modéliser. De plus, les communications de ces candidats sont une des notions que nous voulons intégrer dans nos travaux. Nous partageons également l'idée que la communication inter-personnelle doit se baser sur l'échange de croyances et pas seulement sur les attitudes (opinions pour être plus précis). Plus globalement, l'approche adoptée par [Kottonau et Pahl-Wostl \(2004\)](#) qui consiste à s'inspirer directement des théories psycho-sociales correspond à notre vision sur la recherche en simulation sociale.

Cependant, ce modèle présente plusieurs limites par rapport à notre objectif.

- Le niveau de description du scénario : dans notre cas d'étude, nous ne pouvons pas nous limiter à un scénario décrit par une simple fonction décrivant le taux d'effort des acteurs sociaux. En effet, nous voulons modéliser une variété d'actions ainsi que l'impact de chacune d'entre elles sur les attitudes des individus.
- La notion de groupe social : en reprenant l'exemple précédent, nous pouvons constater qu'il est important de représenter différents groupes sociaux, que ce soit pour la simulation des attitudes politiques (*ex.* : nationalité, classe sociale), des ventes d'un produit commercial (*ex.* : tranches d'âges pour la vente de l'iPhone) ou encore les attitudes envers des Forces armées (*ex.* : rivalités entre les Pashtun et les Tajik en Afghanistan).
- L'accessibilité d'une croyance : bien que le calcul de cette notion s'inspire de plusieurs théories empiriques, il reste *ad-hoc* au cas d'application portant sur les élections. Nous verrons que notre modèle s'appuie sur la Théorie de la Simplicité afin de décrire cette notion d'accessibilité d'une manière plus générique (voir section 2.2.5 ci-après).
- La révision de croyance : ce modèle n'intègre pas de mécanisme permettant de modifier une croyance. Cela provient du fait que l'information relatée par un message se résume à une *marque affective* qui est toujours « vraie ». Or, nous pouvons imaginer que toutes les informations émises ne sont pas vraies, d'autant plus dans le domaine politique (*ex.* : « *L art du mensonge politique* » de [Swift et Courtine \(2007\)](#)). Dans le cas où les acteurs mentent, il peut exister des informations contradictoires. Il est alors nécessaire que

les individus puissent choisir les informations auxquelles croire à l'aide d'un mécanisme de révision de croyances.

Malgré les nombreuses contributions de ce modèle à l'étude de la formation et des dynamiques d'attitudes, ces travaux ne sont pas directement applicables à notre cas d'étude. Les limites se trouvent tant au niveau conceptuel (le niveau de description du scénario et l'absence de la notion de groupe social) qu'au niveau des mécanismes cognitifs (la révision de croyance et le calcul de l'accessibilité). Dans la prochaine section, nous présentons la théorie sur laquelle notre calcul de l'accessibilité d'une croyance se basera.

2.2.5 Théorie de la Simplicité

La Théorie de la Simplicité est une théorie relative aux sciences cognitives. Elle vise à expliquer l'intérêt des situations et des événements pour l'esprit humain. Cette théorie se base sur les travaux de [Dimulescu et Dessalles \(2009\)](#); [Chater et Vitányi \(2003\)](#); [Vitanyi et Li \(2002\)](#). Elle ne relève donc pas directement des théories attitudinales ou émotionnelles et sa formulation n'intègre pas le concept d'attitude. Cependant, elle prend en compte une réponse émotionnelle relative à l'impact du stimuli et une composante relevant de la surprise (une des *émotions de base* de [Plutchik \(1980\)](#) et de [Darwin et collab. \(1998\)](#)). De plus, le concept d'intérêt d'une information est compatible avec la notion d'accessibilité mémoire d'une information qui est notamment mentionnée dans le modèle d'attitude de Fazio. En effet, l'intérêt d'une information correspond à la fois à son intérêt narratif (l'intérêt de parler d'un événement lors d'une conversation sociale) et à l'intérêt « brut » en mémoire : on parle de ce dont on se souvient ([Chater et Vitányi \(2003\)](#); [Dimulescu et Dessalles \(2009\)](#)).

IDÉE GÉNÉRALE Le fondement de cette théorie est de se baser sur l'hypothèse suivante : les situations intéressantes seraient celles ayant (1) un fort impact émotionnel, (2) celles qui surviendraient de manière inattendue / surprenante et enfin (3) les plus simples aux yeux de l'observateur. Ici, la simplicité correspond à la baisse de la complexité au sens de [Kolmogorov \(1956\)](#), c'est-à-dire que la description de la situation est plus courte que celle anticipée par l'observateur. La longueur de la description est à comprendre au sens de la théorie de l'information de [Shannon et Weaver \(1949\)](#) : elle correspond à la longueur du plus court programme permettant de décrire la situation (une des unités de mesure étant le nombre de bits). Par exemple, la description d'un tirage du Loto tel : 22-23-24-25-26-27, est plus courte que celle d'un tirage typique tel : 12-22-27-37-38-42. La première description ne nécessite qu'une instance (le choix d'un numéro parmi tous ceux qui peuvent apparaître au Loto), suivie de cinq numéros, tandis que la seconde nécessite six instances différentes.

FORMULATION *L'impact émotionnel* (1) correspond à la réponse émotionnelle E du sujet lors de son exposition à l'information décrite selon la loi du stimuli proposée par [Fechner \(1832\)](#).

$$E(\text{stimuli}) = \log \left(1 + \frac{|\text{stimuli}|}{\xi} \right) \quad (1)$$

où $\xi \in]0, 1]$ est un paramètre de sensibilité émotionnelle non nul.

La surprise S (2) engendrée chez l'observateur dérive des sentiments :

- d'inattendu brute U_r : la probabilité que cet événement survienne sur la base de ce qu'il a déjà observé auparavant
- d'inattendu personnelle U_p : à quel point l'observateur s'attendait à ce que l'événement le touche personnellement

$$S = U_r - U_p \quad (2)$$

Par exemple, on sait qu'il est rare de trouver un billet de 100 euros sur le trottoir, l'événement est donc surprenant. Mais cette surprise peut être atténuée s'il vous est arrivé d'avoir ramassé plusieurs billets de 100 euros dans le passé. Autre exemple illustré dans la figure 6 : deux touristes à Hawaï, l'un français et l'autre japonais, témoignent d'un tsunami. Cet événement est en soi surprenant, car les tsunamis sont rares. Cependant, la surprise chez le Français, qui vit un tsunami pour la première fois de sa vie sera certainement plus importante que chez le Japonais qui est relativement « habitué » à ce type de situation.



FIGURE 6 – Exemple de différence entre l'inattendu brut et personnel

Ces deux types d'inattendus U sont calculés selon la forme :

$$U = C_w - C_d \quad (3)$$

où C_w représente la **complexité de génération** (la probabilité que l'événement se produise) et C_d la **complexité de description** de l'événement (au sens de Shannon). Cette complexité de description est négative comme on peut le constater dans la formule 2.2.5, elle correspond en fait à la notion de simplicité. Les deux complexités C_w et C_d peuvent être calculées sur plusieurs dimensions : temporelle, géographique, sociale etc. Par exemple, le fait que votre voisin meure est « surprenant » en soi, a priori c'est une information importante (« intéressante » dans les termes de cette théorie). Si ce défunt n'est pas notre voisin mais un habitant du Kiribati, la distance géographique et la complexité de description du lieu sont si importantes que cette information ne nous touchera que très peu : l'information ne sera probablement pas retenue et certainement pas relayée. On retrouve ainsi la règle de « la mort kilométrique » communément appliquée par les reporters (Leyens et collab. (2000)).

Second exemple, illustré par la figure 7 : pourquoi ne connaissez-vous pas la Tour Taxis Fenestralis, pourtant très ressemblante à la Tour de Pise ? Nous conviendrons que la Tour de Pise est célèbre car elle penche. En effet, la complexité de génération qu'une tour soit penchée est élevée (il existe peu de tours penchées). La complexité de description, quant à elle, est identique à toutes les

autres tours : c'est une tour (complexité faible). C'est bien la différence entre ces deux complexités qui distingue la Tour de Pise des autres tours.

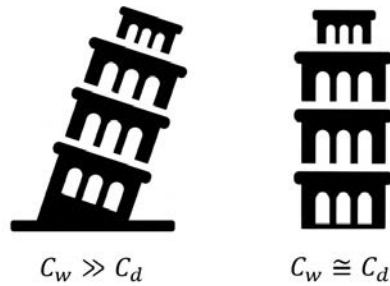


FIGURE 7 – Exemple de différence entre la complexité de génération et de description

2.3 SYNTHÈSE : VERS UN MODÈLE DE LA DYNAMIQUE D'ATTITUDES

Les études sur les travaux en sciences humaines et sociales ainsi qu'en simulation sociale sur l'attitude que nous avons présentés dans ce chapitre ont nourri notre modèle informatique de la formation et de la dynamique des attitudes. Nous défendons l'idée qu'un modèle informatique de l'attitude doit se fonder sur les travaux en sciences humaines et sociales qui sont à l'origine de ce concept. Dans cette section, nous présentons les principaux objectifs de notre modèle mis en rapport avec leurs sources d'inspiration, tout en nous démarquant des travaux existants en simulation sociale.

OBJET, VALENCE ET INTENSITÉ Notre modèle d'attitude présentera au minimum trois attributs caractéristiques qui sont acceptés de manière unanime par la communauté psycho-sociale (cf. section 2.1.1) :

- L'objet : un objet social qui peut être physique, humain ou conceptuel vis-à-vis duquel l'individu possède une attitude. Dans notre cas particulier, les attitudes porteront sur les acteurs sociaux ayant un impact sur la vie sociale de la population et dont nous étudions l'évolution des attitudes.
- La valence : le signe, positif ou négatif de l'attitude.
- L'intensité : l'amplitude de l'attitude qui peut être plus ou moins forte.

ACCESSIBILITÉ Les attitudes ainsi que les croyances qui les constituent sont caractérisées par une valeur d'accessibilité mémoire (voir §2.1.3.2). Cette notion correspond au niveau de réminiscence de l'information (cf. §2.1.3.2 et section 2.3). Contrairement aux travaux de [Kottonau et Pahl-Wostl \(2004\)](#), nous proposerons une implémentation générique du calcul de ce concept basée sur la Théorie de la Simplicité (cf. section 2.2.5). L'accessibilité servira notamment à calculer la probabilité des transmissions inter-personnelles de messages narratifs.

FONCTION Dans notre modèle, l'attitude a une fonction heuristique et schématique (cf. section 2.1.2) :

- Elle sert à l'estimation de l'impact d'une action provoquée par un acteur social (fonction heuristique).

- Nous verrons par la suite que notre modèle d’attitude possède une représentation mémoire organisant les croyances de l’individu (fonction schématique).

DEUX COMPOSANTES En se référant aux travaux en psychologie sociale (cf. section 2.1.4), notre modèle d’attitude se compose d’une dimension cognitive et affective. Nous n’intégrons pas de composante comportementale, d’une part car elle est controversée (cf. section 2.1.4) et d’autre part car nous ne nous intéressons pas aux comportements des individus dans le cadre de nos travaux. La composante cognitive porte sur les croyances de l’individu. La composante affective représente les réponses émotionnelles de l’individu lors de l’acquisition des informations.

INFORMATIONS COGNITIVES Nous permettons au modélisateur de décrire des scénarii composés d’une variété d’actions qui peuvent impacter la population ou des acteurs sociaux. Ce sont les informations relatives à ces actions qui constituent la cognition des agents. Ainsi, les croyances de nos agents ne sont pas limitées à une ou quelques valeurs continues (*ex* : [Urbig et Malitz \(2007\)](#); [Kottonau et Pahl-Wostl \(2004\)](#)) définissant leurs impacts. En effet, les actions dans notre modèle sont définies par un nom (permettant de savoir de quoi il s’agit), une date ou encore par un bénéficiaire.

GROUPES SOCIAUX La population peut être divisée en plusieurs groupes sociaux composés d’agents partageant une affiliation commune telles que l’appartenance à un groupe religieux, ou ethnique. Nous apportons par ce biais une identification sociale aux individus qui leurs permettent de se positionner vis-à-vis d’un événement.

CARACTÉRISTIQUE DE LA SOURCE La probabilité de persuasion par un message sera fonction de la crédibilité de sa source (cf. §2.1.5.1). Nous employons pour cela un mécanisme de révision de croyance proposé par [Thiriou et Kant \(2007\)](#) qui prend en compte cette notion de crédibilité.

RÉVISION DE CROYANCE Nous avons vu qu’un agent cognitif manipulant des croyances doit se doter d’un mécanisme de révision de croyances si nous voulons décrire des scénarii riches. Par exemple, pour décrire des situations où l’impact d’une action (ou d’une information) évolue : « Je n’aime pas la Force Bleue car leurs patrouilles font des dégâts. » vers « J’aime la Force Bleue car leurs patrouilles ne font plus de dégâts et sont sécurisantes. ». Autre exemple, pour simuler des phénomènes de désinformation : un acteur social pourrait tenter de dissimuler une bavure en diffusant le message « L’opération d’hier n’a fait aucune victime civile. » à la radio alors qu’un individu a été témoin du contraire « J’ai vu une victime civile. ». Enfin, la révision de croyance permet aussi le déni d’une information tel que décrit par [Petty et collab. \(2007\)](#) dans le Modèle Méta-Cognitif.

Implémentation

Notre modèle repose sur l’implémentation de deux théories.

ATTITUDE DE FAZIO La structure de notre modèle d'attitude se base directement sur le paradigme proposé par Fazio : l'attitude est un ensemble d'associations mémoire entre l'objet et ses évaluations (cf. section 2.3). Ainsi nous représenterons explicitement les croyances, les évaluations et les associations.

THÉORIE DE LA SIMPLICITÉ Dans notre modèle, l'accessibilité ainsi que la probabilité de communication inter-personnelle sont fonctions de l'intérêt de l'information calculé selon la Théorie de la Simplicité. Elle stipule que les informations intéressantes sont celles ayant un fort impact émotionnel, celles qui surviendraient de manière inattendue / surprenante et enfin les plus simples aux yeux de l'observateur (cf. section 2.2.5). Notamment, cette théorie prend en compte la notion de la répétition d'expositions à un stimulus (phénomène d'habituation et d'ennui) (cf. section 2.1.6).

Notre vision d'un modèle de formations et de dynamiques d'attitudes d'une population ainsi établie se concrétise dans le chapitre suivant.

Notre modèle, que nous nommons Polias, repose sur le principe suivant : une simulation correspond à l'exécution d'actions (discours politique, campagne publicitaire, etc.) par des acteurs (parti politique, Force militaire, enseigne ou autres) qui sont perçues au cours du temps par les individus constituant le terrain humain (section 3.2). Ces perceptions nourrissent les croyances de la population, les individus ont une plus grande connaissance des acteurs. De plus, dans le cas où des informations contradictoires circulent, les individus peuvent être amenés à réviser leurs croyances (section 3.3). Les informations ainsi acquises sont évaluées localement par chaque agent en fonction de leurs connaissances, émotions, sentiment de surprise, attitudes et affiliations à des groupes sociaux (section 3.4). Les informations effectuées par les acteurs peuvent se diffuser au sein du réseau social qui lie la population par le biais d'un mécanisme de communication inter individuelle (section 3.5). En se référant à ces éléments cognitifs, les individus forment et adaptent leurs attitudes vis-à-vis des acteurs (section 3.6).

Avant de présenter ces mécanismes de dynamique d'attitude, nous commençons par décrire la structure de données du modèle avec ses différents objets conceptuels qui sont à la base de notre simulation dans la section suivante (section 3.1).

3.1 STRUCTURE DE DONNÉES

Cette section définit les objets conceptuels constituant notre modèle multi-agent de dynamique d'attitude.

3.1.1 *Protagonistes*

Nous distinguons deux types de protagonistes dans notre simulation. D'une part la population qui constitue le cœur de notre étude. Elle est constituée d'*individus*, chacun représenté par un agent computationnel doté d'une cognition qui permet de percevoir les actions, stocker les informations en mémoire, construire des attitudes et communiquer via le réseau social. Ces individus sont par ailleurs catégorisés en différents groupes sociaux notés *SG* qui divisent le terrain humain.

D'autre part les *acteurs* qui sont à l'origine des actions et constituent la cible des attitudes de la population. Chaque acteur est représenté par un automate dont la séquence d'actions est définie par l'utilisateur. Les acteurs ne font qu'exécuter ces actions au cours de la simulation.

INDIVIDUS DE LA POPULATION Soit Ind l'ensemble des agents, représentant les individus de la population. Chaque individu $i \in \text{Ind}$ est défini par le tuple

$$i = \langle \text{name}, \text{sg}, \text{Blf}, \text{Cnt} \rangle$$

avec :

- **name** l'identifiant unique de l'agent
- **sg** $\in SG$ le groupe social auquel appartient l'individu
- **Blf** l'ensemble des croyances présentes dans la mémoire de l'individu (le modèle de représentation des croyances est décrit dans la section 3.1.3)
- **Cnt** $\subset Ind - \{i\}$ l'ensemble des contacts de l'agent dans le réseau d'interactions (c.f. section §4.3)

ACTEURS Soit **Actor** l'ensemble des acteurs de la simulation. Chaque acteur $actor \in Actor$ est caractérisé par sa liste d'actions $aList_{actor}$ définie en paramètre de simulation. Les actions sont définies dans la section 3.1.3.

3.1.2 Objets sociaux

Nous appelons objet social une entité abstraite ou concrète, humaine ou artificielle vis-à-vis de laquelle des individus (deux a minima) exercent un comportement social (formation d'attitude, échange d'opinion, formation d'une représentation sociale etc.). Dans notre modèle, les objets sociaux correspondent aux objets ciblés par les attitudes, c'est à dire un acteur ou un groupe social. Nous notons $SO = Actor \cup SG$ l'ensemble des objets sociaux. Par exemple, un individu peut avoir une attitude vis-à-vis d'un acteur représentant sa société qui l'emploie, un gouvernement ou une Force militaire. Il peut également avoir une attitude vis-à-vis d'une ethnie, des catholiques ou des Américains en général mais pas vis-à-vis d'un autre individu.

3.1.3 Croyances sur des actions

L'élément fondateur de notre modèle d'attitude, conformément au modèle proposé par Fazio et discuté en section 2.3, est la croyance à propos d'action. Ce sont les croyances des individus qui portent sur les actions effectuées par les acteurs (ex : baisse des impôts, opération de patrouille, publicité etc) qui forment les attitudes. Dans nos travaux, les actions n'existent qu'au travers de croyances, que ce soient des actions perçues, des actions crues ou des actions communiquées. En conséquence, nous indiquerons systématiquement l'origine du point de vue de la perception de l'action. Nous notons $a(i)$ l'action crue par l'agent i , c'est-à-dire l'action relatée par la croyance stockée dans la mémoire Blf_i de l'agent i . Dans la suite du document, le terme croyance sera synonyme d'« action » dans le sens où elle fera référence à la *croyance propos de l'action*. Nous notons **Act** l'ensemble des actions. Une action $a \in Act$ est définie par :

$$a = \langle name, actor, coresp, date, bnf, pyf, \sigma \rangle$$

avec :

- **name** le nom de l'action (unique)
- **actor** $\in Actors$ l'acteur à l'origine de l'action
- **coresp** $\in Actors$ un acteur coresponsable de l'action s'il en existe
- **date** $\in \mathbb{N}$ la date d'occurrence de l'action

- $\text{bnf} \subset \text{Ind}$ les agents impactés par l'action (ou bénéficiaires de l'action) qui appartiennent tous à un même groupe social¹.
- $\text{pyf} \in \mathbb{R}$ la valeur du bénéfice obtenu par le bénéficiaire, négative lorsque l'impact est nocif et positive s'il est bénéfique
- σ la crédibilité de la source de l'information avec
 $\sigma \in \Sigma = \{\sigma_1, \sigma_2, \dots, \sigma_s\}$ ensemble ordonné par la relation $\sigma_1 \prec \sigma_2 \prec \dots \prec \sigma_s$ qui sera utilisée lors de la révision de croyances (voir section 3.3)

Dans la suite du document, i et j feront référence à des individus (*c.-d.* des agents), a désignera une action.

EGALITÉ ET TYPE D'ACTION Dans notre modèle, nous avons besoin de définir deux types d'égalité d'actions : l'une dite « faible » et l'autre « forte ». Nous distinguons ces deux types d'égalité afin que les agents puissent avoir un protocole différent pour l'évaluation des actions d'un même type (ex : « les patrouilles sont généralement rassurantes ») et la révision de croyance qui porte sur une occurrence d'action en particulier (ex : « la patrouille d'hier m'a rassuré »). Deux actions sont dites *similaires* si elles ont le même nom, acteur et groupe social impactés (elles font référence à un même type d'action). Par exemple, deux patrouilles distinctes effectuées à des dates différentes par la Force Bleue dans la même zone sont considérées comme étant similaires. Si les deux croyances d'action ont aussi la même date d'occurrence, elles sont dites *équivalentes* (les deux actions font référence à la même occurrence d'action).

Nous notons $a_1 \sim a_2$ si les actions a_1 et a_2 sont similaires et $a_1 \equiv a_2$ si elles sont équivalentes.

Nous appelons *type d'action* la méta action qui englobe des actions similaires : $\text{at} = \langle \text{name}, \text{actor}, \text{coresp}, \text{bnf} \rangle$. Pour une action a donnée, son type d'action pour l'agent i est défini par $\text{at}(i, a) = \{a'\}$, $a' \in \text{Blf}_i$ tel que $a' \sim a$ est l'ensemble des actions similaires à l'action a selon l'individu i . Par extension nous notons $\text{at}(i)$ la liste des types d'action connue par l'agent i .

ANCIENNES OCCURRENCES ET NOMBRE D'OCCURRENCES Un individu i associe à chaque action a :

- sa précédente occurrence similaire $a_{\text{old}}(a, i) = a' \in \text{at}(a, i)$ tel que $a'.\text{date} < a.\text{date} \wedge \forall a'' \in \text{at}(a, i) \setminus \{a\}, a''.\text{date} \leq a'.\text{date}$
- sa précédente occurrence similaire vécue personnellement par i $a_{\text{perso}}(a, i) = a' \in \text{at}(a, i)$ tel que $a'.\text{date} \leq a.\text{date} \wedge i \in a'.\text{bnf} \wedge \forall a'' \in \text{at}(a, i) \wedge i \in a''.\text{bnf}, a''.\text{date} \leq a'.\text{date}$
- le nombre d'occurrences d'actions similaires
 $\text{nbOcc}(a, i) = \text{card}(\text{at}(a, i))$
- le nombre d'occurrences d'actions similaires par groupe social bénéficiaire
 $\text{nbOcc}_{\text{SG}}(a, \text{sogr}) = \text{card}(\{a' \in \text{at}(a, i), a'.\text{bnf.sg} = \text{sogr}\})$

Toutes ces informations sont basées sur la connaissance locale de l'agent i .

1. Dans le cas où l'action impacte plusieurs groupes sociaux, il suffit de représenter plusieurs actions, une par groupe social

3.1.4 Attitudes

Pour un individu donné, chaque *objet social* est associé à une *attitude*. Nous définissons deux fonctions selon la nature de l'objet social (groupe social ou acteur) :

- $\text{attSG} : \text{Ind} \times \text{SG} \rightarrow \mathbb{R}$ donnant la valeur de l'attitude $\text{attSG}(i, \text{sg})$ de l'individu i vis-à-vis du groupe social $\text{sg} \in \text{SG}$, positive si favorable et négative si défavorable. Cette valeur importe peu : elle ne varie pas au cours de la simulation et seul son signe est utilisé dans le modèle lors de l'évaluation de l'impact d'une action (cf. section 3.6.1) ;
- $\text{att} : \text{Ind} \times \text{Actor} \rightarrow \mathcal{R}$ donnant la valeur de l'attitude $\text{att}(i, \text{actor})$ de l'individu i vis-à-vis de l'acteur $\text{actor} \in \text{Actor}$, positive si favorable et négative si défavorable. Cette valeur représente la raison d'être de ce modèle car elle en est sa principale sortie. Elle est notamment utilisée lors de l'évaluation d'une action impactant un acteur ;

Concernant les groupes sociaux, nous considérons un tableau aTable de dimension $|\text{SG}| \times |\text{SG}|$ définissant *a priori* et de manière figée pour l'ensemble de la simulation les attitudes d'un individu i d'un groupe social $i.\text{sg}$ vis-à-vis d'un groupe social sg' (et, par extension, des individus appartenant à ce groupe) :

$$\text{attSG}(i, \text{sg}') = \text{aTable}(i.\text{sg}, \text{sg}')$$

Par extension, nous définirons $\text{attSG}(i, j) = \text{attSG}(i, j.\text{sg})$ l'attitude de l'agent i vis-à-vis de l'agent j .

Ci-dessous, la figure 8 donnant un exemple de configuration d'attitudes inter-groupes, qui sera réutilisé dans notre évaluation au chapitre 4.

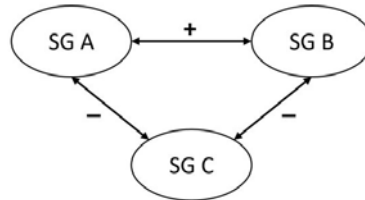


FIGURE 8 – Exemple de configuration d'attitudes inter-groupes sociaux

Le cœur de notre modèle consiste à simuler la dynamique des attitudes de la population vis-à-vis des forces, à partir des perceptions et interprétations des actions. La valeur de $\text{att}(i, f)$ est définie dans la section 3.6 ci-après, comme résultat des communications d'actions et de leurs évaluations par le modèle cognitif des agents.

3.1.5 Messages

Au cours de la simulation, les forces peuvent communiquer sur leurs actions auprès de la population. De plus, les individus relayent les informations acquises (soit parce qu'ils ont observé directement une action, soit parce qu'ils sont sensibles aux communications des forces), propageant ainsi l'information dans la population. Nous présentons le mécanisme de propagation dans la section 3.5. Soit Msg l'ensemble des messages, $m \in \text{Msg}$ est défini par :

$$m = \langle \text{src}, \text{date}, a, \text{Adr} \rangle$$

avec :

- $\text{src} \in \text{SO}$ l'agent émetteur du message
- $\text{date} \in \mathbb{N}$ la date de réception du message
- $a \in \text{Blf}_{m.\text{src}}$ l'action relatée par le message
- $\text{Adr} \subset \text{Ind}$ les destinataires du message

Dans les prochaines sections, nous décrivons les mécanismes cognitifs permettant aux agents de réviser leurs croyances sur les actions des acteurs, de les communiquer et de calculer leurs attitudes.

3.2 ACQUISITION DE CROYANCE

Les actions peuvent être perçues selon trois canaux :

1. Perception directe : l'agent est sujet ou témoin direct de l'action (ex : l'ONU distribue des vivres dans un village et l'agent est membre du village ou se trouvait à proximité lorsque l'action s'est déroulée) ;
2. Communication des acteurs : les acteurs diffusent un message à propos d'une action à la population et l'agent fait partie des destinataires (l'agent fait partie de l'audimat) ;
3. Communication inter-individuelle : l'agent reçoit l'information par un message émis par un autre individu (conversation quotidienne) ;

Les cas 1 et 2 sont scriptés : ils sont définis par chaque actor.actionList , $\forall \text{actor} \in \text{Actor}$. La valeur de crédibilité de l'impact d'une perception directe est fixée à σ_s (la valeur de crédibilité maximale au sens de la relation d'ordre définie dans la section 3.1.3) et celle d'une action de communication des acteurs à σ_1 (la valeur minimale dans notre modèle).

Dans le cas des communications intra-population (cas 3), la valeur de la crédibilité de l'impact communiqué est exactement la valeur présente dans les croyances de l'émetteur du message : les agents communiquent une action selon la connaissance qu'ils en ont. Comme nous le montrerons ci-après, c'est au récepteur de modifier cette valeur au stade de l'ajout de l'information dans sa propre liste de croyances.

3.3 RÉVISION DE CROYANCE

Lorsqu'un agent acquiert une nouvelle information, soit par perception directe, soit par l'intermédiaire d'une communication provenant d'un acteur ou d'un individu de la population, il révisé ses connaissances. Le processus de révision fait intervenir un tableau de probabilité de révision en fonction des crédibilités des informations confrontées. Cette approche est inspirée du modèle COBAN proposé par Thiriot (2009). Il est important de noter que nous avons un ensemble fini et ordonné de valeurs de crédibilité avec un minimum (σ_1) et un maximum (σ_s). Ainsi nous avons $\sigma \in \Sigma = \{\sigma_1, \sigma_2, \dots, \sigma_s\}$ avec $\sigma_s \succ \sigma_{s-1} \succ \dots \succ \sigma_1$.

Considérons qu'un agent i ait acquis une information a_{in} entrante, nous notons $\sigma_{in} = a_{in}.\sigma$ la crédibilité de cette information. Le processus de révision de la base de connaissances Blf_i qui en découle est défini comme suit :

- Si l'information \mathbf{a}_{in} n'existe pas encore dans Blf_i (c.-d. il n'existe pas d'action *équivalente* avec le même nom, acteur et groupe social bénéficiaire et date), l'agent ajoute l'action dans sa base de croyance. Sa valeur de crédibilité est définie par :

$$\mathbf{a}_{in}.\sigma = \begin{cases} \sigma_s & \text{si observation directe} \\ \max(\sigma_{in-1}, \sigma_1) & \text{autrement} \end{cases} \quad (4)$$

Lorsque ce n'est pas une observation directe, la crédibilité est décrétementée d'un niveau : lorsqu'un ami me dit qu'il a observé un événement (σ_s pour lui), cela reste une observation indirecte pour moi (σ_{s-1}).

- S'il existe une action \mathbf{a}_{ex} (avec une crédibilité σ_{in}) *équivalente* à \mathbf{a}_{in} dans Blf_i ($\mathbf{a}_{ex} \equiv \mathbf{a}_{in}$) mais avec des payoff différents, deux situations sont possibles :

1. *Cas compatible, renforcement*

Si $|\mathbf{a}_{ex}.\text{pyf} - \mathbf{a}_{in}.\text{pyf}| \leq \varepsilon$ avec ε une limite fixée², les croyances sont considérées comme étant suffisamment proches pour faire référence à la même occurrence d'action. Dans ce cas, la valeur de crédibilité $\mathbf{a}_{ex}.\sigma$ peut être mise à jour :

$$\mathbf{a}_{ex}.\sigma = \max(\sigma_{ex}, \sigma_{in-1}) \quad (5)$$

Par exemple, considérons que la crédibilité d'une action existante est de $\sigma_{ex} = \sigma_1$ chez le sujet. Un ami lui parle d'une action équivalente qu'il croit à σ_{in} .

Premier cas, l'ami est témoin direct de l'action ($\sigma_{in} = \sigma_s$ pour lui), la crédibilité σ_{ex} qu'accorde le sujet sera renforcée à σ_{s-1} .

Second cas, cette fois l'ami n'a pas été témoin direct mais a entendu parler d'une action ($in < s$). Si $in - 1 > ex$, alors la crédibilité σ_{ex} chez le sujet sera augmentée de σ_1 à σ_{in-1} . Sinon, elle restera inchangée à σ_1 .

2. *Cas incompatible, révision :*

Si $|\mathbf{a}_{ex}.\text{pyf} - \mathbf{a}_{in}.\text{pyf}| > \varepsilon$, les croyances sont incompatibles. Dans ce cas, il y a une probabilité pour que l'agent remplace sa croyance \mathbf{a}_{ex} par \mathbf{a}_{in} . Cela va dépendre de la probabilité de révision donnée par la Table 2, où $\pi_i \in [0, 1]$ $i \in \{0, \dots, p\}$ sont des paramètres du modèle pour la révision basée sur la crédibilité³. Nous effectuons un tirage aléatoire uniforme sur $[0, 1]$: si cette valeur est plus grande que la probabilité de révision, la croyance \mathbf{a}_{ex} est remplacée par \mathbf{a}_{in} .

TABLE DE RÉVISION DE CROYANCES La table 1 décrit la structure d'une table de révision de croyances en fonction des valeurs de crédibilité respectives aux informations confrontées. Nous faisons l'hypothèse que les probabilités de révision sont égales pour un même écart de crédibilité : $\forall C \in \mathbb{N}, \exists! \pi_k$ tel que $\forall i, j$ tel que $i - j = C, p_r(\sigma_i, \sigma_j) = \pi_k$ où $\pi_k \in [0, 1]$. Les diagonales de constantes sont dues à cette formule. De plus, à des fins de simplifications, nous admettons une

2. Dans notre implémentation, nous avons fixé cette valeur à zéro (voir la section 3.7)

3. Notons que σ_3 versus σ_3 est impossible car un agent ne peut pas témoigner deux fois d'une même occurrence d'actions (elles auront au moins des dates différentes).

effet, ces deux valeurs sont équivalentes en science cognitive : les informations que l'on communique le plus sont celles qui sont les plus accessibles dans nos mémoires [Dimulescu et Dessalles \(2009\)](#); [Chater et Vitányi \(2003\)](#). La Théorie de la Simplicité de Dessalles ([Dimulescu et Dessalles \(2009\)](#)) propose de définir ces intérêts d'une information selon deux composantes : la réponse émotionnelle E du sujet face au stimulus et à la surprise S ressentie. Elle met en avant une formule de la forme suivante :

$$I = E + S \quad (6)$$

Nous adaptons cette formule à notre modèle en introduisant la notion d'action et un paramètre de calibration $\alpha \in [0, 1]$ permettant d'équilibrer les deux composantes selon le cas d'étude. Notre valeur d'accessibilité, d'une action \mathbf{a} pour un individu \mathbf{i} est ainsi donnée par :

$$I(\mathbf{i}, \mathbf{a}) = \alpha \times E(\mathbf{a}) + (1 - \alpha) \times S(\mathbf{i}, \mathbf{a}) \quad (7)$$

Il est à noter qu'aussi bien I que $I(\mathbf{i}, \mathbf{a})$ ne peuvent pas être négatifs : l'intérêt d'une information inintéressante est au minimum nulle [Dimulescu et Dessalles \(2009\)](#). Ainsi, dans le cas où $\alpha \times E(\mathbf{a}) < (1 - \alpha) \times S(\mathbf{i}, \mathbf{a})$, $I(\mathbf{i}, \mathbf{a}) = 0$.

3.4.1 *L intensité émotionnelle*

L'intensité émotionnelle E correspond à l'amplitude émotionnelle non-nulle vécue par un individu lorsqu'il est exposé à un événement. [Dimulescu et Dessalles \(2009\)](#) montrent que lorsque le stimulus n'est pas borné, l'intensité émotionnelle suit une loi logarithmique conformément à la théorie du stimuli de Weber-Fechner [Fechner \(1832\)](#) (dans notre cas, la valeur du stimulus correspond au payoff de l'action) :

$$E(\mathbf{a}) = \log \left(1 + \frac{|\mathbf{a} \cdot \text{pyf}|}{\xi} \right) \quad (8)$$

Où $\xi \in [0, 1]$ est un paramètre de sensibilité modulant l'intensité de la réponse émotionnelle. Il est à noter que selon cette formule, une réponse émotionnelle ne peut pas être négative.

3.4.2 *Surprise*

La théorie de la simplicité permet de calculer la surprise selon plusieurs dimensions (ex : la distance géographique, la récence temporelle etc.). La surprise est la somme de ces dimensions. Dans notre modèle nous nous référons à deux dimensions, la distance temporelle et la distance sociale.

$$S = \beta \times S_{\text{time}} + (1 - \beta) \times S_{\text{social}} \quad (9)$$

avec $\beta \in [0, 1]$ un paramètre permettant d'équilibrer les deux dimensions selon le cas d'étude.

Suivant la théorie de Dessalles, la surprise ressentie par un individu lorsqu'il est exposé à un événement dérive d'un sentiment d'inattendu U_{raw} (ex : « Il est étonnant qu'un Taliban vienne en aide à un civil »). Ce sentiment est atténué par un niveau d'inattendu personnel de référence U_{perso} qui se base sur une

expérience personnelle (ex : « Mais j'ai moi-même déjà été aidé par un Taliban auparavant. ») :

$$S_x = U_{\text{raw}}^x - U_{\text{perso}}^x \quad (10)$$

où x la dimension de la surprise (temporelle ou sociale dans notre cas).

En conséquence, nous sommes amenés à quatre valeurs d'inattendu : $U_{\text{raw}}^{\text{time}}$, $U_{\text{raw}}^{\text{social}}$, $U_{\text{perso}}^{\text{time}}$ et $U_{\text{perso}}^{\text{social}}$. Les prochains paragraphes présentent les formules de chacun de ces éléments.

Dans tous les cas, l'inattendu d'un événement (une action dans notre modèle) est défini par le contraste entre sa complexité de génération et sa complexité de description :

$$U^x = C_w^x - C_d^x \quad (11)$$

avec x la dimension. La complexité de génération C_w définit le niveau auquel l'événement était anticipé par l'individu sur la base de ses connaissances courantes. La complexité de description C_d doit être interprétée dans le cadre de la théorie de l'information de [Shannon et Weaver \(1949\)](#), *c.-d.* la taille du plus petit algorithme qui puisse décrire l'événement. De plus, un inattendu est borné à zéro, si $U^x < 0$ alors la valeur est ramenée à $U^x = 0$: en effet, un inattendu ne peut être négatif en psychologie cognitive [Dimulescu et Dessalles \(2009\)](#). Ceci implique qu'une surprise ne peut pas être négative. Ainsi, selon la formule 6, la surprise ne peut pas réduire l'intérêt d'une information. L'intérêt est au minimum égal à la réponse émotionnelle vis-à-vis du payoff.

L'INATTENDU TEMPOREL BRUT ($U_{\text{raw}}^{\text{time}}$) La complexité de génération temporelle fait référence à la probabilité qu'une action survienne à un instant donné. Cette notion peut être interprétée comme l'écart de temps usuel entre deux occurrences d'une action : plus l'action est rare, plus cet écart est grand, moins elle est probable, plus elle est inattendue. Par conséquent, nous définissons l'écart temporel usuel selon la distance entre la date de l'occurrence de l'action *courante* (celle dont on évalue l'intérêt) `a.date` et la date de sa précédente occurrence `a_ol.d.date` :

$$C_w^{\text{time}} = \log_2(\text{a.date} - \text{a_ol.d.date}) \quad (12)$$

La complexité de description temporelle correspond simplement au temps écoulé entre l'action et la date courante t :

$$C_d^{\text{time}} = \log_2(t - \text{a.date}) \quad (13)$$

Il existe un cas d'exception dans le calcul de la surprise temporelle : dans le cas où $\text{a_ol.d} = \emptyset$, c'est-à-dire que l'agent est confronté pour la première fois à un événement, en conséquence la formule n'est pas directement applicable. En effet, si l'on considère que $\text{a_ol.d.date} = 0$, d'une part ce serait erroné car l'événement n'a pas eu lieu à la date 0, d'autre part la surprise temporelle ne serait relative que de la date courante. Pour ces raisons, nous sommes contraints de définir une valeur constante $UT_0 \in \mathbb{R}^+$ définissant la complexité de génération d'un événement qui survient pour la première fois.

Ainsi, l'inattendu sur la dimension temporelle est obtenu par la formule suivante :

$$U_{raw}^{time}(a) = \begin{cases} \log_2(a.date - a_{old}.date) - \log_2(t - a.date) & \text{si } a_{old} \neq \emptyset \\ UT_0 - \log_2(t - a.date) & \text{sinon} \end{cases} \quad (14)$$

En simplifiant la différence entre les logarithmes nous obtenons :

$$U_{raw}^{time}(a) = \begin{cases} \log_2\left(\frac{a.date - a_{old}.date}{t - a.date}\right) & \text{si } a_{old} \neq \emptyset \\ UT_0 - \log_2(t - a.date) & \text{sinon} \end{cases} \quad (15)$$

INATTENDU SOCIAL BRUT (U_{raw}^{social}) La complexité sociale de génération fait référence à la probabilité qu'une action affecte un bénéficiaire appartenant à un certain groupe social (ex : « Il est rare qu'un Pashtun soit victime d'une attaque talibanne »). Un agent i calcul cette complexité selon :

$$C_w^{social} = -\log_2\left(\frac{nbOcc_{SG}(a, i.sg)}{nbOcc(a)}\right) \quad (16)$$

où $nbOcc_{SG}(a, i.sg)$ le nombre d'occurrences de l'action a dont le bénéficiaire est membre du même groupe social $i.sg$ que l'agent i .

La complexité de description C_d^{social} correspondant à la distance sociale entre l'individu et le bénéficiaire de l'action dépend de deux facteurs : la distance dans le réseau social et le degré moyen du graphe (le nombre de voisins moyen par individu). En effet, plus ce degré est élevé, plus l'information est compliquée à décrire à chaque pas dans le graphe (en terme de théorie de l'information) sachant que cette relation est linéaire. Cependant, la distance dans le graphe a un impact exponentiel sur la complexité de la distance sociale (car elle nécessite de décrire toutes les possibilités de voisins pour chaque nœud). Nous prenons la distance minimale entre l'individu et les bénéficiaires de l'action. Par exemple, dans les journaux télévisés, lorsqu'il y a eu un attentat à l'étranger, les journalistes mentionnent toujours le nombre de français qui ont été blessés ou tués. Cela provient du fait que nous nous référons aux personnes dont nous sommes plus proches. Si de surplus, l'attentat a atteint un ami, l'intérêt de l'information sera d'autant plus important.

$$C_d^{social} = \log_2(v^d) \quad (17)$$

où v le degré moyen du graphe et d la distance la plus courte entre i et les membres de $a.bnf$ dans le graphe.

Par conséquent, nous obtenons cette formule :

$$U_{raw}^{social}(a) = -\log_2\left(\frac{nbOcc_{SG}(a, i.sg)}{nbOcc(a)}\right) - \log_2(v^d) \quad (18)$$

INATTENDUS TEMPOREL ET SOCIAL L'inattendu personnel se base sur la dernière occurrence de l'action a qui a personnellement affecté l'individu i , $c.$ - $d.$ la dernière occurrence de l'action pour laquelle l'agent i appartient lui-même à l'ensemble $a.bnf$. Nous notons a_{perso} cette occurrence particulière (cf. section 3.1.3).

Le mode de calcul de cette valeur d'inattendu est le même que pour les précédents calculs, excepté l'identification des occurrences dans la base de croyance qui se limite aux actions dont i est bénéficiaire :

$$u_{\text{perso}}^{\text{time}}(a) = \log_2 \left(\frac{a.\text{date} - a_{\text{perso}}.\text{date}}{t - a.\text{date}} \right) \quad (19)$$

$$u_{\text{perso}}^{\text{social}}(a) = -\log_2 \left(\frac{\text{nbOcc}_{\text{SG}}(a_{\text{perso}}, i)}{\text{nbOcc}(a)} \right) \quad (20)$$

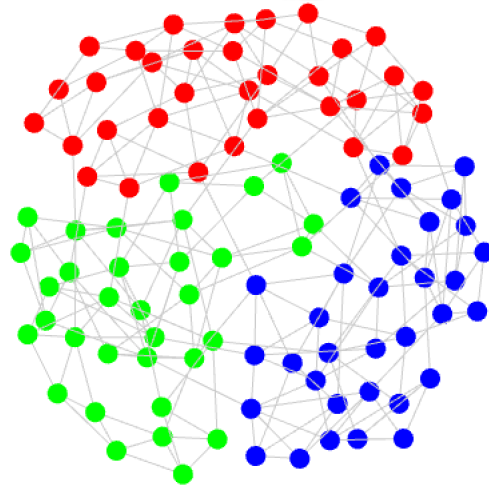
où $a_{\text{perso}}.\text{date}$ correspond à la date de la dernière occurrence vécue par le sujet i et $\text{nbOcc}_{\text{SG}}(a_{\text{perso}}, i)$ le nombre d'occurrences de cette action. Dans le cas où l'individu n'a jamais vécu l'action, son inattendu personnel est nul, $u_{\text{perso}}^{\text{time}} = 0$. En effet, contrairement à l'inattendu temporel brut auquel cas nous avons besoin de définir une constante UT_0 pour provoquer une surprise lorsque l'agent témoigne pour un événement pour la première fois, ici l'agent se réfère à son expérience personnelle. S'il n'a jamais vécu l'action, il n'a pas de raison à atténuer la surprise temporelle avec un sentiment « d'attendu ».

3.5 COMMUNICATION INTRA-POPULATION

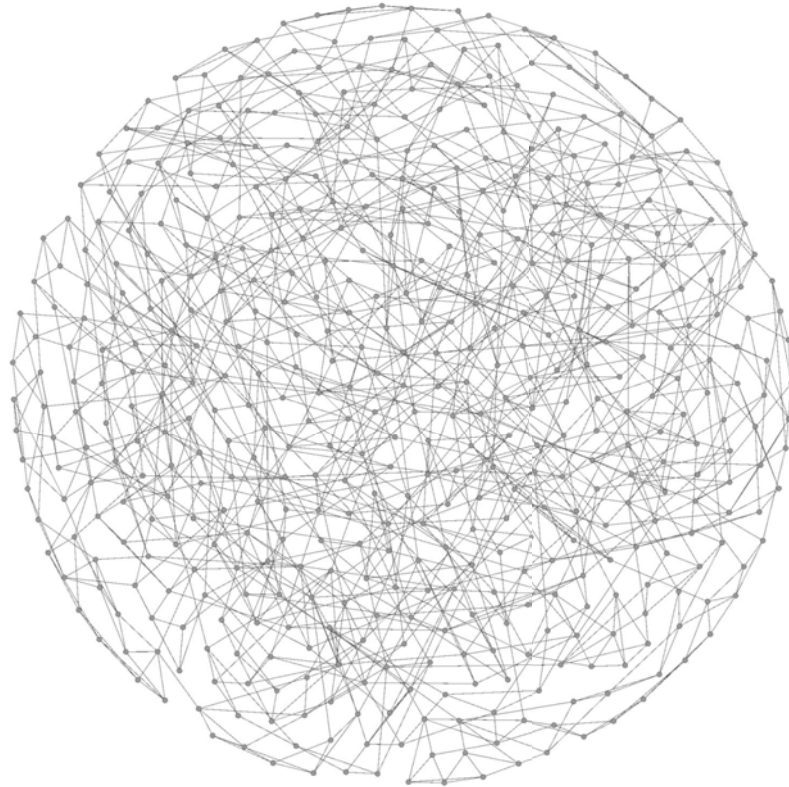
La communication intra-population repose sur la liste des contacts des individus ($i.\text{Cnt}$), c'est à dire ses voisins dans le graphe social. Un agent peut communiquer ses croyances (information sur des actions) à ses contacts. Nous décrirons dans un premier temps la manière dont les listes de contacts sont définies dans notre modèle. Puis nous montrerons comment et quand les agents sélectionnent les croyances qu'ils veulent communiquer.

3.5.1 Topologie du Réseau

La liste de contacts des agents est définie par la topologie du réseau social dans lequel se propage l'information au travers des interactions. La majorité des travaux sur la dynamique des attitudes utilise un réseau d'interactions pour définir les liens entre agents, leurs fréquences et leurs intensités (ex. selon les affinités entre individus). Dans notre modèle, nous considérons les liens comme étant des communications potentielles entre deux individus, indépendamment de leur groupes sociaux ou autres affinités. Cependant, contrairement à ce qui est effectué dans certains travaux (ex : [Nowak et collab. \(1990\)](#)) où les agents ne prennent en compte aucune topologie de réseau, il semble important de disposer d'un réseau tenant compte du contexte de simulation avec différents groupes sociaux : les agents communiquent plus à l'intérieur d'un même groupe social. La tendance de ces dernières années dans ce domaine de recherche est le recours à des réseaux sociaux complexes comme les small-world [Milgram \(1967\)](#) ou encore les graphes à échelle invariante [Wang et Chen \(2003\)](#) permettant d'explorer la dynamique des modèles. Quelques rares recherches utilisent des générateurs de populations synthétiques plus sophistiqués qui intègrent des données socio-démographiques quantitatives et qualitatives (ex. statistique sociale, structure sociale, données économiques etc.) [Gargiulo et collab. \(2010\)](#), [Thiriou et Kant \(2008\)](#). Dans notre cas précis, compte tenu du manque de données sur l'environnement humain pouvant nous intéresser directement (zones de crises), nous construisons notre environnement social sur la base de réseaux small-world. La



(a) Réseau de type small-world pour trois groupes sociaux



(b) Réseau de type small-world pour un groupe social de 500 individus

FIGURE 9 – Exemples de réseaux sociaux employés dans nos simulations

figure 9a offre l'exemple d'un tel réseau pour trois groupes sociaux. Il s'agit de trois small-world interconnectés de manière artificielle. A des fins de simplification, la topologie du réseau est un paramètre de simulation et restera statique au cours d'une simulation.

3.5.2 Algorithme de communication

Dans cette section nous définissons un algorithme permettant aux individus de sélectionner certaines informations nouvellement acquises durant le pas de temps courant à transmettre à certains de leurs voisins. Nous exposons dans un premier temps les spécifications que nous attendons d'un tel algorithme avant de présenter notre algorithme.

SPÉCIFICATIONS Nous voulons que la probabilité de relayage d'une information nouvellement acquise par un individu :

- i soit en fonction de l'intérêt de l'information (cf. section 3.4) ;
- ii soit relative à une information référente d'un intérêt élevé. L'idée ici est qu'un individu va choisir l'information à communiquer en fonction de l'information la « plus intéressante depuis un certains temps ». Cette notion sera approfondie dans l'algorithme ;
- iii soit croissante monotone avec l'intérêt de l'information (plus une information est intéressante, plus elle a de chance à être communiquée) ;
- iv ne soit pas linéaire, en effet les réponses émotionnelles ne sont pas de forme linéaire ;
- v exige que l'intérêt de l'information soit proche de celle de l'information référente : les individus sont exigeants, ils ne communiquent pas les événements d'un intérêt trop faible ou moyen, ils ne parlent que de faits intéressants ;
- vi soit non nulle même pour des informations à intérêts nuls : les individus étant non rationnels, ils peuvent parler de faits anodins avec une très faible probabilité.

ALGORITHME La sélection des informations à relayer pour chaque agent i est définie par le processus suivant :

1. Déterminer de $I_{\max}(i, t)$, la valeur de l'information la plus intéressante depuis $n\text{Step} \in \mathbb{N}$ pas de temps, avec $n\text{Step}$ un paramètre de simulation permettant de simuler la dynamique de l'intérêt référent : $I_{\max}(i, t) = \max_{a.date \in [t-n\text{Step}, t]}(I(i, a))$.

Par exemple, durant une période qui suit un attentat important, les gens et la presse ont tendance à moins communiquer sur les autres faits d'actualité. Mais cette durée est limitée : au bout d'un certain temps, l'information référente est mise à jour. Ainsi, dans notre modèle, l'information perd en intérêt (1) au fur des communications avec le mécanisme de révision de croyance (crédibilité décroissante à chaque transmission) et (2) au fur du temps avec cette fenêtre glissante.

2. La probabilité de transmission d'une information dépend de l'écart entre son intérêt et celui de l'information référente, c'est-à-dire $I(i, a) - I_{\max}(i, t)$. Pour chaque voisin de l'agent i , la probabilité $p_s(i, a, t)$ que i transmette l'information a à ce voisin est calculée par :

$$p_s(i, a, t) = \begin{cases} 2^{I(i, a) - I_{\max}(i, t)} & \text{si } 2^{I(i, a) - I_{\max}(i, t)} > T_{\text{com}} \\ 0 & \text{autrement} \end{cases} \quad (21)$$

où $T_{\text{com}} \in [0, 1]$ un seuil de communication qui permet de filtrer les informations d'un intérêt mineur. Il représente le niveau d'intérêt des autres informations du quotidien qui viennent aussi nourrir la conversation. Une autre interprétation possible de ce seuil serait que les individus sont plus ou moins bavards : plus T_{com} est bas, plus il y a de messages qui passent ce filtre, plus les agents sont bavards. Ce seuil peut très bien être ignoré ($T_{\text{com}} = 0$) par le modélisateur s'il considère que dans son cas d'application, même les informations peu intéressantes doivent avoir une chance d'être échangées.

Le graphique présenté dans la figure 10 trace la fonction $2^{I(i,a) - I_{\text{max}}(i,t)}$ et la compare à une fonction linéaire $\frac{I(i,a)}{I_{\text{max}}(i,t)}$. La ligne en pointillés correspond à un seuil $T_{\text{com}} = 0.15$, valeur calibrée dans le chapitre 5. Nous pouvons observer qu'une information doit au moins avoir un intérêt supérieur à la moitié de l'intérêt maximal de référence.

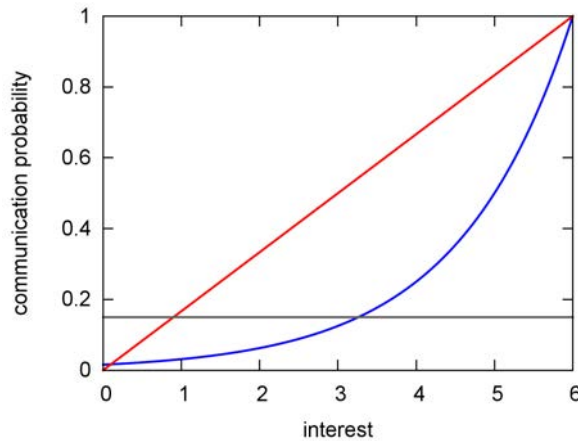


FIGURE 10 – Fonction de probabilité de communication

Dans un futur proche, nous comparerons l'effet de différentes fonctions de probabilité sur la propagation de l'information au sein de la population afin d'appuyer notre choix. Parmi les fonctions possibles, nous pensons au rapport linéaire $\frac{I(i,a)}{I_{\text{max}}(i,t)}$ ou encore son carré $\left(\frac{I(i,a)}{I_{\text{max}}(i,t)}\right)^2$.

REMARQUES D'autres alternatives à ce modèle ont été envisagées lors de sa construction. L'une des difficultés a été de définir I_{max} , la valeur d'intérêt référente à laquelle les intérêts des autres informations sont comparées. Dans un premier temps, cette valeur a été suggérée comme étant locale ou d'une actualité déjà obsolète : les individus se basent sur les faits qu'ils ont observés le jour même pour sélectionner les informations à transmettre. Cependant, cette solution ne permet pas d'obtenir une saturation de la communication par un événement important relaté pendant plusieurs jours. Afin d'y remédier, nous avons essayé dans un second temps un I_{max} global à toute la simulation : les individus comparent les événements à celui qui a été le plus marquant de leur vie. Cette deuxième option a montré rapidement ses limites : avec une action importante en début de simulation, les agents arrêtaient rapidement de communiquer sur les faits postérieurs. Le compromis qui consiste à prendre un I_{max} sur une fenêtre glissante permet de reproduire le comportement voulu au prix d'un paramètre de simulation.

La seconde difficulté a consisté à trouver une fonction de probabilité permettant à la fois aux agents de communiquer intensivement sur les informations fortement intéressantes et de limiter les informations relevant de la banalité. Nous avons envisagé dans une version antérieure de procéder à un tirage aléatoire uniforme sur l'intervalle $[0, \frac{I(i,a)}{I_{\max}(i,t)}]$. Cette solution permettait bien de filtrer les informations banales, cependant les plus importantes ($\frac{I(i,a)}{I_{\max}(i,t)} = 1$) n'avaient qu'une chance sur deux d'être transmises. C'est pourquoi nous sommes restés sur un tirage uniforme avec un seuil T_{com} .

Il est à noter que la communication d'action n'introduit pas de biais dans la description de l'événement, mais seulement dans son degré de crédibilité. En effet, le payoff transmis est toujours égal au payoff de l'action initiale. Par contre, chaque agent, lorsqu'il reçoit l'information depuis un autre agent, déduit sa valeur de crédibilité. Cette crédibilité est importante car c'est elle qui conduit à la révision de croyance (cf. section 3.3).

L'idée proposée ici est qu'une information provenant de l'oncle de l'épouse d'un ami est moins crédible que si elle émanait directement de cet ami. La décroissance de la crédibilité permet notamment de limiter la portée du message.

3.6 CONSTRUCTION D'UNE ATTITUDE

Chaque acquisition d'information au sujet d'une action a réalisée par un acteur $a.actor$ résulte en une modification de croyance dans $Blf(i)$ (insertion d'une nouvelle croyance, révision d'un payoff ou mise à jour du niveau de crédibilité d'une information). L'individu i va alors réviser son attitude envers $a.actor$ en fonction de son nouvel état cognitif.

Notre modèle de construction d'attitude est fondé sur des évaluations de croyances associées à un objet, tel que proposé par Fazio (2007). Suivant ce modèle, une attitude vis-à-vis d'un objet social est calculée à partir d'un ensemble de croyances qui lui sont liées. Chacune de ces croyances sont évaluées du point de vue de l'individu afin d'en estimer leur valeur bénéfique ou nocive. De plus, les croyances ont chacune un certain degré d'accessibilité en mémoire qui pondère l'importance de leur évaluation correspondante. Enfin, l'agrégation de tous ces éléments donne la valeur de l'attitude. La figure 11 illustre la représentation d'une attitude conformément à ce modèle.

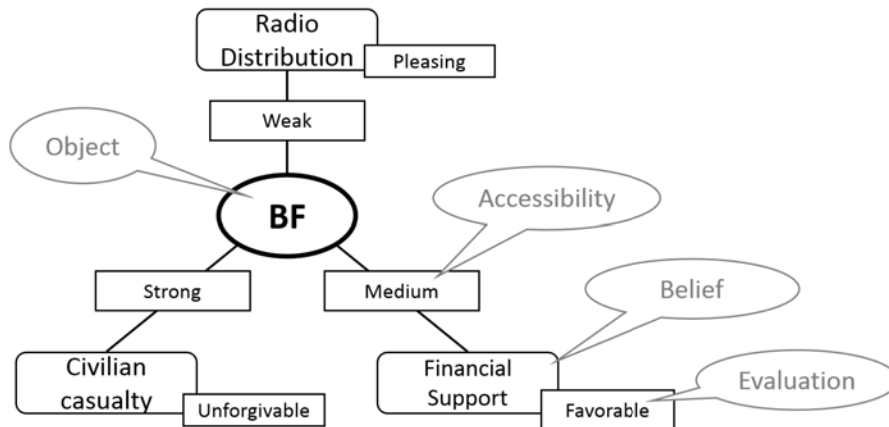


FIGURE 11 – Exemple d'une attitude dans Fazio, 2007.

Dans les prochains paragraphes, nous décrirons chacune de ces étapes de calcul d'une attitude en commençant par l'évaluation du bénéfice subjectif d'une action, puis le calcul de valeur d'accessibilité et enfin l'agrégation finale qui donne la valeur de l'attitude. Nous montrerons aussi comment la notion de co-responsabilité est prise en compte.

3.6.1 *Evaluation du bénéfice d'une action*

Le bénéfice d'une action a est déterminé de manière subjective, c.-à-d. respectivement aux attitudes et croyances de l'agent i . Nous appliquons pour cela le modèle d'évaluation proposé par Fishbein and Ajzen [Ajzen et Fishbein \(2005\)](#). Ce modèle combine le payoff de l'action sur son bénéficiaire impacté et l'attitude de l'individu vis-à-vis de ce bénéficiaire. Ainsi, si un individu juge une action bénéfique pour lui ou pour un « ami » (attitude positive), le bénéfice global sera positif. Inversement, si l'action est bénéfique à un « ennemi » (attitude négative), l'action sera évaluée avec une valeur négative. Dans notre modèle, nous utilisons uniquement le signe de l'attitude vis-à-vis du bénéficiaire. Certains modèles (ex. [Ortony et collab. \(1990\)](#)) proposent d'utiliser aussi la valeur de l'attitude pour pondérer l'évaluation. Mais la notion de distance sociale envers le bénéficiaire est déjà intégrée dans le calcul de la surprise dans notre modèle. De plus, la notion d'amplitude d'impact est aussi incluse dans le concept d'intensité émotionnelle de l'intérêt qui, comme nous le verrons en section [3.6.4](#), est aussi utilisé comme un facteur multiplicatif.

Par ailleurs, nous devons distinguer deux cas de figure selon la nature du bénéficiaire. Si le bénéficiaire de l'action est un individu j de la population, nous utilisons le signe de la valeur de l'attitude du groupe inter-social $\text{attSG}(i, j)$ pour l'appliquer au payoff de l'action. Dans le cas où le bénéficiaire est un acteur, le signe appliqué est celui de $\text{att}(i, \text{actor})$.

De plus, dans le cas où l'action impacte plusieurs individus bénéficiaires, l'évaluation doit être d'autant plus amplifiée. Pour ce faire, nous prenons une fonction logarithmique croissante du nombre de bénéficiaires afin de pondérer l'évaluation. En effet, un attentat qui tue 10 civils est plus important que s'il avait fait 1 victime. Pourtant, la différence entre un attentat faisant 110 ou 100 victimes civiles est quasiment nulle. Cela correspond à des phénomènes retranscrits dans la Théorie de la Simplicité qui applique également des fonctions logarithmiques pour représenter l'impact du stimulus ainsi que l'estimation des différents niveaux de surprise.

$$\text{evaluation}(i, a) = \begin{cases} a.\text{pyf} \times \text{sign}(\text{attSG}(i, a.\text{bnf.sg})) \times \log(1 + n) & \text{si } a.\text{bnf} \in \text{Ind} \\ a.\text{pyf} \times \text{sign}(\text{att}(i, a.\text{bnf.sg})) & \text{si } a.\text{bnf} \in \text{Act} \end{cases} \quad (22)$$

$$\text{avec } \text{sign}(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ -1 & \text{sinon} \end{cases} \quad \text{et } n = |a.\text{bnf}|$$

3.6.2 Accessibilité

Comme cela a été présenté dans la section 3.4, l'accessibilité d'une action \mathbf{a} correspond à $I(\mathbf{i}, \mathbf{a})$.

3.6.3 Attitude de Fazio

Une fois que tous les composants du modèle de Fazio ont été calculés, les individus peuvent enfin construire leurs attitudes. Cela correspond à l'accumulation de toutes les évaluations des actions pour un acteur donné, pesées par leurs accessibilités correspondantes (c.-à-d. la valeur de l'intérêt dans notre cas).

Généralement, les modèles de dynamique d'attitudes agrègent les éléments impactants (ex : caractéristiques d'un objet social, opinions ou parties d'opinion) vus comme des entités indépendantes [Urbig et Malitz \(2007\)](#); [Deffuant et collab. \(2000\)](#). Le modèle d'agrégation correspond alors à une somme ou une moyenne des impacts.

Cependant, comme l'ont montré [Kahneman et collab. \(2003\)](#) dans l'approche rétrospective de l'évaluation basée sur la mémoire, lors de l'évaluation de l'impact des événements (plans sociaux, attentats, aides humanitaires, etc.), l'information est mise en perspective avec les précédentes occurrences des actions du même type. Par exemple, lorsqu'un individu témoigne d'une patrouille effectuée par la Force Bleue, il va mettre à jour son évaluation de l'impact des *patrouilles en général* avant de mettre à jour son attitude et non pas directement modifier son attitude envers les Bleus en se limitant à cette unique occurrence de patrouille.

Soit $\mathbf{at}(\mathbf{i}, \mathbf{actor})$ l'ensemble des types d'action effectuées par l'acteur \mathbf{actor} connu par l'agent \mathbf{i} (au sens de la section 3.1.3). Soit $\mathbf{al}(\mathbf{i}, \mathbf{actor}, \mathbf{t})$ la liste des actions de type $\mathbf{t} \in \mathbf{at}(\mathbf{i}, \mathbf{actor})$ connues par l'individu \mathbf{i} et effectué par l'acteur \mathbf{actor} . L'attitude $\mathbf{attitude}(\mathbf{i}, \mathbf{actor})$ de l'individu \mathbf{i} vis-à-vis de \mathbf{actor} est donnée à chaque instant de la simulation par la formule suivante :

$$\mathbf{attitude}(\mathbf{i}, \mathbf{actor}) = \sum_{\mathbf{t} \in \mathbf{at}(\mathbf{i}, \mathbf{actor})} \left(\sum_{\mathbf{a} \in \mathbf{al}(\mathbf{i}, \mathbf{actor}, \mathbf{t})} \left(\frac{\mathbf{evaluation}(\mathbf{i}, \mathbf{a}) \times I(\mathbf{i}, \mathbf{a})}{|\mathbf{al}(\mathbf{i}, \mathbf{actor}, \mathbf{t})|} \right) \right) \quad (23)$$

La première somme correspond à l'accumulation des impacts des types d'action. Par exemple, l'agent \mathbf{i} peut être informé que la Force Bleue effectue des patrouilles, des ravitaillements, des opérations militaires visant les Taliban. La seconde somme sert à évaluer l'impact d'un type d'action en moyennant tous leurs impacts de chacune de leurs occurrences. Par exemple, l'individu \mathbf{i} peut penser que les patrouilles sont généralement bénéfiques pour lui et son entourage, les ravitaillements ne sont pas suffisants mais que les opérations militaires font généralement trop de dégâts.

REMARQUES D'autres alternatives ont été envisagées avant de parvenir à cette forme d'agrégation. Dans un premier temps nous avons employé une moyenne sur les impacts d'actions sans tenir compte de la notion du type d'action. Cependant, cette méthode diluait au fur et à mesure de la simulation l'impact des

nouvelles actions : le nombre d'actions avait un impact direct sur la sensibilité des agents. En conséquence, la variation des attitudes était faussée car l'impact des actions dépendait du nombre total des actions perçues par la population. Afin d'y remédier, nous avons tenté d'employer l'opérateur OWA [Yager \(1988\)](#) (une moyenne pondérée dans laquelle les poids sont ordonnés et de sorte que leur somme fasse 1) afin de donner plus de poids sur les actions récentes. Cependant, cette solution faisait doublon avec le rôle de l'accessibilité que nous calculons en appliquant la théorie de la Simplicité. De plus, même en utilisant un OWA, avec la moyenne un problème subsistait : les attitudes pouvaient baisser avec un stimulus positif (et évalué positivement). Cela vient du fait qu'une moyenne baisse si l'on ajoute un élément inférieur à la moyenne.

Une autre solution a été d'employer une mise à jour de l'attitude de type markovien, c'est à dire de calculer la nouvelle attitude en fonction de l'ancienne valeur d'attitude et d'une valeur d'impact basée sur les événements témoignés au pas de temps courant. Cette approche était inspirée des travaux sur les modèles de confiances bornées (ex : [Deffuant et collab. \(2000\)](#)). Toutefois, avec cette solution, la répétition d'une même action même très faiblement impactante (ex : distribution de goodies) permet d'augmenter l'attitude à l'infini. Cela vient du fait que la somme tend vers l'infini si les éléments cumulés sont positifs.

La solution retenue correspond alors à la combinaison de ces deux approches. Nous moyennons les actions d'un même type, il est alors normal que l'évaluation du type d'action baisse si l'occurrence témoignée est moins performante que la moyenne des anciennes occurrences. De la même manière, il est normal que l'impact des nouvelles occurrences diminue au fil des répétitions de l'action (ex : l'évaluation d'un attentat se fige avec le temps). Ces évaluations d'impacts des types d'action sont cumulées à l'aide d'une somme : plus un acteur diversifie ses actions positives, plus il est aimé, si elles sont négatives l'attitude baisse.

3.6.4 Coresponsabilité

Dans notre cas d'étude, la Force Bleue endosse le rôle de gardien de sécurité, ils sont alors coresponsables de toutes les actions compromettant la sécurité de la population, incluant les attaques des Talibans, selon la perspective de la population. Cela correspond au concept de la « responsabilité de rôle » d'[Hart \(2008\)](#). Plus généralement, les individus ont tendance à attribuer la responsabilité d'une action à d'autres personnes, groupes ou facteurs situationnels en plus du responsable direct [Kelley \(1973\)](#); [Jones et Harris \(1967\)](#).

Ainsi, nous introduisons un mécanisme de coresponsabilité qui permet aux individus d'attribuer une fraction $\rho \in [0, 1]$, paramètre de la simulation, du payoff d'une action à un ou plusieurs coresponsables. Ce mécanisme se produit quand un individu fait face à une action \mathbf{a} :

1. Dans laquelle il intervient un acteur coresponsable : $\mathbf{a}.\text{coresp} \neq \emptyset$;
2. Dont l'impact est négatif : $\mathbf{a}.\text{pyf} < 0$ (c.-à-d. il n'y a pas de coresponsabilité dans le cas d'actions bénéfiques) ;
3. Dont l'évaluation est négative : $\text{evaluation}(\mathbf{i}, \mathbf{a}) < 0$.

Dans ce cas spécifique, l'individu ajoute une croyance \mathbf{a}' avec $\mathbf{a}'.\text{actor} = \mathbf{a}.\text{coresp}$ et

$$\text{evaluation}(\mathbf{i}, \mathbf{a}') = \rho \times \text{evaluation}(\mathbf{i}, \mathbf{a}) \quad (24)$$

Cette nouvelle croyance conduit à une révision de l'attitude envers l'acteur coresponsable, comme pour une croyance « classique ».

3.7 PARAMÈTRES DE SIMULATION

Nous reprenons dans cette section l'ensemble des paramètres de notre modèle. Nous les catégorisons selon la méthode que nous appliquons afin de les déterminer.

PARAMÈTRES DE CALIBRATION Certains paramètres du modèle n'ont d'autre moyen que d'être fixés à travers un processus de calibration sur la base d'un scénario réel. Ce sont des facteurs cognitifs dont les valeurs sont latentes et dont l'influence sur la dynamique d'attitude n'est pas directement perceptible :

- $\alpha \in [0, 1]$ permet d'équilibrer la composante émotionnelle et celle relative à la surprise dans le calcul de l'intérêt d'une information (cf. équation 7) ;
- $\beta \in [0, 1]$ permet d'équilibrer les dimensions temporelle et sociale dans le calcul de la surprise d'une information (cf. équation 9) ;
- $\xi \in]0, 1]$ définit la sensibilité émotionnelle des individus (cf. équation 8) ;
- $\rho \in [0, 1]$ module le taux d'impact attribué à un acteur coresponsable de l'accomplissement d'une action négative (cf. équation 24).

PARAMÈTRES FIXÉS SELON UNE ANALYSE FONCTIONNELLE Cette approche consiste à identifier les valeurs de paramètres en fonction de leurs effets sur le comportement du modèle et celui attendu a priori. Par exemple, pour la variable `Tcom`, l'utilisateur peut avoir des informations sur le taux de communication au sein de la population concernant les événements qui figurent dans le scénario. Dans ce cas, il va étudier l'effet de ce paramètre sur le taux de communication en le faisant varier sur toutes ses valeurs possibles et sélectionner celle qui permet de reproduire le comportement le plus proche.

- `Tcom` $\in [0, 1]$ filtre les messages communiqués selon que leurs valeurs d'intérêt dépasse ce seuil (cf. équation 27) ;
- `nStep` $\in \mathcal{N}$ intervalle temporel en nombre de jours au bout duquel l'intérêt maximal d'une information expire (cf. section 3.5.2) ;
- `credTable`(σ_{in}, σ_{ex}) table définissant les probabilités de révision en fonction de la confrontation des crédibilités de deux actions. Dans nos travaux, considérons trois niveaux de crédibilité et certaines règles sur la forme de la table (voir section 3.3). Par conséquent, nous avons deux valeurs de probabilités inconnues que nous fixons sur la base d'une analyse de sensibilité (cf. section 4.4).

PARAMÈTRES D'ENVIRONNEMENT Certains paramètres doivent être fixés selon les choix du modélisateur qui eux-mêmes dépendent des cas d'études :

- $\epsilon \in \mathcal{R}$ correspond au seuil de différence de payoff entre deux actions au-delà duquel les individus estiment qu'il s'agit d'informations incompatibles (cf. section 3.3). Ce seuil dépend d'une part des valeurs de payoffs des actions dans la simulation et, d'autre part, de la manière dont ces payoffs sont retranscrites dans la simulation.

Par exemple, dans nos simulations le payoff d'une action correspond exactement à la valeur qui a été définie par l'utilisateur en amont de l'exécution. C'est pourquoi dans notre étude, présentée au chapitre 4, deux actions sont

incompatibles dès lors qu'elles ont des valeurs de payoff différentes. Cependant, nous pouvons imaginer que les valeurs de payoff contiennent une composante aléatoire, par exemple en suivant une loi gaussienne centrée sur la valeur définie de l'utilisateur. Dans ce cas, il est important de définir un ϵ supérieur à zéro afin que toutes les actions ne soient pas toujours estimées comme étant incompatibles.

- **aTable**(*sg, sg*) définit les attitudes inter-groupes sociaux. Comme mentionné précédemment, les valeurs de cette table ne sont pas importantes, seuls les signes de ces valeurs sont utilisés dans notre modèle. Ces signes doivent être renseignés sur la base des observations sociologiques du terrain. Deux groupes sont liés négativement s'il existe une réelle tension au point que l'un se réjouisse lorsque l'autre a subi une action nocive. Dans tous les autres cas, y compris dans le cas où les deux groupes affichent une attitude neutre, le signe est positif.
- Les **payoff** des actions représentant l'impact de ces actions sur les individus. Ces valeurs jouent un rôle décisif sur la dynamique des attitudes car elles sont à la base de notre modèle. A ce jour, il n'existe aucune méthode scientifique permettant d'estimer ces valeurs. Comme nous le montrerons dans le chapitre 5, dans notre étude d'un cas réel, nous avons défini ces valeurs en coopération avec les experts du domaine (en l'occurrence il s'agit d'experts militaires sur les actions d'influence).
- **aList**, les listes chronologiques des actions ou *scénario* du déroulement des actions. Elles définissent les dates d'occurrence auxquelles les acteurs doivent effectuer les actions.

3.8 CONCLUSIONS

Nous avons proposé un modèle de simulation des dynamiques d'attitude fondé sur des théories socio-psychologiques. Ce modèle construit les attitudes sur la base des croyances des individus. Ces croyances sont acquises au travers de la perception directe, la communication narrative inter-personnelle ou par le biais de la diffusion d'information par les acteurs. Selon leur niveau de crédibilité, ces informations peuvent amener les individus à réviser leurs croyances ou à les renforcer de manière probabiliste. Le modèle cognitif des agents leur permet notamment d'estimer l'intérêt de ces informations selon leurs impacts émotionnels et le sentiment de surprise qu'elles génèrent. Ces valeurs d'intérêt ont un double usage : d'une part elles représentent l'accessibilité des informations en mémoire et, d'autre part, permettent de calculer la probabilité de communication d'un message narratif relatant l'information. Notre mécanisme de communication tient compte des relations et des influences sociales au sein de la population. Dans notre modèle, l'attitude correspond à l'agrégation des croyances pondérées par leurs accessibilités selon une approche rétrospective de l'évaluation des événements. Nous avons notamment introduit un mécanisme de co-responsabilité permettant aux individus d'attribuer une part de responsabilité à d'autres acteurs qui ne sont pas à l'origine de l'action.

Ce processus cognitif de révision des attitudes est en accord avec certaines observations empiriques :

1. Inertie de l'attitude : la nature intrinsèque de l'attitude la rend très peu fluctuante. En conservant les dernières évaluations pour le calcul de la nouvelle attitude qui provient de l'acquisition d'une nouvelle information, le modèle retranscrit cette particularité.
2. Cohérence entre les croyances et les attitudes : dans notre modèle, les attitudes étant directement basées sur les croyances, l'harmonie cognitive est préservée comme le propose Heider (1982).
3. Compatibilité avec d'autres théories de l'attitude : notre structure d'attitude est compatible avec d'autres théories d'attitudes comme le modèle tripartite avec les composantes cognitives, émotionnelles et comportementales Rosenberg et Hovland (1960).
4. Une mémoire limitée : en implémentant la théorie de la Simplicité, l'agent est capable de limiter la quantité d'informations utilisée pour réviser les attitudes.
5. Perception subjective : les impacts sont évalués en considérant l'attitude des individus, ce qui rend l'information subjective à son point de vue.

Dans ce chapitre, nous exposons diverses expérimentations visant à (1) analyser la sensibilité du modèle à ses paramètres de simulation et (2) à illustrer les phénomènes descriptibles dans le cadre de Polias. Ces expérimentations nous permettent également de valider certains de nos choix de modélisation, en vérifiant que les mécanismes et paramètres du modèle jouent effectivement leurs rôles attendus dans le processus global. Toutes les expérimentations de ce chapitre mettent en œuvre des scénarii génériques, artificiels. Dans le chapitre suivant, nous illustrerons les capacités descriptives et explicatives de notre modèle, sur la base d'un scénario basé sur des faits réels (chapitre 5).

Dans un premier temps, nous étudierons l'impact des paramètres déterminants l'effet de l'émotion et de la surprise d'une information agissant sur l'évolution des attitudes. Deux scénarii illustreront ces impacts et permettront d'évaluer les intervalles de valeurs pertinents qui entreront en jeu dans la suite des expérimentations (section 4.1). Sur la base des intervalles, restreints pour les paramètres précédents, nous procéderons par la suite à l'analyse de l'impact de la distance sociale sur les attitudes à l'aide d'un scénario mettant en jeu trois groupes sociaux (section 4.2). Enfin, nous étudierons le mécanisme de communication, l'un des vecteurs principaux de l'acquisition des croyances. Dans cette thématique, nous étudierons en premier lieu les attributs des messages qui influent sur leur propagation dans le réseau d'interaction (section 4.3.1) avant d'évaluer l'impact du paramètre de communication (section 4.3.2). Nous terminerons ce chapitre en présentant un scénario dans lequel la population est confrontée à une opération de désinformation (diffusion d'une fausse information par les médias) afin d'étudier l'impact des paramètres de révision de croyances (section 4.4).

4.1 IMPACT DE L'ÉMOTION ET DE LA SURPRISE : ÉTUDE FONCTIONNELLE

Dans cette section, nous étudions l'impact des composantes émotionnelles et d'effet de surprise sur la dynamique des attitudes. Notre objectif est d'identifier les valeurs pertinentes des paramètres relatifs à ces composantes, à savoir α et ξ . Ces valeurs seront employées dans les analyses de sensibilités qui suivront dans les prochaines expérimentations. Nous définirons notamment différents profils cognitifs caractérisés par un ensemble de valeurs de paramètres. Il s'agit de personnes partageant une tendance cognitive commune. Ces profils seront notamment utiles afin de limiter l'espace des paramètres à explorer lors des prochaines expérimentations.

MODÈLE DE L'ÉMOTION ET LA SURPRISE Pour rappel, $\alpha \in [0, 1]$ a pour rôle d'équilibrer la partie émotionnelle et celle relative à la surprise dans le calcul

de l'intérêt d'une information donnée (cf. section 3.4). Ci-dessous, l'équation dans laquelle le paramètre intervient.

$$NI(\alpha) = 2^{\alpha \times E(\alpha) + (1-\alpha) \times S(\alpha)} \quad (25)$$

D'après cette équation, plus α est grand, plus le poids de l'émotion est important, au détriment de la surprise et vice versa.

Le paramètre $\xi \in]0, 1]$ correspond à la sensibilité émotionnelle d'un individu qui intervient dans la partie émotionnelle $E(\alpha)$ de l'équation précédente.

$$E(\alpha) = \log \left(1 + \frac{|\alpha \cdot \text{pyf}|}{\xi} \right) \quad (26)$$

Selon cette formule, conformément à la loi de Weber-Fechner, plus ξ est petit, plus l'individu est sensible et plus l'amplitude de la réponse émotionnelle $E(\alpha)$ sera importante.

4.1.1 Premier protocole

Afin d'étudier l'impact de ces deux paramètres sur la variation des attitudes, nous avons mis en place un premier scénario dans lequel 150 individus sont exposés quotidiennement à des actions durant 15 semaines (105 jours). Il existe 10 types d'actions avec des payoff positifs compris entre 0 et 50^1 qui peuvent représenter la diffusion d'une émission radio, une patrouille, etc. Chaque semaine, une ou plusieurs de ces actions agissent sur 5% de la population (7 agents tirés aléatoirement). La probabilité qu'une action survienne est de 0.1 (une chance sur dix) par semaine. Il survient donc en moyenne une action par semaine et chaque agent perçoit (témoin direct) environ 10 actions au cours d'une simulation.

L'indicateur employé pour cette analyse est la moyenne des attitudes de la population à la fin de chaque simulation (attitude moyenne à $t = 105$). Dans l'objectif de limiter la stochasticité des résultats, nous répliquons 100 fois chaque expérience (définie par un couple (α, ξ)), les résultats présentés sont la moyenne de l'indicateur sur les 100 simulations.

L'espace de paramètre exploré est défini par :

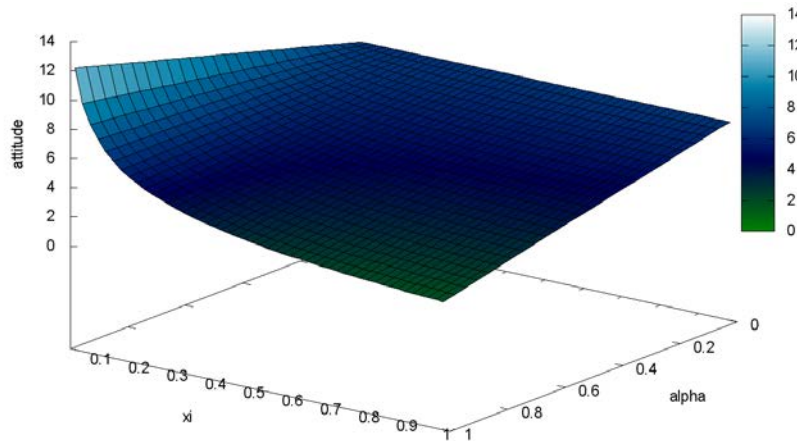
- α allant de 0 à 1 avec un pas de 0.04 ;
- ξ allant de 0.01 à 1 avec un pas de 0.04 (ce paramètre étant en dénominateur, il ne peut pas être égal à zéro).

Les autres paramètres sont fixés à des valeurs constantes neutres (comme pour toutes les expérimentations qui suivront) :

- Réseau social de type small-world avec un voisinage moyen de 4 individus ;
- $T_{\text{com}} = 0$ la communication n'est pas limitée, elle n'affecte en rien l'étude des deux paramètres ;
- $\rho = 0$ la coresponsabilité n'intervient pas dans le scénario, il n'y a qu'une Force ;
- $n_{\text{Step}} = 14$ l'intérêt maximal d'une information expire au bout de deux semaines.

Le graphique présenté dans la figure 12 affiche les résultats obtenus pour les 625 expériences (62 500 instances de simulations) effectuées.

1. L'ajout d'actions négatives aurait eu pour conséquence de ramener l'attitude vers zéro, ce qui ne faciliterait pas l'analyse.

FIGURE 12 – Attitude moyenne de la population en fonction de α et ξ

Dans la figure 12, nous observons que :

1. sur le plan $\alpha = 1$, la courbe présente un profil logarithmique décroissant avec ξ croissant ;
2. sur le plan $\alpha = 0$, la courbe est constante : la valeur de ξ n'a aucune influence ;
3. avec α croissant, l'attitude baisse, excepté pour $\xi < 0.05$.

Les observations (1) et (2) sont des conséquences directes de la définition de notre modèle :

1. Avec $\alpha = 1$, seule la composante émotionnelle intervient dans le calcul de l'attitude, d'où un profil logarithmique sur le plan $\alpha = 1$. La sensibilité émotionnelle croît avec un ξ décroissant, comme le confirme l'attitude croissante avec ξ décroissant.
2. Avec $\alpha = 0$, seule la composante relative à la surprise intervient dans le calcul de l'attitude, donc la valeur de ξ n'exerce aucune influence.

Enfin, l'observation (3) confirme notre objectif de définition de l'attitude en fonction de la surprise :

3. L'attitude est impactée par α , en témoigne la pente décroissante sur l'axe des α . Cette décroissance est due au scénario employé dans lequel l'impact de la surprise est plus important que les payoff des actions. Toutefois, avec une sensibilité émotionnelle suffisamment importante, la réponse émotionnelle est plus grande que la surprise ($\xi < 0.05$).

ξ est un facteur d'échelle du logarithme dans la formule de l'impact émotionnel du stimuli. α pondère les deux parties de l'intérêt (stimuli et surprise). Par la suite, nous travaillerons avec des valeurs $\xi < 0.05$ afin que les agents soient sensibles aux payoff des actions quel que soit le niveau de surprise. Par exemple, un individu doit être sensible à un attentat suicide, même si cet événement est ancien.

4.1.2 *Second protocole*

Pour étudier plus précisément l'impact de α et ξ , nous nous appuyons sur une seconde expérience. Nous enrichissons le précédent scénario à l'aide d'une action supplémentaire dont le payoff est négatif (-20). Cette action survient une seule fois à la 7^{ème} semaine et affecte également 5% de la population.

La figure 13 décrit sur une seule simulation un exemple d'évolution temporelle des attitudes de la population avec $\alpha = 0.5$ et $\xi = 0.01$.

Nous constatons que l'attitude augmente au fur et à mesure des occurrences d'actions positives et qu'elle baisse brutalement de Δ lorsque l'action négative intervient. La moyenne des attitudes remonte ensuite très légèrement avec les occurrences d'actions positives qui suivent.

Dans ce second protocole, nous étudions l'impact de α et ξ sur la différence Δ des valeurs d'attitudes entre l'instant précédant l'intervention de l'action négative et la fin de la simulation.

L'espace de paramètre exploré est le même que dans le premier protocole.

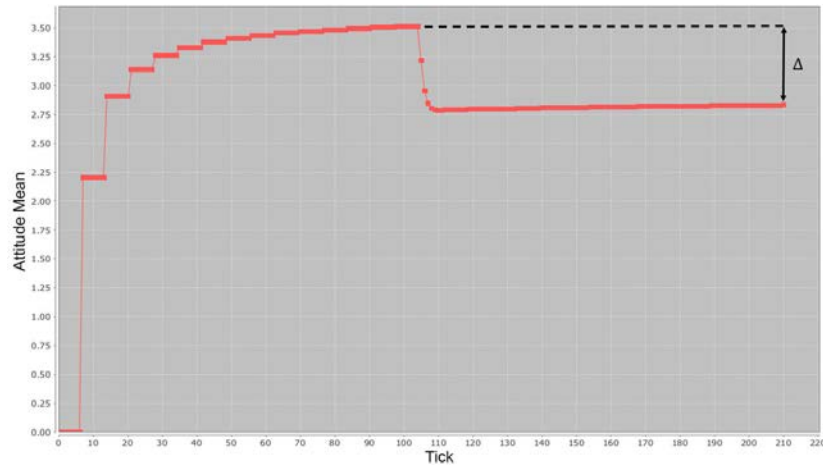


FIGURE 13 – Evolution de la moyenne des attitudes dans le scénario perturbé

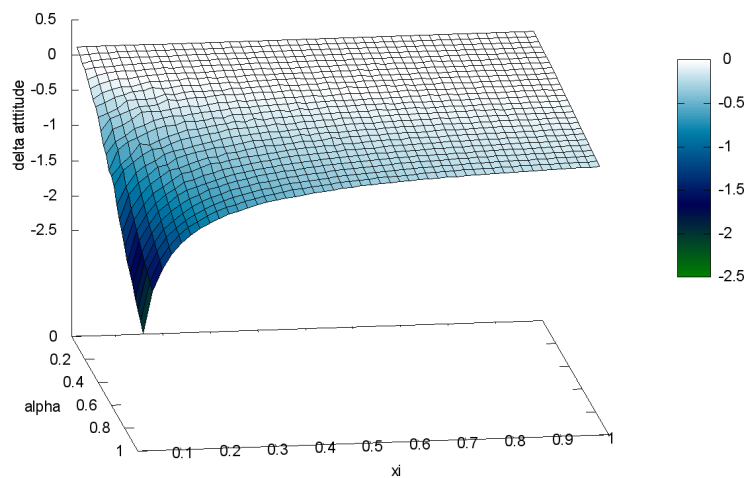
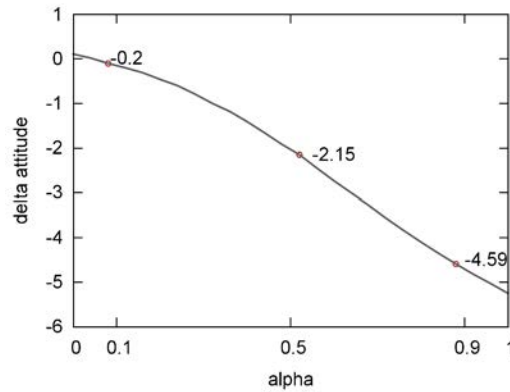


FIGURE 14 – Δ en fonction de α et ξ

FIGURE 15 – Δ en fonction de α avec $\xi = 0.01$

Dans la figure 14 nous observons sur le plan $\alpha = 1$ que la courbe présente un profil logarithmique croissant avec ξ croissant. L'impact de α est identique à la précédente expérience. Avec $\alpha \leq 0.32$ les attitudes ne sont pratiquement plus affectées ni par α , ni par ξ (à 0.02 près). De la même manière, la sensibilité du modèle à ces deux paramètres est quasiment nulle à partir de $\xi \gtrsim 0.26$.

Observations complémentaires : nous avons mesuré, pour chaque point (α, ξ) de notre expérience, l'écart-type entre les 100 simulations. Cet écart-type est au maximum de 0.37, mais sur l'ensemble, il est en moyenne de 0.02. La variance est donc très faible, cette expérience présente des résultats cohérents entre les simulations pour un même point.

Cette fois, l'indicateur Δ étant négatif, la courbure est inversée sur l'axe des z . Plus les agents sont émotionnels (α grand) et sensibles (ξ petit), plus l'impact négatif de l'action néfaste est amplifié sur les attitudes de la population. Cette seconde expérience confirme le rôle de ξ et α identifié lors de la première expérience.

ξ étant un facteur d'échelle de la sensibilité au stimuli, il n'est pas voué à être modulé pour des valeurs de payoffs données. Par contre, α pondérant les deux parties de l'intérêt peut servir de facteur d'arbitrage afin de différencier plusieurs types de profils d'individus. Nous décidons ainsi de travailler avec $\xi = 0.01$ et sélectionnons trois valeurs de α (voir la figure 15) qui nous paraissent caractériser trois types de profils cognitifs :

- profil d'agents *émotionnels* : $\alpha = 0.9$, $\xi = 0.01$; 90% de l'intérêt se base sur l'impact du stimuli au détriment de la surprise ;
- profil d'agents *peu émotionnels* : $\alpha = 0.1$, $\xi = 0.01$; seulement 10% de l'intérêt se base sur l'impact du stimuli au profit de la surprise ;
- profil d'agents *équilibré* : $\alpha = 0.5$, $\xi = 0.01$; l'impact du stimuli et celui de la surprise sont équivalents.

CONCLUSIONS DES DEUX EXPÉRIENCES Nous avons constaté qu'une faible valeur de ξ permet d'obtenir une réponse émotionnelle importante vis-à-vis des payoff des actions quel que soit le niveau de surprise engendré par l'événement. Cette sensibilité émotionnelle et le niveau de surprise peuvent notamment être modulés par le biais du paramètre α , en favorisant plus ou moins l'une des deux

composantes. Enfin, nous avons pu définir, sur la base des résultats obtenus, trois profils cognitifs avec lesquels nous poursuivrons nos expérimentations.

4.2 IMPACT DES GROUPES SOCIAUX SUR LA DYNAMIQUE DES ATTITUDES

Dans cette section nous illustrons l'effet des attitudes inter-groupes sociaux sur la perception et l'évaluation des actions selon l'affiliation des individus. En effet, le modèle prévoit des scénarii mettant en scène plusieurs groupes sociaux qui peuvent être antagonistes tels que les Pashtuns et les Tajiks en Afghanistan (les deux principales ethnies dans cette région). Dans de tels cas, les attitudes doivent évoluer de manière opposée chez les groupes antagonistes (attitude inter-groupe négative). Ce mécanisme est permis par l'usage du signe de l'attitude inter-groupe dans l'équation de l'évaluation d'une action (cf. section 3.6.1). De plus, la surprise intègre une dimension sociale qui dépend de la distance entre les individus dans le graphe d'interaction (cf. section 3.4.2). Notre objectif dans l'expérimentation qui suit est d'illustrer l'impact de ces deux mécanismes.

Nous avons établi un scénario dans lequel chaque groupe social bénéficie, à chacun son tour, d'une action positive. La population est constituée de trois groupes sociaux SGA, SGB et SGC comportant chacun 100 agents. SGA et SGB sont alliés (ils sont liés par une attitude positive) contre SGC, antagoniste des deux autres (ils ont une attitude négative vis-à-vis de SGA et SGB). Ces relations entre les groupes sociaux sont décrites dans la figure 8 (section 3.1.4).

Le réseau d'interaction permettant aux individus de communiquer entre eux est constitué de trois graphes de type small-world. Chacun de ces graphes est généré avec un degré de voisinage de 4. Enfin, nous avons ajouté aléatoirement des arcs afin de connecter ces trois sous-graphes en un graphe. 10% des individus de chaque réseau sont aléatoirement connectés aux deux autres graphes. Ainsi, nous obtenons un graphe dans lequel les membres d'un même groupe sont fortement connectés entre eux (connectivité de 4) mais aussi avec quelques liens entre les groupes.

Le scénario se compose de trois phases. Durant chaque phase un groupe social est stimulé par une répétition d'actions positives aléatoires (10 types d'actions, à l'instar des précédentes expériences). Les phases se succèdent dans l'ordre : SGA, SGB puis SGC.

Nous répétons cette expérience sur les trois profils cognitifs définis précédemment.

L'indicateur employé est à nouveau la moyenne des attitudes, par groupe social.

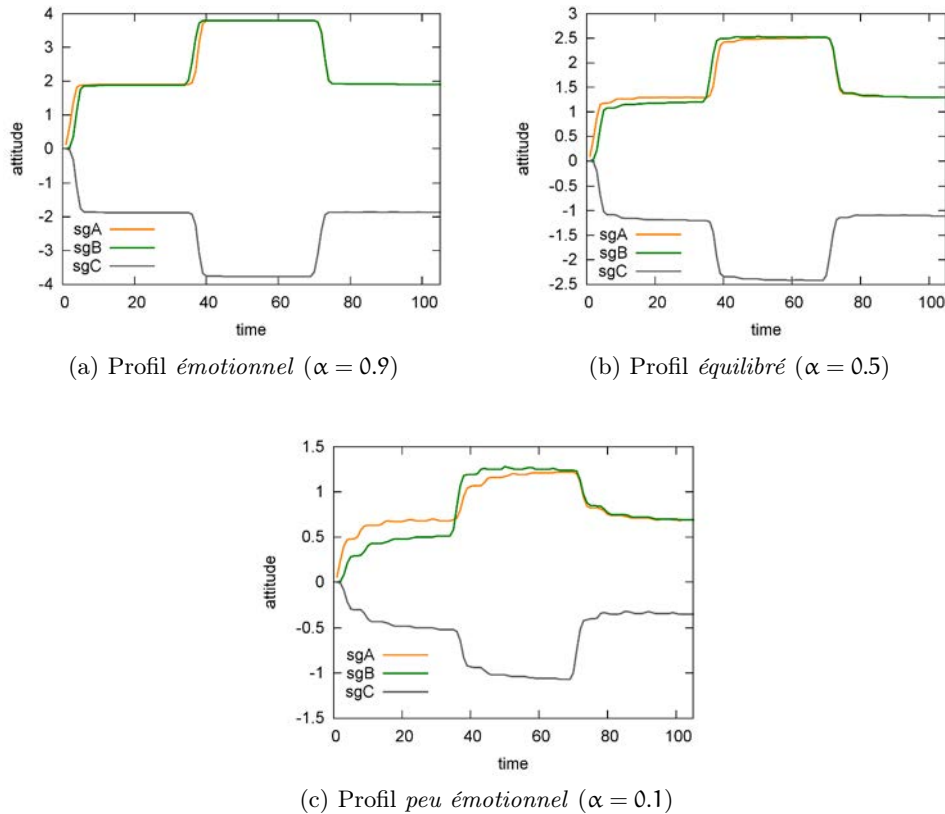


FIGURE 16 – Evolution de la moyenne d'attitude des groupes sociaux dans le temps selon les trois profils cognitifs

Dans les figures 16, nous observons que :

1. Les trois groupes réagissent à chaque phase de stimuli, bien que tous les groupes ne soient pas bénéficiaires d'une action. De plus, nous pouvons percevoir un décalage dans les temps de réaction des groupes ;
2. Lorsque la moyenne du SGA augmente, celle de SGB augmente également et celle des SGC baisse ;
3. Lorsque la moyenne du SGB augmente, celle de SGA augmente également et celle des SGC baisse ;
4. Lorsque la moyenne du SGC augmente, celle de SGA baisse, ainsi que celle de SGB ;
5. Avec $\alpha = 0.9$, lorsque les moyennes sont stabilisées, les moyennes de SGA et SGB sont égales, la moyenne de SGC a une valeur opposée ;
6. Avec $\alpha = 0.5$, les moyennes mettent davantage de temps à se stabiliser ;
7. Avec $\alpha = 0.1$, les moyennes de SGA et SGB ne convergent pas à la première phase, puis convergent à la fin de la seconde phase.

La première observation s'explique du fait que, bien qu'à chacune des phases, les actions ne stimulent qu'un unique groupe social, les autres groupes sociaux sont informés de ces actions par le biais de la communication. En effet, les différents groupes sont reliés par le réseau d'interaction.

Les observations (1), (2) sont dues à l'alliance entre les groupes SGA et SGB (attitude inter-groupe positive entre ces deux groupes). En effet, comme nous

l'avons rappelé précédemment, l'évaluation d'une action positive pour un bénéficiaire « ami » est positive.

L'observation (3) témoigne du même mécanisme, cette fois avec un signe opposé pour SGC : l'évaluation d'une action positive pour un bénéficiaire « ennemi » est négative. Cela explique la dynamique d'attitude symétrique de SGC aux dynamiques de SGA et SGB.

Les trois dernières observations (5, 6 et 7) relèvent de la surprise sociale. La distance dans le graphe d'interaction au bénéficiaire impacte l'intérêt de l'information (et donc la probabilité de communication ainsi que l'attitude). Or, selon le graphe que nous avons défini, deux individus de différents groupes sociaux sont plus éloignés que s'ils faisaient partie du même groupe. Ainsi, l'attitude est plus fortement impactée lorsque le bénéficiaire est membre du même groupe que dans le cas contraire. C'est pourquoi la vitesse de convergence diffère en fonction de α : plus α est petit, plus la surprise sociale est importante, plus l'intérêt dépend de la distance sociale, moins les attitudes sont impactées dans le cas où cette distance est grande.

CONCLUSIONS SUR L'IMPACT DES GROUPES SOCIAUX Cette expérience a permis d'illustrer le mécanisme d'évaluation des actions en fonction des alliances entre les groupes sociaux. De plus, elle a mis en évidence l'impact de la dimension sociale dans l'évaluation de l'intérêt d'une information. Nous avons notamment établi que la communication joue un rôle décisif dans l'évolution des attitudes. Notre modèle est capable de décrire des dynamiques d'attitudes dans des scénarii qui mettent en jeu différents groupes sociaux sur la base (1) des attitudes inter-groupes et (2) de la distance sociale.

4.3 ETUDE FONCTIONNELLE DU MÉCANISME DE COMMUNICATION

La communication inter-personnelle permet aux agents d'acquérir des informations portant sur des événements dont ils n'ont pas été les témoins directs. Le modèle d'attitude étant fondé sur des croyances, il est important de maîtriser ce vecteur d'information qui impacte directement la connaissance acquise par des agents.

Pour ce faire, nous proposons d'étudier les paramètres du modèle cognitif et les attributs d'information qui affectent la communication selon deux scénarii. Dans un premier scénario nous étudions la propagation d'une information perçue par un unique individu au milieu d'un environnement dans lequel d'autres actions sont générées aléatoirement. Dans le second scénario, un individu témoin d'une « vraie » information est confronté à une opération de désinformation : une communication de masse diffuse une fausse information à la population, contredisant celle qui a fait l'objet d'un témoignage.

RAPPEL SUR LE MODÈLE DE COMMUNICATION La probabilité qu'un agent transmette une certaine information a_e à son voisin dépend de différentes actions et de différents attributs de ces actions.

Elle dépend de plusieurs actions car l'action a_e dont on étudie la probabilité d'être communiquée est mise en concurrence avec d'autres informations connues par l'agent. D'après la formule ci-dessous définissant la probabilité $p_s(i, a_e, t)$

de communication dépend de $I_{\max}(t)$, l'intérêt maximal d'une action observée depuis un certain temps.

$$p_s(i, a, t) = \begin{cases} 2^{I(a,i)-I_{\max}(i,t)} & \text{si } 2^{I(a,i)-I_{\max}(i,t)} > T_{\text{com}} \\ 0 & \text{sinon} \end{cases} \quad (27)$$

Chacun des I et I_{\max} dépend de deux actions : l'occurrence de l'action la plus récente et celle qui la précède. En effet, dans le calcul de l'intérêt, les agents estiment un degré de surprise temporelle et sociale en fonction de deux occurrences de l'action (comme présenté dans la section 3.4.2).

Mais la probabilité de transmission dépend aussi de plusieurs attributs d'action car, comme nous venons de le rappeler, la surprise dépend des dates et du nombre d'occurrences sur chaque groupe social. De plus, l'intérêt d'une information est également relatif à une composante émotionnelle liée au payoff de l'action.

$$I(a) = \alpha \times E(a) + (1 - \alpha) \times S(a) \quad (28)$$

En résumé, pour chaque action intervenant dans le calcul de la probabilité, les attributs qui impactent la communication sont :

- le payoff ;
- la date ;
- le nombre d'occurrences pour chaque groupe social.

Enfin le paramètre T_{com} intervient comme seuil dans la décision de communication.

Par conséquent, nous proposons d'étudier l'impact de ces paramètres (T_{com} , payoff, date et nombre d'occurrences) sur la diffusion d'une action dans le réseau social.

PROTOCOLE GÉNÉRAL Afin d'étudier l'impact de ces variables sur la propagation de l'information dans la population, nous nous appuierons sur le taux d'individus détenant une certaine information que nous nommons *awareness*.

La figure 17 ci-dessous illustre cette notion d'*awareness* : la courbe représente le taux de la population informée d'une action témoin par un unique individu au début de la simulation. Nous observons que ce taux augmente au fil du temps, en formant une courbure en S typique de la diffusion d'information (comme l'a montré par exemple Rogers (2010)). Au début de la diffusion, seuls le témoin et ses voisins sont au courant (de $t = 49$ à $t = 50$). Peu après, le taux augmente exponentiellement en fonction du nombre de voisins (période $t = 50$ à $t = 55$). En effet, lorsque peu d'agents sont au courant d'une information, le nombre d'agents au courant au pas de temps suivant de la simulation sera environ $N + N \times V$ où N est le nombre d'agent au courant et V le nombre de voisins moyen dans le graphe. Nous avons donc une suite géométrique $N(t) = N(t-1) \times (1 + V)$, de croissance exponentielle. Vers la fin de la simulation, il ne reste plus que les agents peu connexes, plus lents à réceptionner l'information ($t = 55$ à $t = 59$).

4.3.1 Impact du payoff et de la date de l'information sur sa diffusion

Comme nous l'avons mentionné, les deux attributs d'un message qui influencent sa portée sont le payoff et la date d'occurrence de l'action relatée. Ces deux

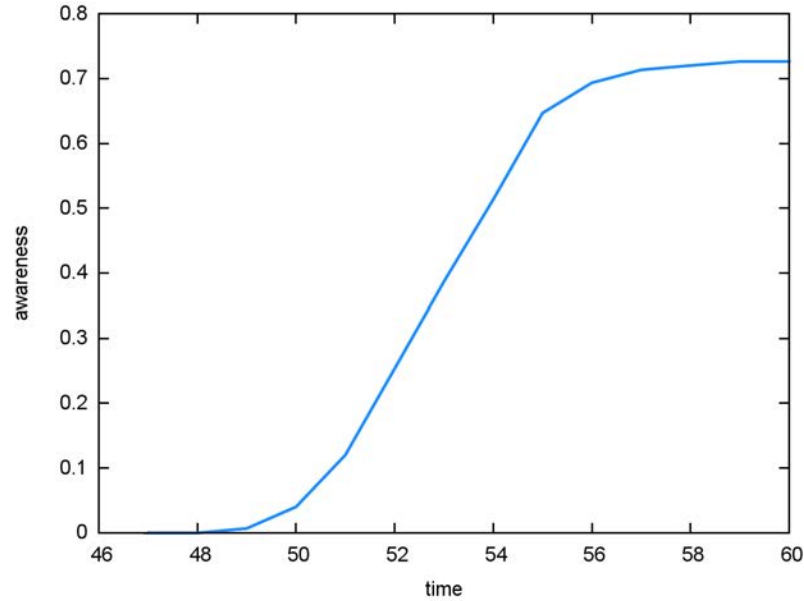


FIGURE 17 – Awareness en fonction du temps

attributs impactent respectivement la composante émotionnelle et celle de la surprise. L'objectif de l'expérience suivante consiste à analyser la sensibilité du modèle de communication à ces deux attributs dans un environnement bruité (c-à-d. dans lequel les agents sont soumis à d'autres informations).

De manière similaire aux précédentes expérimentations, la population composée de 150 agents est stimulée par 10 actions de payoff aléatoires comprises entre 0 et 50. Chacune de ces actions survient à un rythme hebdomadaire, avec une probabilité de 0.1. A chaque occurrence elle est témoinnée par 5% de la population. Nous avons montré dans la précédente section, au travers de différentes expériences, que ces actions ont un impact sur l'attitude de la population. Mais ces actions ont également un impact sur la communication : les agents vont relayer ces informations. Ces événements constituent le bruit de fond du scénario.

A la 7^{ème} semaine nous introduisons une action dont nous contrôlons le payoff et la dernière date d'occurrence. Elle affecte un unique individu tiré aléatoirement. Le but consiste à étudier la diffusion de cette information en fonction des deux attributs contrôlés. L'indicateur est la proportion de la population informée de cette action survenue sur un unique individu, c'est l'awareness de cette action.

L'espace des paramètres exploré est défini par :

- payoff allant de 0 à 50 ;
- UT_0 allant de 0 à 10 (log de la distance temporelle entre l'occurrence courante et la précédente, voir la section 3.4.2).

Chaque point de l'espace est moyenné sur 200 simulations. La communication étant probabiliste, les résultats sont davantage sujets à la stochasticité.

Nous procédons à cette analyse sur les trois profils de population définis précédemment (cf. section 4.1.2).

Dans les figures 18, nous pouvons observer que :

1. Quel que soit le profil, l'awareness augmente avec les attributs croissants ;

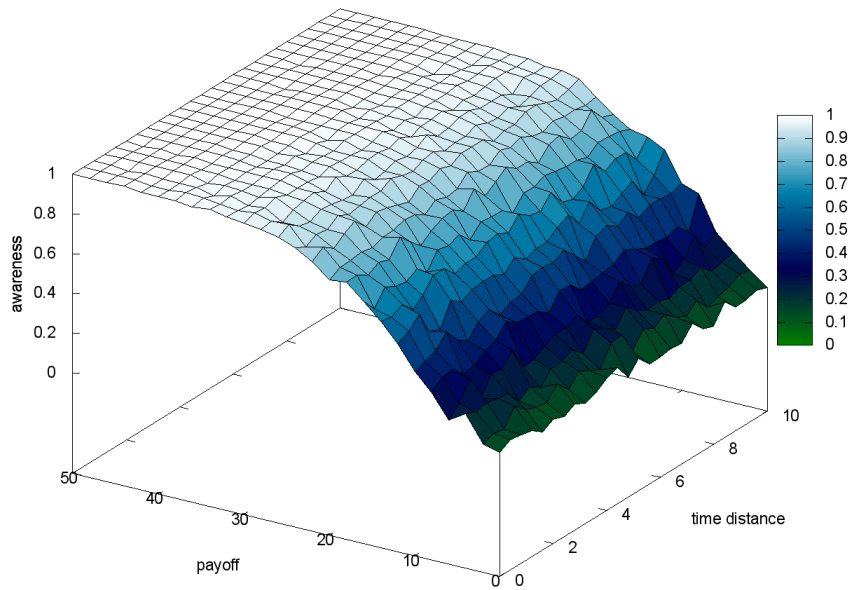
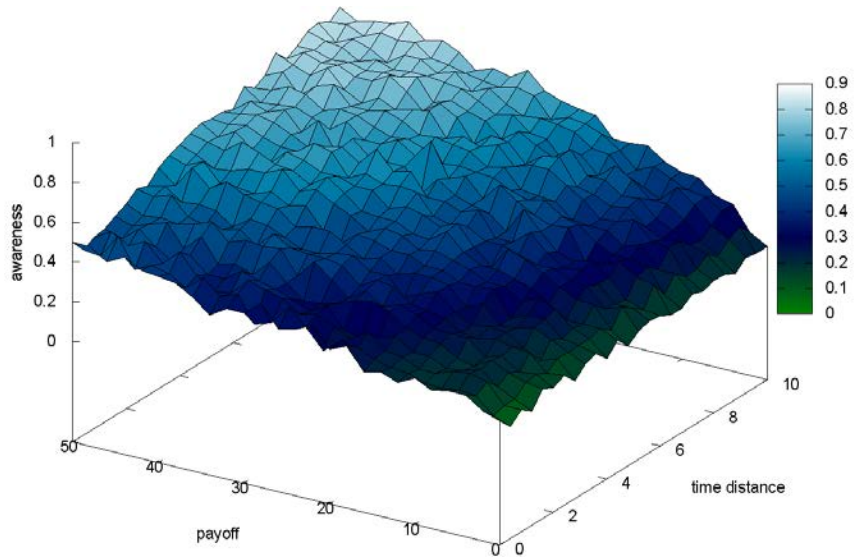
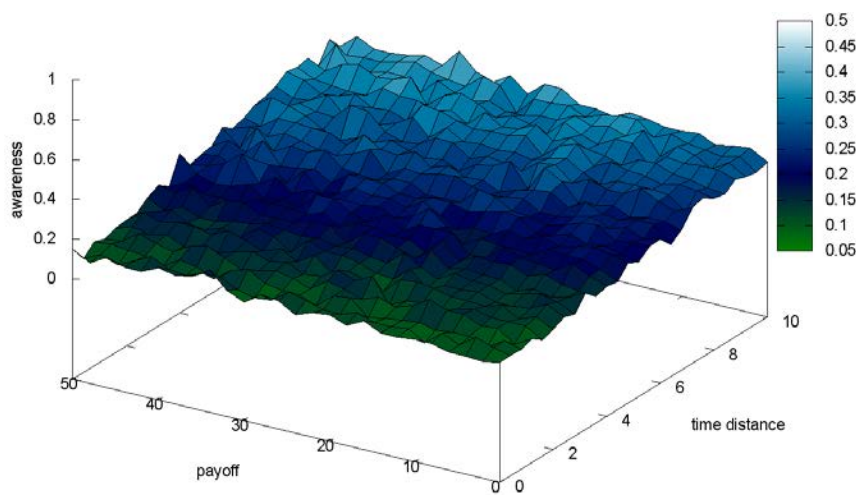
(a) Profil de population *émotionnel* ($\alpha = 0.9$)(b) Profil de population *équilibré* ($\alpha = 0.5$)(c) Profil de population *peu émotionnel* ($\alpha = 0.1$)

FIGURE 18 – Awareness en fonction de la surprise temporelle et le payoff de l'action selon trois profils cognitifs

2. L'impact de la distance temporelle est quasiment nul avec $\alpha = 0.9$ (figure 18c) ;
3. L'impact de la distance temporelle et celle du payoff sur l'awareness sont similaires avec $\alpha = 0.5$ (figure 18b) ;
4. L'impact du payoff est très faible avec $\alpha = 0.1$ (figure 18a) ;
5. L'awareness maximal (100% de la population) est atteint relativement rapidement ($\text{payoff} > 32$) avec $\alpha = 0.9$ (figure 18c) ;
6. L'awareness est au maximum de 78% avec $\alpha = 0.5$ (figure 18b) ;
7. L'awareness est au maximum de 47% avec $\alpha = 0.1$ (figure 18a) ;

Observations complémentaires : l'écart-type moyen entre les awareness sur les 200 réplifications de chaque expérience est de 0.32. Elle est au maximum à 0.44. il est tout à fait normal d'avoir une variance élevée car nous étudions ici un phénomène de communication probabiliste.

La probabilité de relayage d'une information croît avec son payoff et sa rareté (distance temporelle entre l'occurrence courante et la précédente). C'est pourquoi l'awareness augmente avec les attributs croissants (observation 1). Notre modèle est donc capable de décrire le phénomène de propagation d'informations narratives au sein d'un réseau social en fonction de ces deux attributs majeurs (selon la Théorie de la Simplicité).

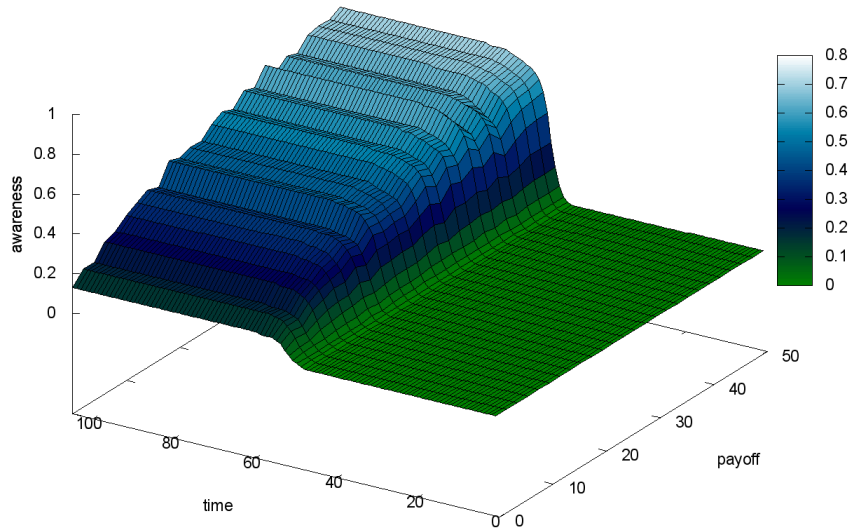
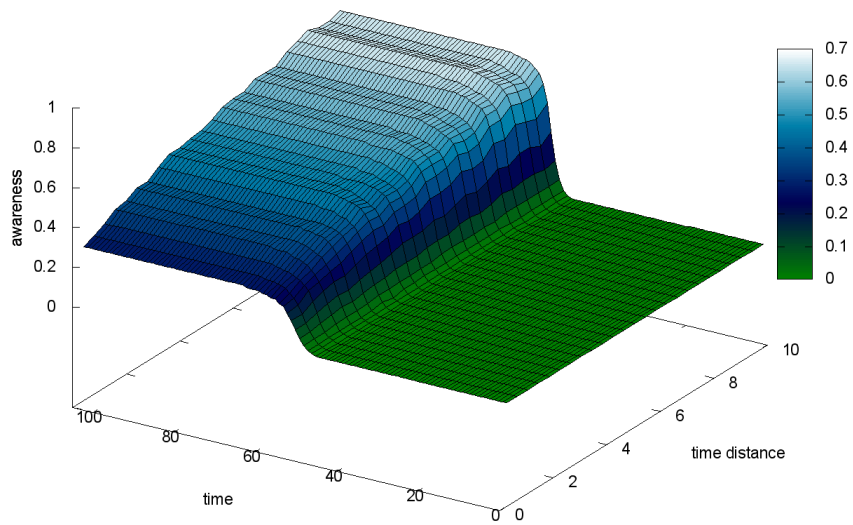
Les observations 2, 3 et 4 ont motivé notre choix du profil *équilibré* pour les études qui suivront. En effet, avec un profil cognitif trop émotionnel, la surprise n'intervient quasiment pas dans l'estimation de l'intérêt d'une information. De la même manière, avec un profil trop axé sur la surprise, l'impact émotionnel dépendant du payoff est inhibé. Or, par intuition, nous pouvons penser que ce ne soit pas le cas dans la plupart des circonstances réelles de propagation d'information narrative. Nous éviterons donc de travailler sur ce type de profil pour les expériences suivantes, nous privilégierons un profil *équilibré*.

Le poids de l'impact émotionnel est plus important que celui de la distance temporelle dans l'estimation de l'intérêt, en témoigne les observations 5, 6 et 7.

Afin d'approfondir cette analyse, nous proposons des vues temporelles pour l'expérience avec un profil cognitif *équilibré*. Ainsi, les figures 19a et 19b décrivent respectivement l'évolution de l'awareness dans le temps en fonction du payoff (avec $UT_0 = 5$) et de la distance temporelle (avec $\text{payoff} = 25$).

Nous observons dans les figures 19 que :

1. L'awareness maximal augmente avec un payoff ou une distance temporelle croissante ;
2. Le temps t_s entre l'introduction de l'information étudiée ($t = 50$) et le moment où l'awareness se stabilise (c-à-d. $\text{awareness}_t - \text{awareness}_{t-1} = 0$) augmente avec payoff.

(a) En fonction du payoff de l'action ($UT0 = 5$)(b) En fonction de la surprise temporelle ($pyf = 25$)FIGURE 19 – Evolution temporelle de l'awareness selon un profil cognitif *équilibré*

payoff	aw_{\max} (%)	t_s (jours)
0	13	60
10	33	63
20	45	65
36	65	67
50	71	69

TABLE 3 – Table de t_s de la figure 19a

L'observation (1) est la conséquence directe du modèle d'intérêt qui prend en compte notamment le payoff et la distance temporelle.

Cependant, l'observation (2) témoigne d'un phénomène moins intuitif. En effet, d'après les figures 19, le temps de convergence (temps entre l'introduction de l'information étudiée et le moment où l'awareness maximal est atteint) semble augmenter avec l'awareness maximal. Ce phénomène est rendu plus lisible dans la table 3 qui expose certaines valeurs numériques du graphique 19a. Cette table confirme la croissance du temps de convergence en fonction du payoff et de l'awareness maximal. Nous pouvons expliquer cette tendance par le fait que plus la portée de l'information est grande, plus la distance que l'information doit parcourir dans le graphe social est grande, plus l'awareness met du temps à atteindre sa valeur maximale. Bien entendu, cela implique une conséquence directe sur la dynamique des attitudes : les attitudes convergent plus lentement lorsque l'information se propage sur une durée allongée.

4.3.2 Le r le du seuil de communication T_{com}

Nous nous intéressons maintenant à l'impact du seuil de communication T_{com} ayant pour rôle de limiter le taux d'échange d'informations.

Nous reprenons le scénario précédent avec les 10 types d'actions en bruit de fond et l'action contrôlée intervenant à la moitié de la simulation. L'indicateur reste le même : l'awareness de cette action contrôlée. Cette fois, nous faisons varier le seuil T_{com} ainsi qu'un des deux attributs de l'action afin de varier la probabilité de communication, comme ce fut le cas dans l'expérience précédente. Nous avons fait le choix de moduler le payoff pour cette expérience (nous avons vu que le payoff et la distance temporelle ont le même impact sur la diffusion de l'information).

L'espace des paramètres exploré est défini par :

- la valeur du payoff de l'action contrôlée allant de 0 à 50 ;
- la valeur T_{com} allant de 0 à 1 avec un pas de 0.04.

Les autres paramètres fixes sont :

- Profil cognitif *équilibré* ;
- $UT_0 = 5$.

Nous observons dans la figure 20 que :

1. L'awareness augmente avec T_{com} décroissant ;

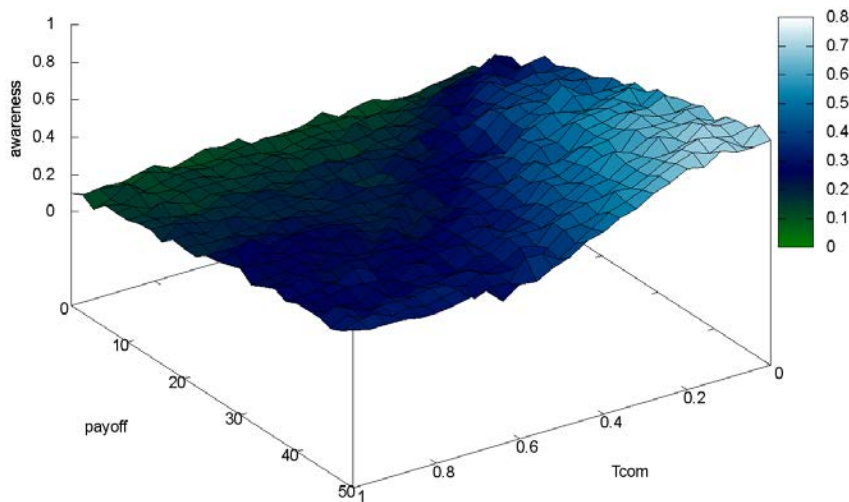


FIGURE 20 – Awareness en fonction du seuil de communication T_{com} et le payoff de l'action

2. L'awareness augmente avec payoff croissant ;
3. L'awareness maximal est de 74%.

Le paramètre T_{com} remplit bien son rôle de limitateur du taux de transmission de messages. En effet, nous pouvons interpréter ce paramètre comme un seuil de tolérance des agents à des informations peu intéressantes. Plus les agents sont tolérants, plus ils sont bavards, plus la portée de l'information augmente.

Ce paramètre est utile dans le cas où le modélisateur dispose de données réelles sur la portée d'un message relatant une action contrôlée. Il suffit alors de calibrer le modèle en modulant T_{com} afin de retrouver dans les simulations la portée mesurée. En l'absence de telles données, $T_{com} = 0$ convient par défaut, ce paramètre n'impacte rien d'autre que la portée des messages dans la simulation.

CONCLUSIONS SUR LE MODÈLE DE COMMUNICATION Dans la première expérience, nous avons exposé la capacité de notre modèle à reproduire le phénomène de diffusion de l'information dans une population tenant compte des deux paramètres d'une action (payoff et date), comme proposé par la Théorie de la Simplicité. La sensibilité du mécanisme de communication à ces deux attributs peut être modulée en fonction du paramètre α . Par ailleurs, la courbe d'évolution du taux de la population informé d'une action suit une courbure en S comme l'a montré Rogers (2010). Nous avons notamment détecté que le temps de convergence de cette courbe augmente avec le payoff (plus l'information est intéressante, plus il y a d'individus qui seront amenés à la connaître, plus le temps de diffusion est long), ce qui conduit à ralentir la convergence des attitudes. Enfin, dans la dernière expérience, nous avons démontré que le paramètre T_{com} permet de seuiller la tolérance des individus aux intérêts des messages afin de réguler le taux de communication inter-personnelle.

Dans la prochaine section, nous montrerons comment la population est capable de résister à une campagne de désinformation grâce à la communication inter-personnelle et au mécanisme de révision de croyance.

4.4 RÉVISION DE CROYANCE ET RÉSISTANCE À LA DÉSINFORMATION

Dans cette section, nous testons un scénario dans lequel l'acteur s'engage dans une opération de désinformation. Notre modèle cognitif étant doté d'un mécanisme de révision de croyance, la population devrait être capable de ne pas se faire duper par cette supercherie. Cette résistance est néanmoins régie par quelques paramètres que nous rappellerons dans le paragraphe suivant. Nous présenterons ensuite une expérience mettant en évidence l'impact de ces paramètres sur la propagation et l'acceptation de l'information (ou la résistance à la désinformation). À l'issue de cette première expérience, nous verrons que la topologie du réseau social exerce un impact sur cette dynamique. Nous concluons cette section par une analyse de sensibilité de la résistance au réseau d'interaction.

MODÈLE DE RÉVISION DE CROYANCE (RAPPEL) Le modèle propose un mécanisme de révision de croyances permettant aux agents de sélectionner les informations auxquelles ils veulent croire, en fonction des niveaux de crédibilité des sources. En effet, lorsqu'un agent acquiert une information incompatible à ses connaissances (présence dans sa base de croyances d'une croyance correspondant à l'information entrante, mais avec un payoff différent), il peut remplacer sa croyance par celle entrante. La probabilité de cette révision de croyance est donnée par un tableau qui confronte les crédibilités des deux informations (cf. la table 4 ci-dessous). π_0 est la probabilité qu'un sujet, peu convaincu de son information, révisé sa connaissance lorsque son interlocuteur croit fortement à son propos (l'interlocuteur a été témoin de son information). Dans le cas où la conviction de l'interlocuteur est supérieure d'un niveau à celle du sujet, la probabilité que le sujet révisé sa croyance est π_1 . Si les crédibilité des deux protagonistes sont équivalentes, nous fixons la probabilité $\pi_2 = \frac{1}{2}$. De plus, nous avons fait le choix de simplifier le tableau, en admettant une symétrie centrée sur la diagonale des crédibilités équivalentes. De ce fait, dans le cas où les rôles sont inversés (mêmes crédibilités, en inversant le sujet et l'interlocuteur), la probabilité correspond au complémentaire de la probabilité initiale. Exemple de symétrie de $\pi_0 : 1 - \pi_0$.

$\alpha_{ex.\sigma} \setminus \alpha_{in.\sigma}$	σ_3	σ_2	σ_1
σ_3	X	$1 - \pi_1$	$1 - \pi_0$
σ_2	π_1	π_2	$1 - \pi_1$
σ_1	π_0	π_1	π_2

TABLE 4 – Table de révision de croyance employée dans nos travaux.

Le mécanisme de révision de croyance dépend donc, dans notre cas, de deux paramètres : π_0 et π_1 . Dans cette première expérience, nous mettrons en évidence l'impact de ces paramètres sur la capacité des individus à résister à un campagne de désinformation.

4.4.1 *Première expérience*

Le premier jour, un individu (tiré aléatoirement) témoigne d'une action négative effectuée par un acteur donné (payoff = -20). Le même jour, cet acteur

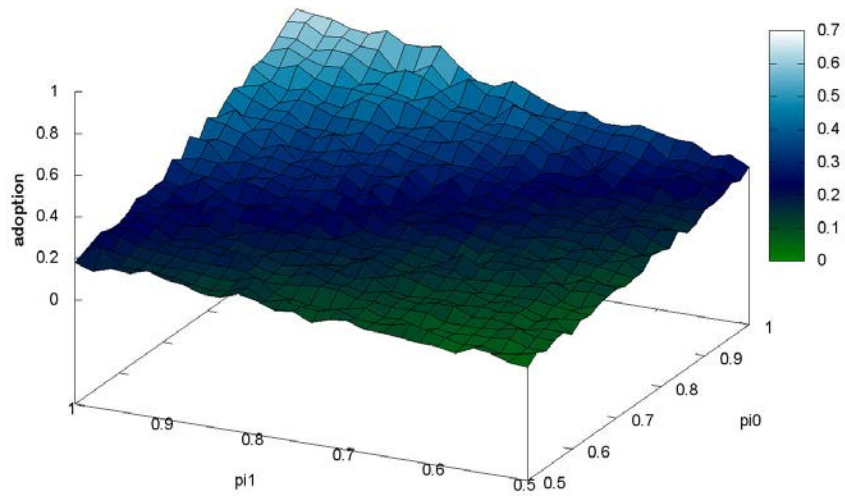
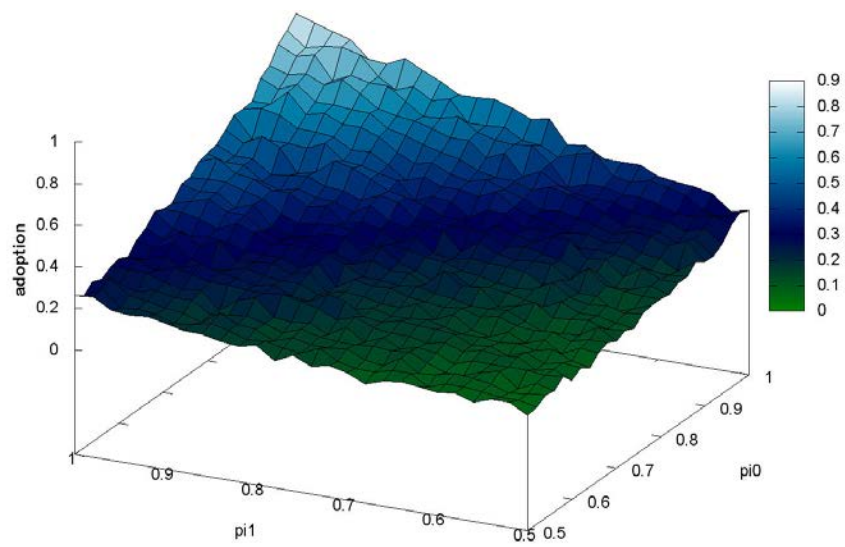
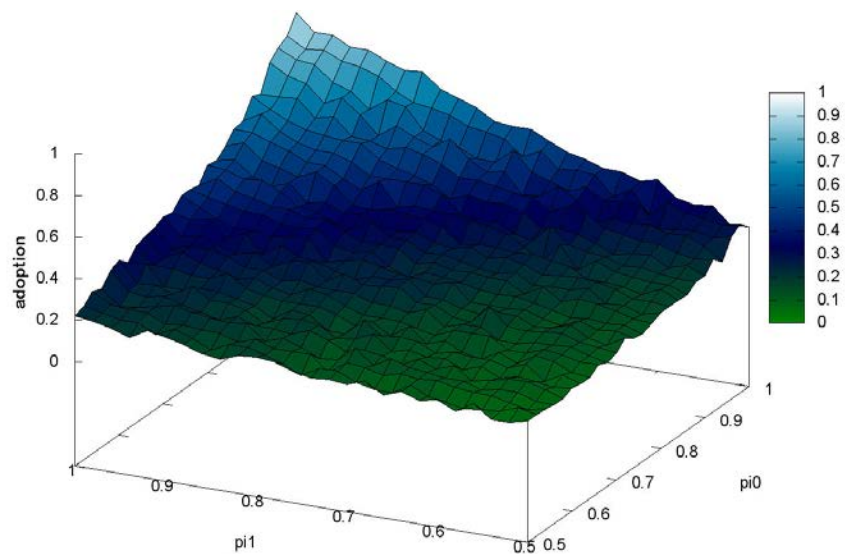
(a) $k = 4$ (b) $k = 6$ (c) $k = 8$

FIGURE 21 – Awareness en fonction des probabilités de révision selon trois réseaux sociaux

entame une opération de désinformation, afin de dissimuler cette vérité. Cette opération consiste à diffuser à toute la population (ex : par radio) une information contredisant l'action ayant eu lieu et dont il a été l'auteur. Cette fausse information relate que l'action qui a eu lieu avait un impact positif ($\text{payoff} = +20$) et atteint 90% de la population quotidiennement. Cependant, la crédibilité de l'acteur étant plus faible que celle de l'individu, le mécanisme de révision de croyance devrait limiter le nombre de personnes dupes de cette fausse information.

L'objectif de cette expérience est de suivre l'évolution de la proportion de la population amenée à croire la vérité témoignée (l'action à -20 de payoff) que nous nommons **adoption**. Cette évolution est fonction des valeurs des paramètres π_0 et π_1 .

L'espace des paramètres exploré est défini par :

- π_0 allant de 0.5 à 1 avec un pas de 0.02 ;
- π_1 allant de 0.5 à 1 avec un pas de 0.02.

Remarques : π_0 et π_1 démarrent à $\frac{1}{2}$ car notre table de révision de croyance admet une symétrie sur la diagonale $\pi_2 = \frac{1}{2}$. Par ailleurs, la moitié de cet espace de paramètre n'a pas de sens : lorsque $\pi_0 < \pi_1$ un individu a plus de chance de se faire convaincre par son voisin (réseau social) du contraire de ce qu'il croit et dont il a été le témoin direct. Nous ne considérerons donc que l'espace où $\pi_0 > \pi_1$.

Nous employons le profil cognitif *équilibré*.

Nous observons dans la figure 21a que :

1. L'adoption augmente avec π_0 croissant ;
2. L'adoption augmente avec π_1 croissant ;
3. Au maximum, l'adoption est de 65%.

Les observations (1) et (2) sont dues aux probabilités de révision liées aux crédibilités des informations confrontées :

1. Plus la probabilité de réviser sa croyance faiblement crédible par une information fortement crédible (π_0) augmente, plus les agents voisins du témoin sont susceptibles de le croire. De plus, la table de révision étant symétrique, plus π_0 augmente, moins le témoin a de chances de croire à la fausse information transmise par l'acteur. En effet, dans notre modèle, même si un individu a été témoin d'une information, il peut être amené à modifier cette croyance par une information contradictoire, selon une probabilité de $1 - \pi_0$. En outre, plus ses voisins lui affirmeront le contraire, plus il sera enclin à les croire. Ce phénomène correspond au processus de *conformité* par lequel un individu accepte de modifier une croyance sous l'influence de la pression sociale, comme l'ont montré Crutchfield (1955) ou encore Kelman (1958).
2. Plus la probabilité de réviser sa croyance par une information plus crédible d'un niveau (π_1) augmente, plus les voisins V_2 des voisins V_1 du témoin ont de chances de croire à la vérité : tous les individus croient à la fausse information à σ_1 (crédibilité de l'acteur) sauf le témoin ; si le témoin croit à σ_3 (crédibilité max), certains de ses voisins V_1 vont le croire, et la crédibilité sera de $\sigma_{3-1} = \sigma_2$; certains voisins de V_2 de ces voisins V_1 vont croire à la vérité selon la confrontation entre σ_1 et σ_2 , donc selon une probabilité

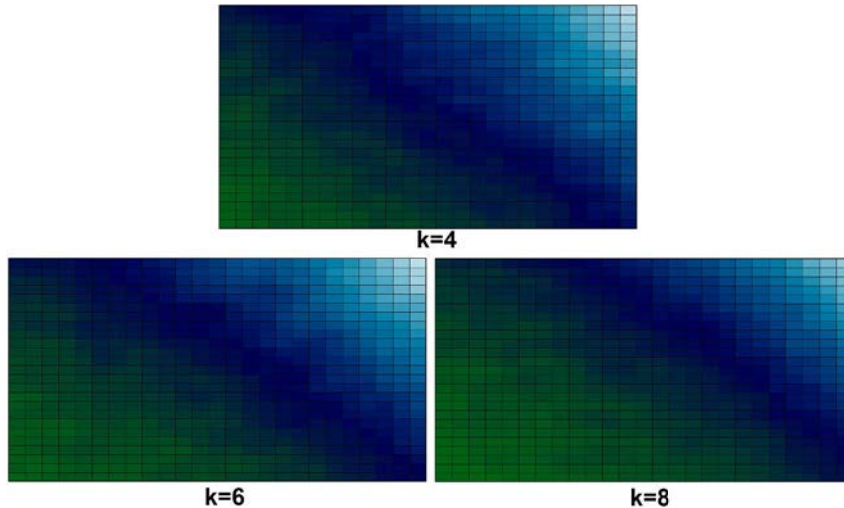


FIGURE 22 – Vues alternatives des figure 21.

Adoption en fonction de π_0 et π_1 avec $k = \{4, 6, 8\}$

de π_1 . De plus, la table de révision étant symétrique, plus π_1 augmente, moins ceux qui ont changé d'avis rebasculeront en faveur de l'acteur.

L'observation (3) témoigne que même avec des probabilités maximales (100% de chance d'être convaincu pour toutes informations), l'adoption ne dépasse pas 65%. Or nous nous attendions à ce que l'adoption atteigne à peu près 78% comme lors des expériences précédentes portant sur la diffusion de l'information (voir la section 4.3). En effet, tous les paramètres sont les mêmes (nombre d'individus, réseau, profil cognitif moyen, payoff etc.) sauf le fait que l'on ait introduit une fausse information. De plus, le payoff de l'information vraie et celui de la fausse information sont opposés (20 et -20) : les impacts émotionnels de ces deux informations sont égaux (car l'équation prend l'absolu du payoff). Par ailleurs, la surprise temporelle de l'action véritable est plus forte que celle de la fausse information et sa crédibilité est plus forte. Pour toutes ces raisons, nous pouvions penser que l'adoption de l'information vraie aurait atteint les mêmes résultats que pour les expériences de l'awareness, ce qui n'a finalement pas été le cas (observation 3). Cette baisse de l'adoption maximale de plus de 10% s'explique selon deux facteurs :

- la crédibilité de l'information témoinnée décroît de transmission en transmission, la portée de l'information est limitée par la fausse information ;
- le nombre d'interactions entre les individus est moins important que le nombre de communications de l'acteur.

Sur ce dernier point, le nombre d'interactions inter-personnelles dépend du nombre de voisins par individu (plus ils ont de voisins, plus ils parlent). En d'autres termes, la topologie du réseau influence la propagation de l'information témoinnée. Par conséquent, nous illustrons ce phénomène en procédant à une seconde expérience dans laquelle nous analysons l'impact du réseau.

4.4.2 Seconde expérience

Nous répétons la précédente expérience en faisant varier la connexité du réseau social. L'espace des paramètres exploré est défini par :

- π_0 allant de 0.5 à 1 avec un pas de 0.02 ;
- π_1 allant de 0.5 à 1 avec un pas de 0.02 ;
- connexité du réseau : $k = 6$ et $k = 8$.

Dans les figures 21, nous observons que :

1. L'adoption maximale αw_{\max} augmente avec k croissant : $\alpha w_{\max}^{k=4} = 65\%$, $\alpha w_{\max}^{k=6} = 87\%$, $\alpha w_{\max}^{k=8} = 93\%$;
2. L'espace de paramètres avec une adoption faible ($< 25\%$) grandit avec k croissant, ce phénomène est plus visible dans les figures 22 : la surface verte s'élargit avec k ;
3. L'adoption est relativement stable de 34% pour $k = 4$ et 40% pour $k = 6$, $k = 8$.

L'observation (1) témoigne que plus les voisins sont nombreux, plus il y a d'échanges d'informations inter-personnels, plus la probabilité que les individus croient à l'information témoinée augmente. Ce phénomène aurait pu également être vérifié en faisant décroître la portée ou le nombre de communications de l'acteur.

Plus concrètement, d'après cette analyse, afin de réussir une campagne de désinformation, il est préférable de sélectionner le canal médiatique ayant le meilleur audimat pour le saturer avec la fausse information. Une seconde option qui peut être couplée à la première est de limiter les interactions inter-personnelles, par exemple en coupant les vecteurs tel qu'internet ou en obstruant les propos défavorables à l'acteur (la véritable information) par le biais d'une surveillance totale.

Plus généralement, nous venons d'explicitier la capacité de notre modèle à reproduire des scénarii de désinformation que l'on peut observer dans certaines zones géographiques.

La seconde observation est contre-intuitive : nous nous attendions à ce que l'adoption s'élève rapidement avec un réseau plus connexe. En réalité, plus les voisins sont nombreux, plus les interactions se multiplient, plus ceux qui croient à l'information témoinée s'exposent à se faire convaincre du contraire. C'est pourquoi nous obtenons un effet de seuil au-delà duquel l'adoption augmente brutalement, aux contours de la diagonale $\pi_0 = 1$, $\pi_1 = 1$. Sur la base de ces observations, nous décidons de définir $\pi_0 = 0.9$ et $\pi_1 = 0.75$ pour nos prochaines simulations.

CONCLUSIONS DES DEUX EXPÉRIENCES DE DÉSINFORMATION Nous avons démontré par ces expériences que notre modèle est capable de retranscrire le phénomène de la propagation de l'information et de la persuasion qui dépend (1) de la crédibilité de la source de l'information et (2) du rapport entre le nombre d'échanges d'informations inter-personnelles et du nombre de communications liées à la diffusion par l'acteur (nombre de diffusions \times portée de la diffusions).

4.5 BILAN DE L'ANALYSE FONCTIONNELLE DU MODÈLE

Dans les premières expériences, nous avons montré l'influence de la composante émotionnelle et celle de la surprise sur la dynamique des attitudes d'une population soumise à des actions positives et négatives. Ces résultats nous ont

permis d'établir trois profils cognitifs définis par des couples de paramètres α et ξ .

Nous avons illustré l'impact de la dimension sociale sur les attitudes dans des simulations mettant en scène différents groupes sociaux. Ces études ont notamment permis de démontrer la capacité de notre modèle à générer des dynamiques d'attitudes qui tiennent compte des rivalités entre les groupes sociaux ainsi que la notion de distance sociale entre deux individus dans le graphe d'interaction.

Les études sur le mécanisme de communication inter-personnelle ont mis en évidence sa faculté à reproduire des phénomènes de diffusions d'informations au sein de populations, telles que nous pouvons témoigner dans la société (ex. : [Rogers \(2010\)](#)). D'autre part, ces expériences ont montré que la portée ainsi que la vitesse de propagation des messages dépendent des deux attributs (payoff et date) des actions relatives.

Enfin, les dernières expériences simulant une campagne de désinformation ont fait émerger plusieurs qualités de notre modèle. Le modèle de révision de croyance permet aux agents de ne pas être leurrés par les fausses informations diffusées par un acteur. Cette résistance a cependant une limite qui dépend d'une part des probabilités de révision de croyance et, d'autre part, du rapport entre le nombre de transmissions inter-personnelles et le nombre de diffusions émises par l'acteur. Ce phénomène a été illustré en employant différentes topologies de réseau social avec un nombre de voisinages moyen croissant. Par ailleurs, nous avons pu observer le phénomène de *conformité* par lequel l'individu témoin pouvait se laisser convaincre du contraire par ses voisins.

Ces premières études fonctionnelles ont permis de s'assurer que le modèle Polias était capable de décrire un certain nombre de phénomènes liés aux attitudes ainsi qu'à la communication. Une seconde contribution de ces expériences a été d'apporter une analyse de sensibilité du modèle à ses paramètres. Cependant, ces simulations sont toutes basées sur des scénarii générés de manière artificielle et aléatoire. Bien que cela permette de témoigner une certaine solidité du modèle, ces études ne font qu'esquisser l'applicabilité de notre approche sur des cas réels. Dans le prochain chapitre, nous exposerons des résultats expérimentaux de notre modèle appliqué à un cas d'étude réel.

Dans le chapitre précédent, nous avons étudié le comportement du modèle Polias en faisant varier ses paramètres sur des scénarii abstraits. Ces analyses nous ont permis d'une part d'analyser l'influence des paramètres sur les sorties du modèle et, d'autre part, d'illustrer les phénomènes descriptibles dans le cadre de Polias. Les phénomènes qui ont émergé dans ces études tels que l'habituation aux stimuli, les similarités d'évolutions d'attitudes entre les groupes sociaux, la courbure en S de la diffusion d'information ou encore la résistance à la désinformation et le processus de conformité sont des comportements qui peuvent être observés dans la société en sciences humaines et sociales (ex. : [Crutchfield \(1955\)](#); [Rogers \(2010\)](#)).

Cependant, ces travaux nous disent peu quant à la capacité du modèle à reproduire, voire à décrire, des phénomènes réels quantifiés. Nous rappelons que notre approche de modélisation consiste à se référer systématiquement, dans la mesure du possible, à des données précises que nous pouvons trouver dans la littérature. C'est dans ce sens que chaque partie du modèle cognitif de nos agents se fonde sur des théories issues de la psychologie sociale ou des sciences cognitives (voir la section [2.3](#)). En poursuivant cette approche, nous tentons de reproduire des phénomènes sociaux concrets et mesurés en appliquant notre modèle sur un cas d'étude précis. Nous pouvons qualifier cette méthode de lien macroscopique (correspondance entre les comportements du modèle au niveau global et les phénomènes réels quantifiés) ainsi que celui entre le modèle cognitif et les théories en sciences sociales en tant que lien microscopique ([Squazzoni \(2008\)](#)). Afin d'y parvenir, notre modèle a besoin de deux types de données : (1) une séquence d'actions datées, caractérisées par des payoffs, exécutées par un ou plusieurs acteurs sociaux impactant une population et (2) des données représentant, ou donnant un aperçu, de l'évolution des attitudes de la population vis-à-vis de cet ou de ces acteurs sociaux sur la même période.

Dans ce chapitre, nous appliquons notre modèle au contexte des opérations militaires de stabilisation. Notre cas d'étude porte sur l'intervention des Forces françaises en Kapisa (Afghanistan) entre février 2011 et septembre 2012. Les données relatives à ce cas d'étude sont : (1) les actions des Forces de stabilisation (Forces Bleues) ainsi que les *actions asymétriques* des Talibans dans la région sur la même période ; (2) les résultats de sondages sur la population concernée vis-à-vis de ces deux Forces. Nous présentons dans un premier temps le nouveau contexte militaire des *menaces asymétriques* dans lequel s'inscrivent ces opérations ainsi que ses conséquences sur les nouvelles approches militaires occidentales (section [5.1](#)). Par la suite, nous décrirons nos données relatives à notre cas d'étude : la méthodologie de leurs récoltes ainsi que leurs analyses (section [5.2](#)). Enfin, nous détaillerons le processus de calibration du modèle sur la base de ces données ainsi que les résultats au niveau macroscopique (agrégé au niveau de la population) et microscopique (au niveau des individus) de la simulation (section [5.3](#)).

5.1 CONTEXTE MILITAIRE : OPÉRATIONS DE STABILISATION

Les opérations de stabilisation qui ont eu lieu ces vingt dernières années (*e.g.* ex-Yougoslavie ou en Afghanistan), ont modifié le paradigme de la guerre, de celui des guerres industrielles à celui de « *conflits armés perpétrés au sein des populations* » [Smith \(2007\)](#). Aujourd'hui, la majorité des opérations militaires impliquant les Forces occidentales fait face à des opposants *irréguliers*.

LES MENACES ASYMÉTRIQUES Ces nouveaux types d'opposants (ex. : insurgés, guérillas, émeutiers ou terroristes) pratiquent des tactiques *asymétriques* (bombes artisanales, embuscades, prises d'otage, menaces etc.) et se fondent dans « *l'environnement humain* » où ils peuvent trouver certains appuis selon les affinités ethniques, politiques ou religieuses de certains individus. Par exemple, les Talibans, qui interviennent notamment dans notre cas d'étude, sont relativement faibles dans un contexte de guerre conventionnelle face à des armées régulières occidentales. De ce fait, ils évitent les conflits directs et cherchent des approches où leurs atouts sont adaptés aux faiblesses de leurs adversaires. De plus, les particularités locales telles que les coutumes, la mentalité (les attaques suicides représentent un exemple de tactique asymétrique qui fait appel à un état d'esprit très particulier) ou le passé historique (les anciens conflits de la région) peuvent causer des plans d'action singuliers. Notamment, ce type d'opposant procède à des actions asymétriques ciblant la population. En plus des conséquences directes, ces actions ont des objectifs sous-jacents tels que l'intimidation, la dissuasion, la démonstration de force, ébranlement de l'autorité des Forces occidentales ou encore forcer l'obéissance de la population. Ces tactiques, dont les victimes directes sont les civils, augmentent notamment les risques d'effets secondaires sur le conflit. En effet, une population désespérée devient instable, rendant la situation plus encline à une perte de contrôle. Une autre stratégie consiste à manipuler la population afin qu'elle soit contre la Force occidentale et supporte les insurgés en matériel, en informations ainsi qu'en mains d'œuvre. Il est donc important pour les Forces de stabilisation de protéger et surtout de suivre les actions et leurs conséquences sur l'état mental de la population.

L'APPROCHE GLOBALE Les pays occidentaux ont dû réviser leurs approches militaires pour faire face à ces nouvelles menaces. Ce fut notamment le cas pour la France qui a publié en 2008 le « *Livre blanc sur la défense et la sécurité nationale* » avançant que les interventions militaires conventionnelles, bien qu'elles soient toujours essentielles, ne sont qu'un outil parmi d'autres dont a besoin la Communauté Internationale pour la résolution des conflits. Ainsi, ce document appuie le besoin d'une *approche globale*, capable de s'intégrer dans un plan d'actions multinational ainsi que sur plusieurs niveaux stratégiques. Les objectifs développés dans l'approche globale consistent à empêcher ou résoudre rapidement et durablement une crise en coordonnant les actions de tous les départements impliqués et déployés sur le terrain afin de rétablir la gouvernance, la sécurité, l'économie et le développement social de la zone. Cette approche a été particulièrement conçue pour la stabilisation des zones ayant vécu une crise provoquée par des menaces asymétriques. Dans ce contexte, l'armée française a développé la doctrine de contre-rébellion. Elle concerne le niveau tactique des opérations et insiste sur l'importance des **actions de communication qui peuvent altérer**

les perceptions, les attitudes et les comportements au sein de la population ainsi qu'entraver les dynamiques pro-insurgées. Le succès d'une opération ne passe plus seulement par l'occupation d'un territoire, mais par la capacité à « *gagner les cœurs et les esprits* » de la population et de ses éléments favorables aux insurgés.

NOUVEAUX BESOINS : SIMULATEURS POPULATION-CENTRÉS Les systèmes de commandement et d'entraînement militaire actuellement mis en œuvre ont été conçus dans une optique d'opérations conventionnelles (guerres classiques inter-étatiques dites *symétriques* par opposition). Ils ne sont pas centrés sur la population civile et sont mal adaptés à capturer ou à simuler des phénomènes sociaux tels que l'émergence de groupes pro-insurgés. Plus précisément, les unités spécialisées dans les actions *non-cinétiques* (actions ne faisant pas usage de la force telles que les opérations civilo-militaires ou de communication) ne disposent d'aucun outil numérique leur permettant de planifier ou d'évaluer leurs opérations. De même, ils sont dépourvus de systèmes d'entraînement informatisés. De plus, les unités de combat conventionnelles ainsi que leur commandement manquent d'expertise concernant le terrain humain lors de la planification et la mise en œuvre des opérations à approche globale. Il existe néanmoins quelques projets industriels basés sur la technologie multi-agent simulant des phénomènes sociaux dans des cadres d'insurrection afin de répondre à ces nouveaux besoins. Ces travaux étudient diverses problématiques telles que la prédiction de l'émergence de conflits irréguliers dans certaines zones du globe, l'entraînement des soldats afin de mieux appréhender la population ou encore la modélisation de l'économie souterraine dans les zones de conflits *e.g.* Geller et Alam (2010); Silverman et collab. (2009); Alt et collab. (2009). Cependant, le besoin crucial de modéliser la dynamique perception-attitude-comportement de la population face aux actions des Forces (et en particulier face aux actions de communication) demeure non résolu. Or, la première partie de ce problème correspond à la formation et à la dynamique des attitudes. Le modèle Polias intervient sur ce point précis : simuler l'évolution des attitudes vis-à-vis des Forces en fonction de leurs actions dans un objectif de formation des militaires spécialisés sur les questions d'influences.

Dans cette section, nous avons présenté le changement de paradigme militaire caractérisé par l'émergence des menaces asymétriques et l'adaptation des approches occidentales qui en résulte. C'est dans ce contexte que s'inscrit notre cas d'étude sur les opérations de stabilisations française en Kapisa. Dans la prochaine section, nous décrivons ce cas d'étude et les données que nous avons récoltées.

5.2 UN CAS D'ÉTUDE : L'INTERVENTION FRANÇAISE EN AFGHANISTAN

Suite aux attentats du 11 septembre, les États-Unis se sont engagés dans une « lutte anti-terroriste », à commencer par une intervention militaire en Afghanistan à partir d'octobre 2001. Cette guerre avait initialement pour but de renverser le régime Taliban, allié direct d'Al-Qaeda.

En parallèle, les Nations Unies ont mandaté l'OTAN à s'engager en Afghanistan avec pour objectif d'éviter une nouvelle guerre civile ainsi que de stabiliser

le pays durant la reconstruction de l'Etat Afghan. Suite à la résurgence des offensives talibanes en 2006, les Forces Alliées de Sécurité Internationales (ISAF) menées par l'OTAN ont dû reprendre leurs opérations contre les Talibans en appui des Forces américaines. Une des contributions de la France en tant que membre de l'ISAF a été d'assurer la sécurité dans les régions de Kapisa et Surobi entre 2008 et 2012.

C'est dans ce contexte que les membres du Centre Inter-Armée des Actions sur l'Environnement (CIAE) en charge des actions civilo-militaires et des OMI ont été déployés sur le terrain d'octobre 2010 à septembre 2012. Dans le cadre de l'approche globale, comme mentionné précédemment, leurs actions étaient complémentaires aux actions conventionnelles de sécurisation : préparer l'environnement humain en amont des opérations (ex. : prévenir la population, l'évacuer etc.); appuyer le déroulement des opérations (ex. : participer aux patrouilles pour entrer en contact ou maintenir les affinités avec la population); se charger des conséquences des opérations (ex. : les promouvoir en cas de succès, contenir les effets négatifs sur l'image de la Force en cas d'échec).

Nous avons appliqué notre modèle Polias sur une période de cette intervention française en Kapisa dans le but de reproduire les résultats de sondages afin de valider le comportement du simulateur ainsi que pour analyser les causes de cette évolution. Pour cela, nous avons reconstitué avec la participation du CIAE un scénario militaire décrivant les actions françaises ainsi que celles des Talibans. De plus, nous avons récolté des résultats de sondages sur l'opinion des habitants de la Kapisa vis-à-vis des deux Forces sur la période correspondant au scénario. Le but de l'expérience consiste à reproduire de manière artificielle ces sondages d'opinions à partir du scénario établi en appliquant le modèle Polias. Dans les prochaines sections, nous décrivons ces deux types de données : la manière dont ils ont été recueillis ainsi que leurs interprétations.

5.2.1 *Reconstitution du scénario militaire*

Toutes les opérations et actions menées par une armée nationale sont archivées à des fins évidentes telles que le retour sur l'expérience, les éventuelles enquêtes *a posteriori* etc. Cependant, pour une raison aussi triviale que la première, ces informations ne sont pas accessibles au public de par leur caractère classifié. Néanmoins, une grande partie de ces informations ne sont pas secret Défense. Malheureusement, elles ne sont pas archivées sur différents supports en fonction de leur degré de classification. Ainsi, le seul moyen de reconstituer un scénario militaire décrivant les actions non-classifiées ou plutôt les informations non-classifiées sur les actions (l'existence même de certaines opérations est tenue secrète, cependant beaucoup d'autres ont un contenu classifié mais leur existence peut être mentionnée) a été de les récolter auprès des officiers en charge de ces actions.

En suivant cette approche, nous avons effectué quatre interviews auprès des officiers du CIAE situé à Lyon qui étaient en charge de leur cellule située en Kapisa au cours de la période octobre 2010 à septembre 2012. Chacun de ces officiers nous a résumé les opérations qu'il a dirigées durant son tour d'opération. Un tour d'opération correspond à une durée pendant laquelle les officiers en charge de la cellule ont été déployés sur le terrain. Cette durée varie en fonction de la saison, environ 7 mois en été et 5 mois en hiver. L'avantage de cette enquête auprès de ce centre est double : (1) les opérations d'influences font partie des



FIGURE 23 – Carte de la Kapisa, une province de l’Afghanistan peuplée d’environ 360 000 habitants selon une densité de 195 habitants/km² principalement situés dans des villes et villages sur l’axe de Vermont et dans les vallées (voir la figure 24)

actions les plus importantes dans les conflits asymétriques (voir le §5.1) ; (2) de par leur fonction de complémentarité aux actions conventionnelles, les membres du CIAE sont autant informés de ces opérations que leurs exécutants (le CIAE participe physiquement à pratiquement toutes les opérations exécutées par les autres départements qui ont un impact sur la population).

Le résultat de ces interviews prend la forme d’un scénario présenté dans la figure 25 à la page 90. L’échelle de temps dans cette figure est la semaine, bien que dans la simulation nous simulons l’évolution du temps selon un pas de temps quotidien. La raison originelle de cette granularité hebdomadaire vient du fait que les officiers ont dû se baser principalement sur leurs mémoires et quelques notes personnelles (les interviews ont été effectuées en 2014, soit environ deux ans après les opérations) limitant ainsi la précision des dates. Certaines opérations importantes ainsi que les attentats talibans sont précisément datés car ils ont été mentionnés par la presse internationale.

Dans un premier temps, nous définissons les actions constituant le scénario en partageant celles effectuées par les Français et celles des Talibans. Par la suite, nous décrirons l’évolution du scénario par tour d’opération, tel que nous l’avons récolté.

5.2.1.1 Définitions des actions

Dans cette section nous décrivons les actions que nous avons récoltées et codées dans nos simulations. Dans le cadre de notre modèle, nous avons besoin de définir les actions par : une date ou une fréquence d’occurrence, un payoff par bénéficiaire

(victime si payoff négatif) de l'impact et une portée. La portée d'une action correspond ici au pourcentage de la population totale qui est directement affectée par occurrence de l'action.

Les payoffs sont définis de manière arbitraire. En effet, notre modèle ne définit aucune échelle ou unité pour l'attitude : les attitudes peuvent être infiniment grandes ou négatives. Par conséquent, la définition des valeurs des payoffs n'est importante qu'en comparaison entre les actions. C'est la distance entre les payoffs des actions qui leur donne un sens. Par exemple, définir le payoff d'une patrouille et d'une écoute de la radio à respectivement +30 et +10, au lieu de +3 et +1 n'aura qu'un impact de grandeur d'échelle. Les attitudes qui en résultent seront des valeurs différentes mais ce ne sont pas leurs valeurs qui nous intéressent mais la relativité entre les attitudes (une patrouille a trois fois plus d'impact qu'une émission radio).

Néanmoins, nous devons convenir que la définition de ces relativités reste problématique. Par exemple, nous pensons qu'il est impossible de définir de manière absolue l'impact de la mort d'un individu en fonction de la valeur d'impact d'une émission radio. Malgré cette contrainte, nous avons défini, avec la participation du CIAE, des valeurs d'impact pour chacune des actions en s'assurant que la hiérarchie (la relation d'ordre) entre les actions soit cohérente.

Lorsque les actions sont qualifiées d'occasionnelles, leurs dates d'occurrences ont été soit communiquées par le CIAE, soit basées sur des articles de la presse internationale. Aucune de ces dates n'est arbitraire. Les nombres de victimes des actions négatives sont définis en fonction de ce qui nous a été relaté lors des interviews ainsi que par la presse.

ACTIONS DE LA FORCE BLEUE

- Radio : l'armée française a mis en place une station de radio destinée au grand public. La radio Omid, signifiant « espoir » en pachto et dari (les deux langues locales principales dans cette région), a été diffusée dans ces deux langues durant toute la durée de l'intervention. Elle est animée par des intervenants locaux, ses émissions portent sur les faits d'actualité, la musique, la culture et la religion (intervention de figures islamiques). Son but est d'informer et de divertir la population locale comme toutes radios conventionnelles. Cependant, elle a aussi une fonction opérationnelle : faire de la contre-propagande talibane et améliorer la perception des Forces alliées. Dans la simulation nous lui avons défini une fréquence quotidienne, une portée de 10%, un payoff de +10 ;
- Patrouilles : ce sont des opérations de niveau tactique menées par les équipes terrestres. Mis à part leur fonction de sécurisation et de dissuasion, elles sont aussi l'occasion pour les Français de nouer des liens avec la population, de « prendre la température » sur le terrain, d'être au courant des faits locaux, d'être informé des besoins particuliers, de faire des distributions (ex. : vivres, jouets, postes de radio etc.). Leur payoff a été défini à +30, une fréquence variable selon les tours d'opération (de 1 à 3 occurrences par semaine), une portée de 1% (correspondant à peu près à 10% d'un village) ;
- Reconstruction : ces actions englobent les opérations de reconstructions de bâtiments civils, l'aménagement de puits pour l'approvisionnement en eau potable, installation de divers équipements, le dédommagement financier des dégâts causés par les Talibans ou les Français (ex. : dans certains cas

- les Français ont dû, avec le consentement de la population, détruire des murs ou élargir des rues afin de faire passer des véhicules blindés). Nous avons défini un *payoff* de +10, une fréquence d'une occurrence par semaine et une portée de 1% ;
- Opération de présence : elle n'a lieu qu'une seule fois lors du premier tour d'opération. Elle consiste en une présence permanente d'une unité française dans un ou plusieurs villages sur une durée relativement longue (de l'ordre de quelques semaines). En temps normal, les soldats sortent rarement de leurs bases militaires pour des raisons de sécurité. Durant ces opérations de présence, les soldats restent et dorment dans les villages. Elles ont pour but d'occuper physiquement le terrain pour prouver leur suprématie par rapport aux Talibans ainsi que de prouver leur bonne volonté et de resserrer les liens avec la population. Ces opérations ont généralement un impact important sur la perception des Forces. Ces actions ont une fréquence quotidienne durant la durée de l'opération, un *payoff* = +20 avec une portée de 1% (au bout de 10 jours, tout le village aura témoigné de cette opération) ;
 - Tracts (leaflets) : ce sont des largages de tracts par avion sur des zones spécifiques (niveau village). Ils servent à informer la population sur un fait particulier. Leur impact sur la population est relativement élevé car les individus sont impressionnés par la précision de leurs mises en œuvre. *payoff* = +30, *reach* = 1%, occasionnels ;
 - Arrestation : arrestations de Talibans. Nous n'avons pu en répertorier qu'une seule sur la période étudiée. *payoff* = -50 sur les Talibans et la portée de 0.05%, elle concerne ici le nombre de civils qui ont directement témoigné de l'action ;
 - Opération de niveau brigade : elles font référence aux opérations de grande ampleur qui ont principalement pour objectif de débusquer les Talibans dans leurs zones ou leurs camps militaires. Ces opérations résultent en des neutralisations des opposants. Parfois, bien que ce ne soit pas le cas sur notre période étudiée, des dommages collatéraux sur les civils peuvent survenir. Ces opérations sont effectuées avec la participation de tous les départements présents sur le terrain. Leurs impacts sont extrêmement négatifs pour les Talibans : *payoff* = -100. Elles sont témoignées directement par 0.1% de la population et sont occasionnelles ;
 - Opération spéciale : leurs objectifs sont similaires à celles des opérations de niveau brigade mais elles sont effectuées de manière plus discrète (assauts de nuits, équipement furtifs ...). Paradoxalement, elles sont les plus relatées au sein de la population qui les trouvent particulièrement impressionnantes. Nous avons fait le choix de retranscrire cette particularité en augmentant la portée de l'action par rapport aux opérations de brigades : *payoff* = -100 pour les Talibans, *reach* = 0.2%, occasionnelles ;
 - Légations de COP/FOB : les COP et FOB (Combat Ouptpost et Forward Operation Base) sont des avant-postes et des bases militaires avancées contrôlés par les Français situés en avant-poste. Ils permettent d'apporter de la sécurité dans la zone ainsi que de contrôler les allers et venues des opposants. Leurs légations consistent à donner leurs commandements aux militaires afghans. *payoff* = 50, *reach* = 0.1%, occasionnelles.

- Pressions : ces actions englobent toutes les actions oppressant la population locale telles que le racket, les menaces, les brimades (coupure d'eau ou d'électricité), les actes de violences (ex. : tabassage). $\text{payoff} = -10$, $\text{reach} = 0.1\%$, quotidien ;
- Prêches : prêches religieux dans les lieux publics. $\text{payoff} = +15$, $\text{reach} = 0.05\%$, quotidien ;
- IED : explosions de mines anti-personnelles ou anti-char fabriquées à la main. Elles peuvent causer des dommages à la population ou aux Forces Bleues (les deux cas sont distingués selon le scénario). $\text{payoff} = -100$, $\text{reach} = 0.005$, occasionnel ;
- Attaques suicides : attaques kamikaze à la bombe, parfois avec l'usage d'un véhicule piégé. $\text{payoff} = -100$, $\text{reach} = 0.01$, occasionnelles.
- Green on Blue : attaques menées par des soldats afghans sur les Forces alliées. Ces soldats sont souvent des adhérents à la cause talibane qui se sont infiltrés, parfois durant plusieurs années avant de commettre leurs crimes. Dans la majorité des cas, le soldat se suicide après son action. $\text{payoff} = -100$ sur la Force alliée, $\text{reach} = 0.1\%$.

Nous avons défini de manière statique les actions que nous voulons simuler. Dans la prochaine section, nous présentons l'évolution de ces actions qui constitue notre scénario.

5.2.1.2 *Evolution du scénario au cours des tours d'opération*

L'intervention militaire française dans les régions de Kapisa et Surobi se constitue de six mandats de l'ISAF, effectués par la brigade La Fayette, se succédant de fin 2009 à fin 2012. A chaque mandat correspond un tour d'opération dirigé par un officier, dont le grade varie entre commandant et colonel, déployé sur le terrain. Ils ordonnent les opérations tactiques de leurs unités appartenant à la cellule OMI en se coordonnant avec les autres départements appartenant à la même brigade, mais aussi avec les autres intervenants tels que l'armée américaine ou encore des organisations non-gouvernementales.

Le découpage de notre scénario se base sur les dates de sondages d'opinions notés P₁ (février 2011), P₂ (septembre 2011), P₃ (mars 2012) et P₄ (septembre 2012). En effet, nous devons mesurer les sorties du modèle aux dates correspondant aux sondages afin de les évaluer. Or, les dates de sondages ne correspondent pas aux transitions de mandats. De plus, le premier sondage a été effectué durant le troisième mandat. Par conséquent, notre scénario porte sur les quatre derniers tours d'opérations. La première partie de notre scénario démarre au cinquième mois de la brigade La Fayette 3 (février 2011) et se termine au dernier mois de la brigade La Fayette 4 (septembre 2011), soit un total de 8 mois.

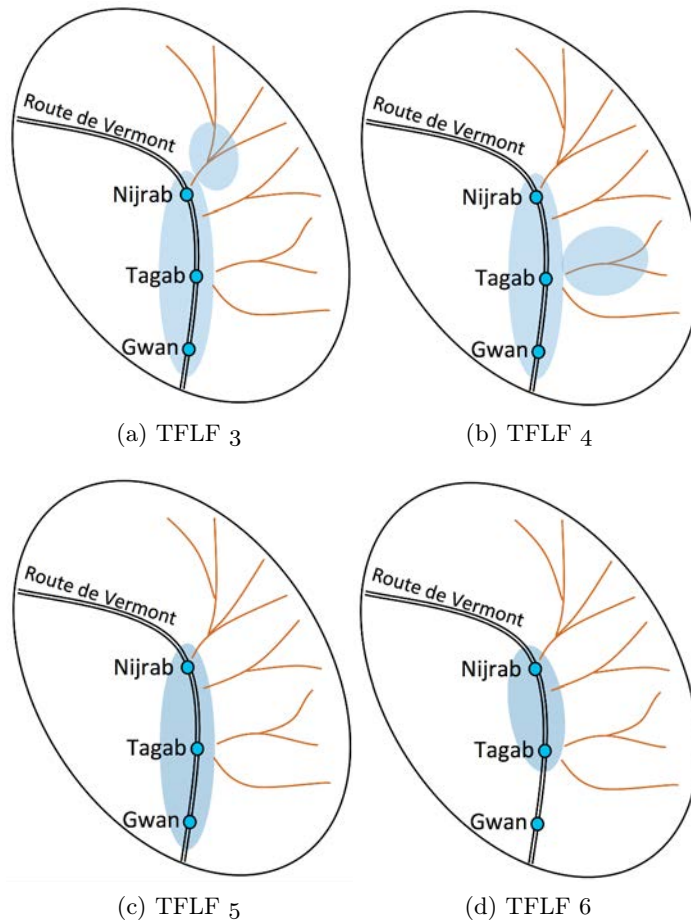


FIGURE 24 – Schématisation de l'évolution des axes d'efforts de la Task Force La Fayette (TFLF) sur la période étudiée au fil des tours d'opérations - Les points bleus représentent les principales villes où des COP/FOB sont installés, les zones bleues désignent les axes d'effort et les lignes marron caractérisent les principales vallées.

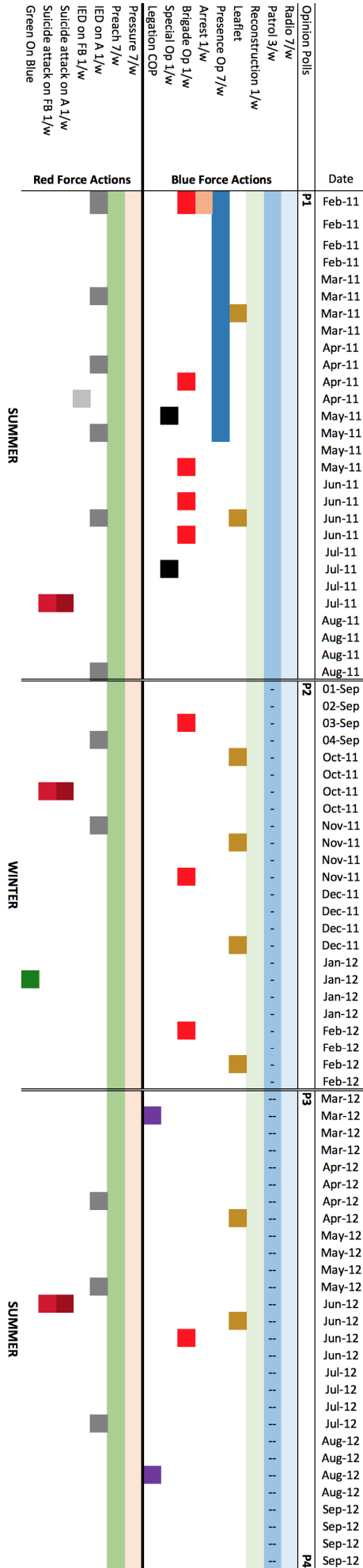


FIGURE 25 – Scénario militaire (actions des Forces Bleu et Rouge en Kapisa entre 2010 et 2011)

PREMIÈRE PHASE (P1 À P2) : PRÉPARATION ET ENGAGEMENTS L'axe d'effort de la brigade La Fayette 3 se focalisait sur le regain de la liberté d'action dans la vallée de Nijrab au détriment de Tagab qui est moins maillé (voir la figure 24 illustrant l'évolution des opérations). Les actions permanentes de la Force Bleue qui sont constantes à travers tous les mandats sont les patrouilles dans les villages, les émissions radio et la reconstruction. Il est à noter que durant ce mandat, la fréquence des patrouilles était particulièrement élevée, de l'ordre de 3 occurrences par semaine. La seconde moitié du mandat (début de notre simulation) est marquée par une opération de présence permanente au sein de quelques villages de la vallée de Nijrab. Cette opération profite de l'absence des leaders Talibans durant cette phase, facilitant ainsi le contact avec les habitants. L'hiver étant particulièrement rude, beaucoup de Talibans quittent le pays pour se replier au Pakistan. Un second facteur limite les activités talibanes durant cette saison. L'usage des IED (une des principales actions Rouge) décroît en hiver en raison du feuillage des arbres : la pose des explosifs se fait souvent à l'ombre des arbres pour éviter de se faire remarquer par les caméras de surveillance ainsi que par les drones.

L'axe d'effort de la brigade La Fayette 4 se recentre sur la vallée de Tagab. Les actions permanentes sont maintenues. Au retour des Talibans en mai 2011, la Force Bleue profite de sa bonne implantation dans la zone de Nijrab obtenue grâce aux efforts du précédent mandat pour les attaquer lourdement. En effet, ce tour d'opération a été le plus cinétique de tous : trois opérations de brigade à grande ampleur ont été menées sur trois mois de suite. C'est à l'issue de ces opérations qu'est survenu l'attentat de Joybar le 13 juillet 2011, selon certains, il s'agissait de représailles pour cette recrudescence des activités Bleues. L'attentat commence par l'explosion d'un kamikaze intervenant lors d'une choura¹ organisée dans le nord de Tagab, tuant 5 militaires français du GTIA Kapisa (Groupe Terrestre Inter-Armé de la brigade La Fayette 4), blessant grièvement quatre autres soldats et quatre civils. L'attentat suicide est suivi d'une attaque embusquée d'un groupe de Talibans armés de fusils d'assaut et de lance-roquettes anti-char. Suite à l'engagement de plusieurs unités terrestres en renfort, de deux hélicoptères français, des véhicules terrestres et d'un avion de chasse américain, les Talibans ont décroché sans commettre davantage de dégâts. Toutefois cet attentat a marqué les esprits, surtout chez l'opinion publique française, renforçant ainsi le souhait d'un retrait des troupes françaises. En conséquence, le Président de la République en date avait décidé d'accélérer le processus de retrait, impactant grandement la suite des actions Bleues sur le terrain.

SECONDE PHASE (P2 À P3) : VERS UN RETRAIT La seconde phase de la simulation démarre avec la brigade La Fayette 5. Cette fois, en conséquence directe de la décision prise par le gouvernement français suite à l'attentat de Joybar, l'axe d'effort consiste à préserver au maximum les vies françaises. Cela se traduit par une baisse générale de la fréquence des activités Bleues et surtout une sécurisation renforcée de toutes les opérations. Chaque sortie est accompagnée d'une unité d'artillerie lourde, de tireurs embusqués et d'engagements de plusieurs unités terrestres quadrillant les itinéraires. Autre impact, les contacts avec la population sont limités au maximum lors des patrouilles. Néanmoins, trois

1. Une choura correspond à un concile de juges musulmans. Ici, le terme fait référence, de manière plus générale, à une réunion faisant intervenir des personnalités importantes locales.

opérations de niveau brigades auront été effectuées durant ce mandat, toutes accompagnées d'un largage de tracts (opération à risque faible mais avec un impact élevé) pour informer la population des victoires de la Force Bleue. La fin du mandat est marquée par le premier (et le seul à notre connaissance) Green on Blue subi par la brigade La Fayette. Un soldat afghan (infiltré pro-Taliban) ouvre le feu dans une base militaire à Gwan, tuant 5 soldats français et en blessant 14 autres.

Comme pour l'année précédente, les Talibans sont moins actifs en hiver à partir de novembre 2011. Leurs actions de pression sur la population affectent moins de personnes, les IED diminuent.

TROISIÈME PHASE (P3 à P4) : RETRAIT Cette dernière phase concerne les derniers mois de la brigade La Fayette 5 et 6 avant le retrait de la brigade. Elle démarre avec une légation d'une COP au second mois suivie de la légation d'une FOB en août 2012. Parallèlement, les patrouilles sont d'autant plus réduites, beaucoup sont annulées en raison d'un manque de niveau de sécurité requis. Le niveau de sécurité ayant été relevé à la suite des dernières pertes humaines du Green on Blue. Une unique opération de brigade aura été menée.

De leur côté, les Talibans poursuivent leurs actions et les multiplient au printemps (quasiment autant que pendant la première phase). Les IED continuent à menacer la population, ils opèrent notamment un attentat suicide suivi d'une attaque sur une base de l'armée afghane. Elle ne causera aucune perte ni blessé du côté des militaires/forces de l'ordre afghans.

Nous venons de présenter le scénario militaire qui s'est déroulé en Kapisa sur l'intervalle étudié. Dans la prochaine section, nous exposons les conséquences de ces actions sur l'opinion publique de la population locale vis-à-vis des deux Forces.

5.2.2 *Sondages d'opinions*

Le projet « Suivi de l'impact de la présence française sur les perceptions de la population de Kapisa et de Surobi » a été commandité par le Ministère de la Défense auprès de la société de sondage Altai Consulting en novembre 2010. Ses objectifs étaient d'établir l'état des opinions des populations de Kapisa et Surobi et de suivre son évolution dans le temps. Il avait également d'autres objectifs plus analytiques, tels que la mise relation des causes et de leurs conséquences sur l'opinion ou encore l'émission de recommandations pour corriger les éventuelles erreurs des militaires français. Cependant, les documents produits au cours de ce projet étant classifiés, nous n'avons pas pu avoir accès à l'intégralité du contenu du projet. Dans les faits, nous n'avons pu obtenir que les supports de présentations (planches de type PowerPoint) résumant l'avancée des travaux après chaque phase. Nous rappelons que les phases de ces sondages correspondent aux phases constituant notre simulation, comme présentées précédemment.

INTERVENANTS Le dispositif du projet a été développé et mis en place par Altai Consulting et le directeur général adjoint en date de l'institut de sondage CSA. La société a été fondée en 2003 à Kaboul par trois Français, elle comptait en 2010 plus de 15 consultants internationaux, 40 chefs de projets locaux et 50

enquêteurs locaux. L'équipe a travaillé pour la plupart des acteurs internationaux présents en Afghanistan : la Commission Européenne, les agences des Nations Unies, l'AFD, le gouvernement britannique, l'OTAN/ISAF, USAID, la Banque mondiale etc.

THÉMATIQUES Le travail d'enquête s'appuie sur des interviews à domicile. Un questionnaire d'une durée d'environ 45 minutes a été élaboré en lien avec différentes équipes des Forces françaises. Il couvre les thèmes suivants :

- Profils socio-démographiques et niveau de vie ;
- La reconstruction et le développement ;
- La gouvernance (affiliation politique, attentes aux niveaux des provinces et districts) ;
- La sécurité et la perception des différentes Forces ;
- Les sources d'informations.

MÉTHODOLOGIE Les enquêtes ont été menées sur un échantillon de 720 individus pour une taille de population de 364 900 habitants en Kapisa. La méthode employée repose sur des enquêtes terrain avec interviews en face à face à domicile, avec sélection de zones d'enquête et de foyers par tirage au sort. Le tirage au sort se base sur un protocole de tirage proportionnel à la taille des populations respectives en se basant sur les données du dernier recensement disponible à l'époque : CSO 2003.

Pour des raisons de sécurité, certaines zones d'enquête, notamment dans les districts de Tagab et Alasai, n'ont pas pu être sondées au cours de toutes les phases. La situation sécuritaire à Alasai ayant été particulièrement défavorable, l'équipe a pris contact avec les habitants des villages voisins pour les faire venir dans le chef-lieu afin de procéder aux recueils. Les données présentées dans cette étude ont été pondérées afin de prendre en compte les écarts de population entre les différents districts. A titre d'exemple, l'intervalle de confiance à un degré de confiance de 95% est de +/-3.65%.

Au sein de chaque zone d'enquête, les équipes sont parties des mosquées et ont sélectionné des foyers toutes les six maisons et ainsi de suite, rue par rue, à partir de leurs points de départ. La sélection des individus a été effectuée en appliquant les grilles de Kish (1965) afin de maintenir le principe aléatoire au niveau du foyer. Seuls les adultes de plus de 18 ans ont été interrogés.

RÉSULTATS DE SONDAGE Bien que l'étude ait engendré des résultats très affinés au niveau des villages, les documents sur lesquels nous nous basons ne nous permettent pas de reconstituer leurs évolutions : les zones présentées ne sont pas consistantes dans les résumés que nous détenons. Par ailleurs, les questions posées dans les formulaires n'ont pas été constantes au cours des phases. Beaucoup de questions posées durant la première phase ont disparu au cours de l'étude, certaines sont réapparues d'autres ont été ajoutées en cours de route.

Parmi toutes les questions, nous avons retenu deux questions qui ont été posées à chaque phase, aux mêmes échantillons durant toute la durée de l'étude et en lien avec notre problématique :

- Q1 : « Est-ce que les insurgés sont le principal vecteur d'insécurité ? »
- Q2 : « Les Forces étrangères participent-elles à la sécurité ? »

Les réponses se mesurent en pourcentage d'approbation : le rapport entre le nombre d'individus ayant répondu positivement à la question et l'échantillon total. La table 5 montre les résultats de sondages à ces questions.

Questions	Polls Dates			
	2/11(P1)	9/11(P2)	2/12(P3)	9/12(P4)
"The Blue Force contribute to security" (Q1)	40	32	24	19
"The Red Force is the principal vector of insecurity" (Q2)	27	60	27	37

TABLE 5 – % de la population ayant répondu favorablement aux questions à différentes dates

Afin de calibrer notre modèle sur la base de ces résultats, nous avons étendu le sens de ces questions :

- La favorabilité aux insurgés (Rouges) est déduite en prenant l'opposé du résultat de Q1 ;
- La favorabilité aux Forces Bleues est directement déduite des résultats à la question Q2.

En effet, notre scénario porte strictement sur des actions militaires et sécuritaires en adéquation avec la thématique de ces questions.

Nous sommes cependant conscients que des données de sondages ne sont pas les mieux adaptées pour estimer des attitudes (voir la section 2.1.3). De plus, la première question ne peut pas être directement transposée, même en prenant son opposé, à une attitude vis-à-vis des Rouges. En effet, il existe d'autres vecteurs d'insécurité tel que les chômeurs, les drogués ou même la police locale (groupes recueillis par Altai). Nous ne pouvons pas directement déduire les attitudes à partir de la position des Talibans dans le classement des vecteurs d'insécurité. Néanmoins, les données portant sur de tels cas d'application restent rares et classifiées, nous poursuivons notre application du modèle avec cette approximation des attitudes.

Les figures 26 et 27 montrent l'évolution de ces favorabilités qui constituent les objectifs du simulateur.

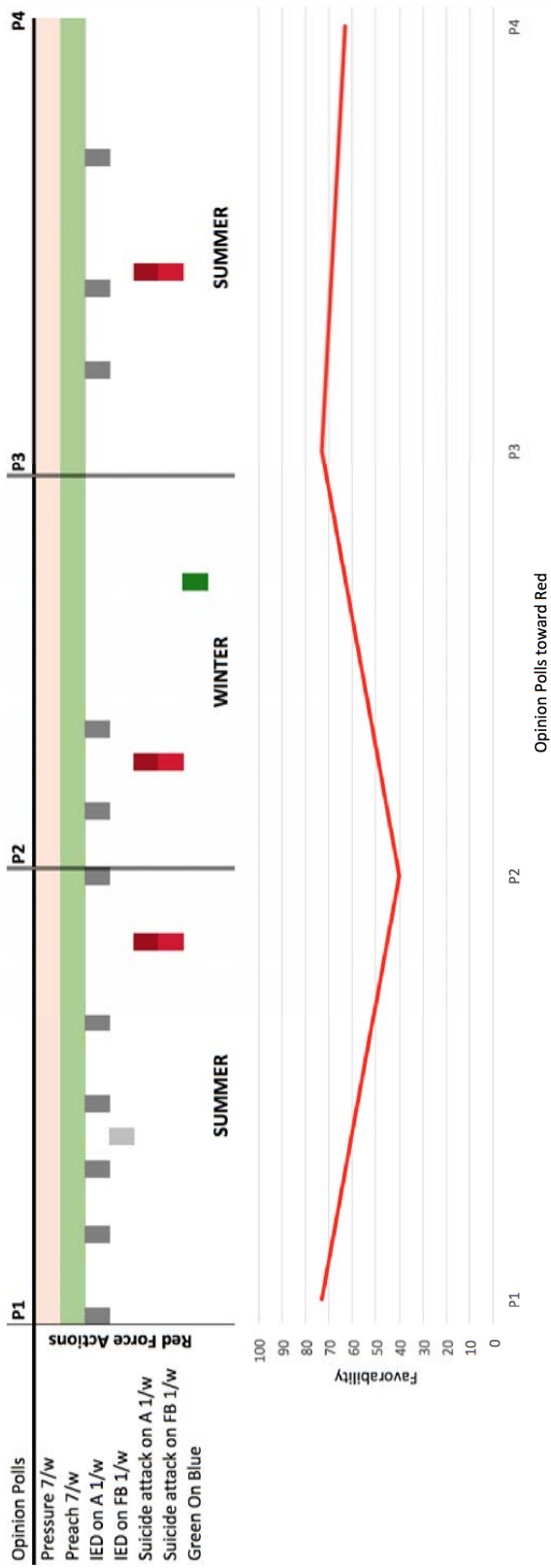


FIGURE 26 – Scénario militaire (actions des Forces Bleu et Rouge en Kapisa entre 2010 et 2011)

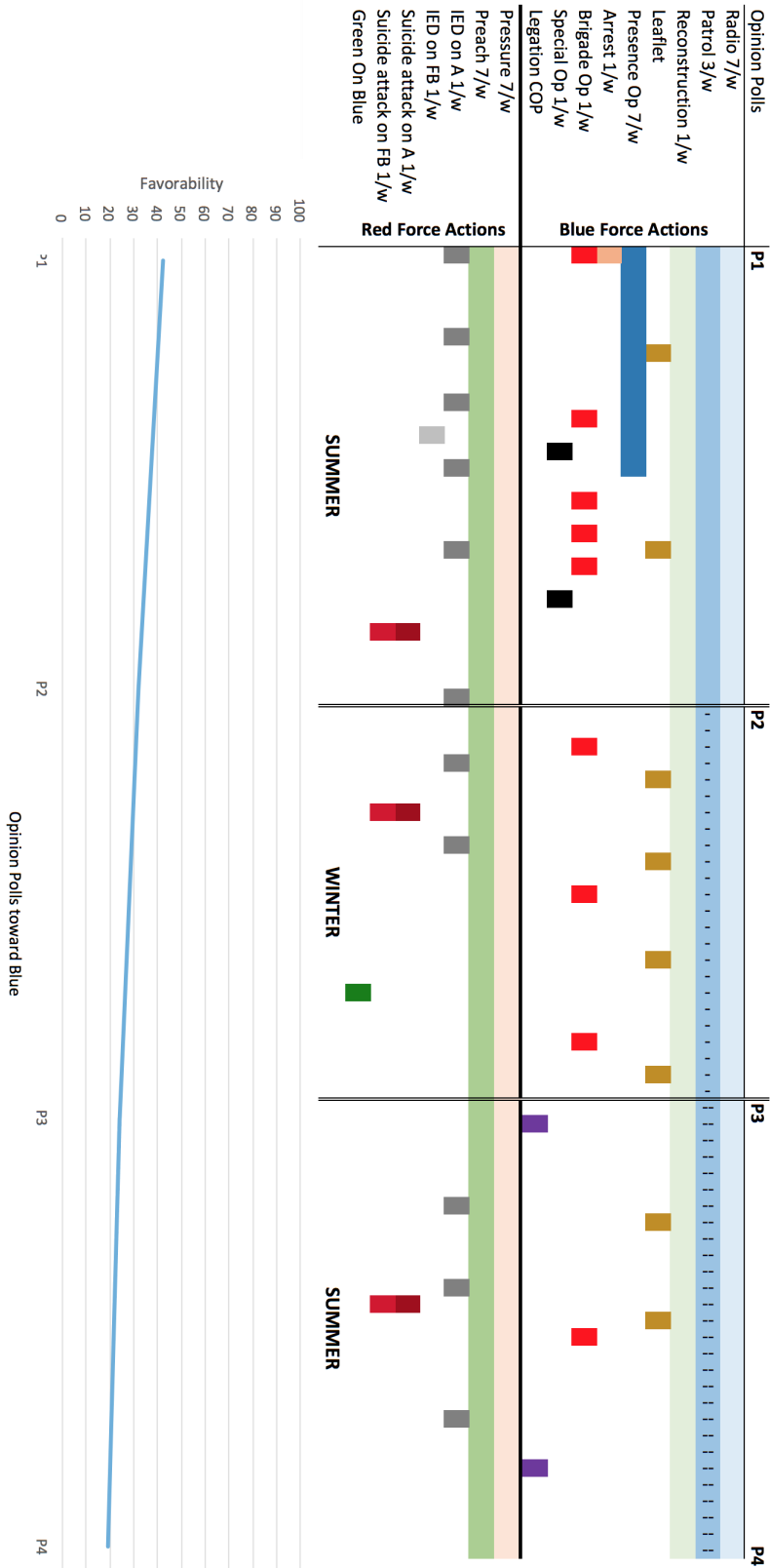


FIGURE 27 – Scénario militaire (actions des Forces Bleu et Rouge en Kapisa entre 2010 et 2011)

Nous avons présenté le scénario militaire des deux Forces antagonistes présentes sur le terrain, constituant les données d'entrées du simulateur, ainsi que les résultats de sondages d'opinions sur ces acteurs, représentant les objectifs du simulateur. Maintenant que nous avons réuni tous les éléments afin d'appliquer notre modèle aux opérations françaises en Kapisa, nous procédons à la calibration du modèle que nous détaillerons dans la prochaine section.

5.3 CALIBRATION DU MODÈLE ET RÉSULTATS MACROSCOPIQUES

La calibration d'un modèle est un processus automatique permettant de déterminer des valeurs de paramètres afin que les résultats de simulations parviennent à reproduire ou à se rapprocher des valeurs données en tant qu'objectifs. Dans cette section nous décrivons notre méthodologie de calibration pour le modèle Polias et les résultats obtenus. L'objectif consiste à simuler le scénario militaire établi précédemment et à faire correspondre les sorties du simulateur aux résultats de sondages que nous avons récoltés.

Dans un premier temps, nous décrivons l'algorithme de calibration que nous avons employé avant de présenter notre protocole de calibration. Enfin, nous exposerons et analyserons les résultats de calibration ainsi que ceux des simulations.

5.3.1 *Algorithme de calibration : CMA-ES*

Le CMA-ES (Covariance Matrix Adaptation Evolution Strategy) de Hansen et collab. (2003) est un algorithme évolutionnaire pour les problèmes d'optimisations non-linéaires, non-convexes dans un espace continu. Il est adopté en tant qu'outil standard pour l'optimisation continue.

Le principe général consiste à répéter des simulations avec différents jeux de paramètres, calculer une distance entre les résultats de la simulation et les objectifs (selon une fonction dite de « fitness ») et adapter l'espace des paramètres en fonction, avant de réitérer le processus. La calibration s'achève selon un critère d'arrêt défini par l'utilisateur, souvent dépendant d'un seuil de tolérance sur la fitness ou le nombre d'itérations durant lesquelles l'algorithme n'a pas pu améliorer la fitness. Notons que le CMA-ES est mono-objectif : dans le cas où plusieurs points sont à atteindre (notamment dans notre cas), il est nécessaire d'agrèger les distances dans la fonction de fitness.

L'algorithme se fonde sur deux principes piliers afin d'adapter l'espace de distribution des paramètres. Premièrement, le principe de la vraisemblance maximale permet d'augmenter la probabilité d'obtenir des solutions candidates et des étapes de recherches valides. La moyenne de la distribution d'un paramètre est mise à jour de telle sorte que sa vraisemblance (notion statistique qui peut se traduire par le contraire de la variance) avec la précédente génération de solutions candidates valides soit maximisée. De la même manière, la matrice de covariance de la distribution est mise à jour de manière à augmenter sa vraisemblance avec la dernière matrice valide. Ces deux mises à jour peuvent être interprétées comme une descente de gradient.

Secondement, deux évolutions temporelles de la moyenne des distributions sont enregistrées, elles sont appelées trajectoires d'évolution (ou traces). Ces trajectoires contiennent des informations sur la corrélation entre les étapes consécutives.

L'une des trajectoires sert à adapter le processus de mise à jour de la matrice de covariance en permettant d'accélérer la variance vers une direction favorable par rapport à une approche qui ne considérerait qu'une unique étape de recherche. L'autre trajectoire sert à exécuter un processus de contrôle de la vitesse de convergence permettant d'éviter les convergences prématurées et de ne pas stagner dans un optimum local.

Dans la prochaine section, nous décrivons la manière dont nous avons appliqué cet algorithme à notre problème de calibration.

5.3.2 *Protocole de calibration*

L'objectif de la calibration consiste à rapprocher au maximum les sorties de notre simulateur aux résultats de sondage mesurant la favorabilité de la population vis-à-vis des Forces aux instants P_1 , P_2 , P_3 et P_4 . Pour ce faire, nous avons implémenté le scénario militaire présenté dans la figure 25 dans le simulateur Polias. Chaque jour (chaque pas de temps), les deux Forces procèdent à leurs actions décrites dans le scénario qui dure 554 jours. La population est constituée de 500 individus, connectés par un réseau d'interactions de type small-world avec un voisinage moyen de 4 comme pour les expériences de l'analyse fonctionnelle.

INITIALISATION Avant de démarrer le scénario, nous procédons à une **phase d'initialisation** de la population en donnant un vécu personnel pour chaque individu dans le but d'atteindre les premiers points des objectifs (P_1). Le scénario étudié démarre lorsque (1) l'initialisation des croyances de la population est accomplie, (2) les attitudes envers les Forces sont initialisées telles que les favorabilités des population correspondent aux points P_1 des objectifs et (3) le système est stable (les attitudes ne bougent plus).

L'une des originalités de notre modèle réside dans le fait que les attitudes de nos agents dépendent de leurs états mentaux, caractérisés par des croyances, des évaluations et des accessibilités. En conséquence nous ne pouvons pas initialiser directement les attitudes des agents comme pour les simulateurs d'attitudes/opinions existants. Nous sommes contraints d'initialiser les croyances de nos agents selon une certaine distribution d'actions avec une variété de valeurs de payoffs afin que la population ait une attitude envers les Rouges et les Bleus. Ces croyances peuvent être interprétées comme les interactions entre la population et les Forces qui précèdent le début de la période simulée. Autre point sensible, les individus sont surpris lorsqu'ils témoignent d'actions complètement nouvelles, entraînant une surestimation de leurs impacts. Afin de les habituer aux actions régulières (telles que la radio, les patrouilles, les prêches etc.) nous avons besoin de les intégrer dans la période d'initialisation. Il est à noter que les actions de type brigade ou leaflet ne font pas partis de ce scénario d'initialisation car nous n'avons aucune donnée les concernant.

Afin d'atteindre les valeurs de favorabilité de P_1 (2), nous devons ajouter dans le scénario d'initialisation une action positive effectuée par la Force Rouge afin d'augmenter l'attitude de la population vis-à-vis d'elle que nous notons AIRP (Action d'Initilisation Rouge Positive). De la même manière, une action négative effectuée par la Force Bleue AIBN doit être intégrée dans l'initialisation. Ces actions permettent de contre-balancer les actions d'initialisation d'habituation (généralement négatives pour les Rouges et strictement positive pour les Bleus)

de la population aux actions régulières. AIRP et AIBN sont caractérisés leurs payoffs et leurs portées.

Cette période d'initialisation dure environ 200 jours, elle est décrite dans la figure 28.

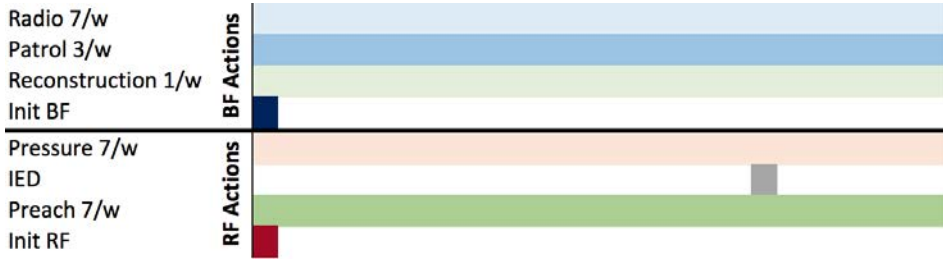


FIGURE 28 – Scénario d'initialisation

PARAMÈTRES DE CALIBRATION Nous avons 4 paramètres du modèle à calibrer pour 8 objectifs (quatre points de sondages pour chacune des deux Forces) :

- α équilibrant la partie émotionnelle et la surprise ;
- ξ la sensibilité émotionnelle au stimuli ;
- ρ le facteur de coresponsabilité ;
- T_{com} le seuil de probabilité de communication.

La valeur de U_0 est définie à 10, ce qui correspond à surprise temporelle équivalente à 3 ans environ (voir la section 4.3). Comme nous l'avons décrit dans le modèle, tous les paramètres sont partagés par toute la population : nos agents sont homogènes.

Notons que dans ce scénario, toutes les informations sont justes et compatibles : il n'y a pas d'opération de désinformation (c.f. section 4.4). Ce qui implique que toute la partie du modèle portant sur la révision de croyance n'entre pas en jeu dans ce cas d'application car la condition d'entrée dans cet algorithme est un écart entre les payoffs de deux actions similaires (voir la section 3.3). Par conséquent, il n'est pas nécessaire de spécifier les valeurs des paramètres relatifs à ce mécanisme (voir la section 3.7).

Nous avons également besoin de calibrer les attributs des actions d'initialisations AIRP et AIBN afin d'atteindre les favorabilités vis-à-vis des deux Forces au point P_1 . Les attributs de ces actions à calibrer sont leurs payoffs et leurs portées (soit un total de 4 valeurs). Nous obtenons ainsi un total de 8 valeurs à calibrer pour 8 objectifs. Cependant ces attributs des actions d'initialisation ne portent que sur la phase qui précède la simulation.

Nous définissons notre fonction de fitness (distance à l'objectif) selon la formule :

$$\text{fitness} = \sum_{p \in \{1,2,3,4\}} (\text{popAtt}(\text{RedForce}, p) - \text{polls}(\text{RedForce}, p))^2 + (\text{popAtt}(\text{BlueForce}, p) - \text{polls}(\text{BlueForce}, p))^2 \quad (29)$$

Où RedForce et BlueForce correspondent respectivement aux groupes sociaux des Forces Rouges et Bleus de l'ensemble SG . $\text{popAtt}(sg, p)$ une fonction donnant la favorabilité de la population vis-à-vis du groupe social $sg \in SG$ définie par le

pourcentage d'individus ayant une attitude positive à la phase $p \in 1, 2, 3, 4$. De la même manière, $\text{polls}(sg, p)$ la favorabilité de la population donnée par les sondages.

A chaque itération de l'algorithme de calibration, nous effectuons 20 répliques de simulations avec les mêmes jeux de paramètres afin de moyenniser la fitness. Cela diminue l'effet stochastique des simulations permettant ainsi de mieux orienter la recherche des paramètres et d'accélérer la calibration.

Notre critère d'arrêt consiste à terminer la calibration au bout de 100 itérations successives qui n'ont pas abouti à une amélioration de la fitness.

Nous avons procédé à ce protocole de calibration sur une grappe de calcul composée de 20 processeurs de type Intel Xeon cadencés à 4GHz. Les simulations sont lancées en parallèles (1 réplique par cœur). A chaque itération d'exploration, les résultats de ces répliques sont agrégés pour calculer la fitness de la génération courante. Le programme s'est arrêté au bout de 1500 itérations, soit environ 36 heures de calculs.

5.3.3 Résultats macroscopiques de calibration

Les résultats obtenus sont présentés dans la figure 29. Nous pouvons observer que la tendance générale des résultats de sondages est reproduite pour les deux Forces.

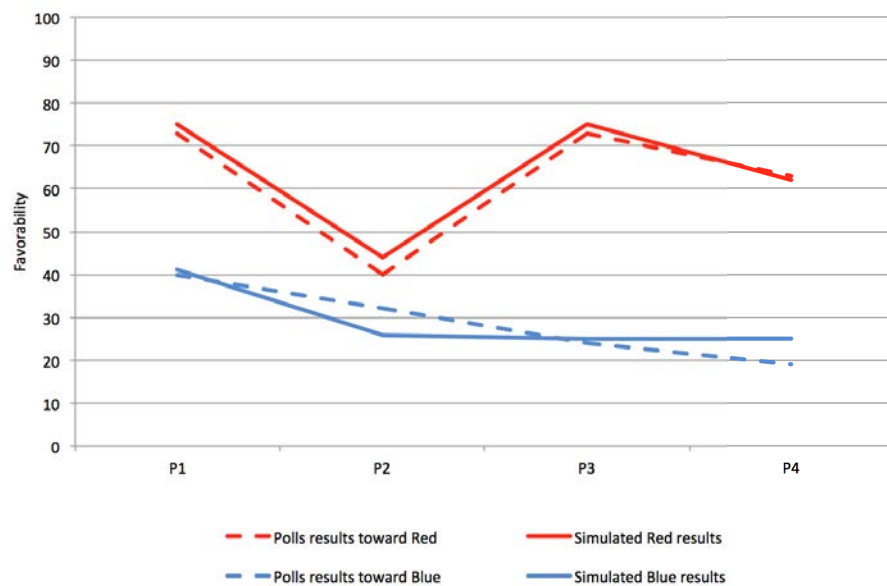


FIGURE 29 – Résultats de calibration

TENDANCE ROUGE Nous observons une baisse importante de l'attitude envers les Rouges sur la phase P1. Cette évolution retranscrit la période la plus meurtrière pour la population qui est marquée par l'attentat de Joybar.

La phase P2 correspond à une période d'hiver durant laquelle les Talibans sont moins actifs. Cette tendance se répercute dans les résultats de simulation avec une favorabilité croissante sur cet intervalle de temps. La raison de cette pente s'explique par le fait qu'en hiver le principal vecteur d'insécurité ne représente pas les Talibans, absents du terrain en période hivernale.

La baisse de l'attitude à la phase 4 correspond à la reprise des activités des Rouges au printemps. Cette chute est moins importante que lors de la première phase car elle est moins néfaste (moins d'attentats et d'IED).

TENDANCE BLEUE Malgré leurs efforts et l'absence d'actions négatives sur la population, les attitudes envers les Bleus ne cessent de décroître. Cette tendance peut s'expliquer selon deux facteurs.

Les opérations de brigades et les opérations spéciales qui ciblent les Rouges ne sont pas bénéfiques aux yeux d'une grande partie de la population qui soutient (attitudes positives) les Rouges. En effet, au cours de la majorité de la période simulée, la population est plutôt favorable aux Rouges. Les actions négatives sur les Rouges ont alors un impact négatif sur l'attitude envers les Bleus.

D'autre part, les Bleus subissent l'impact de la coresponsabilité (nous verrons dans le prochain paragraphe que $\rho = 0.21$). Ce mécanisme attribue une partie des impacts négatifs des Rouges aux Forces étrangères.

PARAMÈTRES Les valeurs de paramètres obtenues par la calibration :

- $\alpha = 0.77$ nous obtenons un profil d'individus que nous avons qualifié d'émotionnel dans le chapitre 4. Ainsi environ 25% de l'intérêt est constitué de l'effet de surprise ;
- $\xi = 0.01$ cette valeur appartient bien à l'intervalle discuté dans la section 4.1.1 ;
- $\rho = 0.21$ environ 20% des effets négatifs sur la population est attribuée à la Force Bleue. Les individus rejettent une petite partie de la faute de leurs malheurs à ceux qui ont pour rôle d'assurer leur sécurité ;
- $T_{com} = 0.15$ une information doit au moins être intéressante de la moitié de l'information référente pour pouvoir être relayée, tel que nous l'avons exposé dans la section 3.5.2.

MESURE DE L'IMPRÉCISION La valeur de fitness à la fin du processus de calibration donnée par l'équation 29 est de 0.08. Cependant, cette valeur n'est pas très explicite. C'est pourquoi nous avons évalué l'imprécision de nos résultats sur la base de l'erreur absolue. Elle se calcule selon une simple différence entre la valeur simulée et la valeurs donnée. Dans notre cas, nous devons agréger les erreurs de tous les points correspondant aux sondages. Nous prenons la valeur absolue de ces différences afin d'éviter que deux erreurs de signes opposés se compensent et baissent ainsi l'erreur au niveau agrégé. Ainsi, notre formule calculant l'erreur entre les valeurs simulées et les résultats de sondages pour une Force donnée est de la forme :

$$\text{error}(\text{force}) = \sum_{p \in \{1,2,3,4\}} \frac{|\text{popAtt}(\text{force}, p) - \text{polls}(\text{force}, p)|}{4} \quad (30)$$

Nous avons obtenu une erreur moyenne de 2.25 pour les Rouges avec un maximum de 4 à la phase 2 et une moyenne de 3.5 pour les Bleus avec un maximum de 6 aux phases 2 et 4. Ce qui fait une moyenne globale de 2.875 points d'erreur pour les deux courbes sur des valeurs comprises entre 20 et 80% de favorabilité.

5.4 ANALYSES MICROSCOPIQUES

Les mesures macroscopiques sur la favorabilité de la population vis-à-vis des Forces ont permis de calibrer les paramètres du modèle sur la base des résultats de sondages. Cependant, ces mesures ne représentent qu'une partie des sorties du simulateur. En effet, un des atouts des systèmes multi-agent est leur capacité de description au niveau microscopique, ce que ne permettent pas, par exemple, les simulations à base d'équations statistiques. Dans cette section, nous présentons des mesures de sorties du simulateur à un niveau plus bas. Nous commençons à cet effet par une classification des individus en fonction de l'évolution de leurs attitudes vis-à-vis des Forces que nous étudierons par la suite.

5.4.0.1 *Classification des agents*

Nous avons appliqué les cartes auto adaptatives de Kohonen (1998) (Self-Organising Map, SOM) afin de classifier les individus selon l'évolution de leurs attitudes au cours de la simulation, en se référant à leurs attitudes à chaque date de sondage (P₁, P₂, P₃ et P₄). Les individus sont alors caractérisés par 8 valeurs d'attitudes (4 dates, 2 Forces). Cette méthode de classification appartient à la classe des réseaux de neurones artificiels, fondée sur des méthodes d'apprentissages non supervisés et permet de produire des représentations discrètes (souvent en deux dimensions) des données d'entrées.

PRINCIPE DE L'ALGORITHME La carte est composée de neurones (ou cellules) qui correspondent également aux classes qui contiendront les données d'entrées, dont le nombre et la topologie sont donnés par l'utilisateur. Par exemple, dans notre cas, nous avons 9 neurones disposés en 3 lignes trois colonnes sur une grille hexagonale (l'hexagone est souvent préféré au carré car une cellule a plus de voisins). La période d'entraînement démarre avec des neurones dont les valeurs sont initialisées de manière aléatoire (les valeurs sont des vecteurs de la même forme que les données d'entrées). Chacune des données est soumise aux neurones et va être classée dans une de ces cellules, celle qui est la plus proche d'elle (la distance minimale entre les deux vecteurs). Lorsque la meilleure cellule est identifiée, elle sera gratifiée en se rapprochant de la valeur de la donnée nouvellement acquise. Par la même occasion, les cellules voisines (selon la topologie donnée) sont elles aussi gratifiées (leurs vecteurs sont modifiés) selon un facteur multiplicatif inférieur à 1. Ainsi, c'est toute la région de la carte autour du neurone gagnant qui se spécialise. L'entraînement se termine lorsque les neurones ne modifient plus leurs valeurs (ou très peu, à un paramètre près). Enfin, une fois l'apprentissage terminé, les données sont mappées sur les cellules.

RÉSULTAT DE CLASSIFICATION La figure 30 montre les poids des vecteurs des cellules à la fin de l'apprentissage. Les lignes noires (en gras) délimitent les cluster de classes obtenus en appliquant une analyse de cluster hiérarchique sur les distances entre les classes. Ainsi, nous pouvons constater que les classes 2 et 3 sont similaires et appartiennent au même cluster car elles sont les seules à se baser uniquement sur les valeurs de p_{1r} (attitude de l'individu à la Phase 1 vis-à-vis de la Force Rouge), p_{3b}, p_{1b} et p_{3r}. De la même manière, les classes 8 et 9 sont similaires dans le sens où elles considèrent uniquement les valeurs p_{2r}, p_{4b}, p_{2b}

et p_{4r} . En contraste, les autres classes se servent de toutes les valeurs avec des poids plus ou moins importants. Cette représentation ne nous informe que sur les propriétés des classes et ne nous permet pas de comprendre la classification.

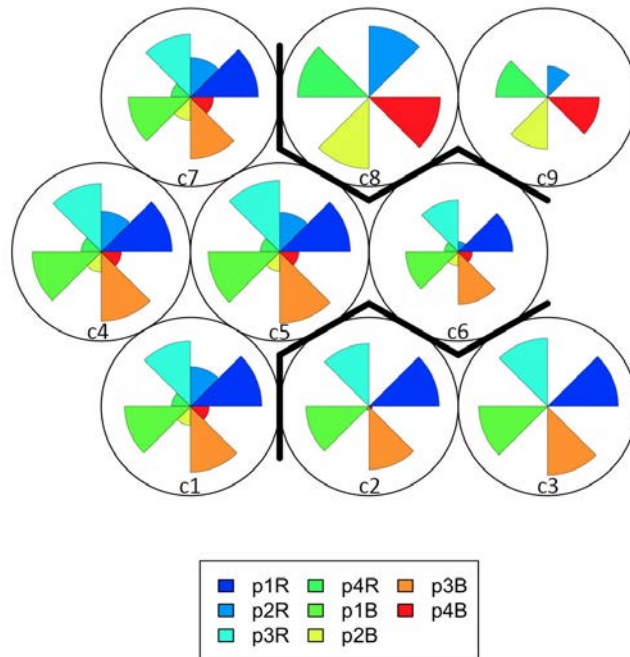


FIGURE 30 – Carte de Kohonen avec les poids des vecteurs du modèle de classification

La figure 31 offre un aperçu des évolutions des attitudes des individus appartenant à chacune des classes. Les courbes bleues et rouges représentent la moyenne d'attitude au sein de la classe vis-à-vis de la Force Bleue et Rouge respectivement. Les lignes grises horizontales correspondent aux axes des abscisses ($y = 0$). Nous remarquons une similarité entre les classes 2 et 3 ainsi qu'entre les classes 8 et 9, comme cela avait été anticipé sur la figure 30. Nous pouvons qualifier le profil des classes 2 et 3 comme étant *anti-Bleus* car les valeurs d'attitudes vis-à-vis des Bleus sont les plus négatives de toutes les classes. De la même manière les classes 8 et 9 correspondent aux *pro-Bleus*. Les autres classes (1, 4, 5 et 7) sont également similaires entre elles, avec des valeurs d'attitude proches de zéro. Nous pouvons les qualifier de *neutres* par rapport aux deux autres profils. Autre observation importante, la classe 6 n'a pas de courbe. Nous verrons dans le prochain paragraphe que cette classe est vide.

La figure 32 expose des propriétés de la carte de classification : l'effectif ou le nombre d'individus appartenant à la classe ainsi que la qualité des classes, fonction de la dispersion des individus au sein d'une même classe. Dans la figure sur les effectifs, nous remarquons que la classe 6 (en gris) est vide. En effet, l'algorithme de Kohonen n'interdit pas l'existence de classes vides. Par ailleurs, nous pouvons observer que les classes 2 et 3 (les *anti-Bleus*) représentent à peu près la moitié de la population totale (299 agents pour une taille totale de 500 agents). La troisième classe la plus peuplée est la classe 9 avec 86 agents. En ajoutant les individus de la classe 8, nous obtenons un effectif de 122 agents de profil *pro-Bleus*. Les 79 agents restants appartiennent donc aux *neutres*. La dispersion au sein des classes est très faible de manière générale, à l'exception

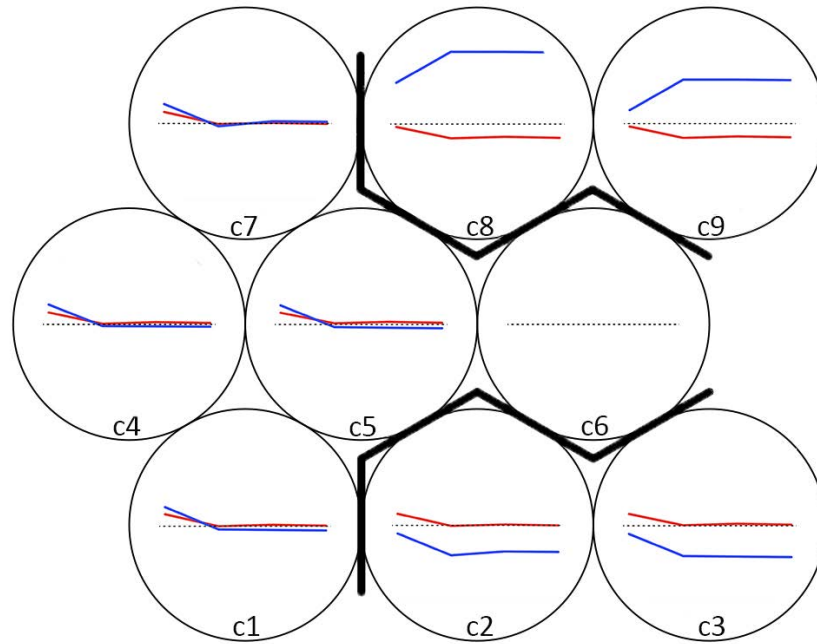


FIGURE 31 – Carte de Kohonen avec un aperçu des moyennes d’attitudes vis-à-vis des Forces

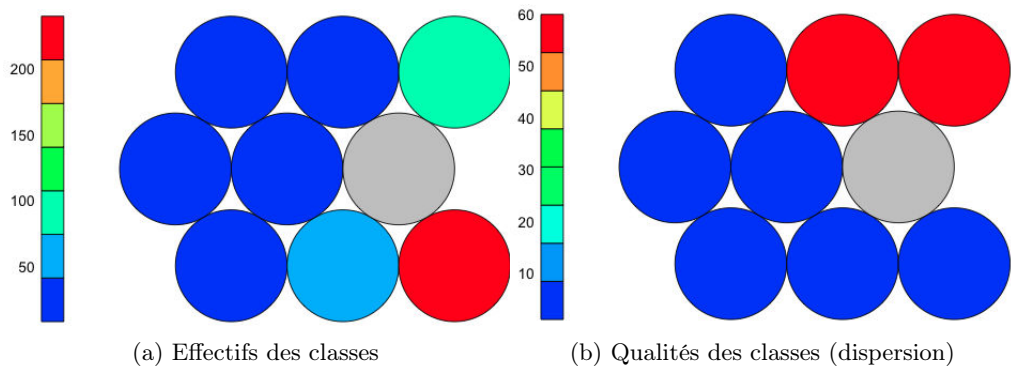


FIGURE 32 – Mesures sur la classification de Kohonen

des classes 8 et 9. Cela est dû, comme nous le montrerons plus tard, aux valeurs élevées des attitudes au premier point de sondage (p1b).

5.4.0.2 Analyses des clusters

Dans cette section, nous allons voir plus en détails les trois clusters de classes définis précédemment regroupant les *anti-Bleus*, les *pro-Bleus* et les *neutres*. Pour cela, nous affichons l’évolution des attitudes moyennes au sein des classes correspondantes ainsi que leurs écart-types afin de représenter la dispersion des individus.

CLUSTER ANTI-BLEUS La figure 33 présente les moyennes d’attitudes et écart-types des classes 2 et 3 appartenant au cluster des *anti-Bleus*. Nous observons que :

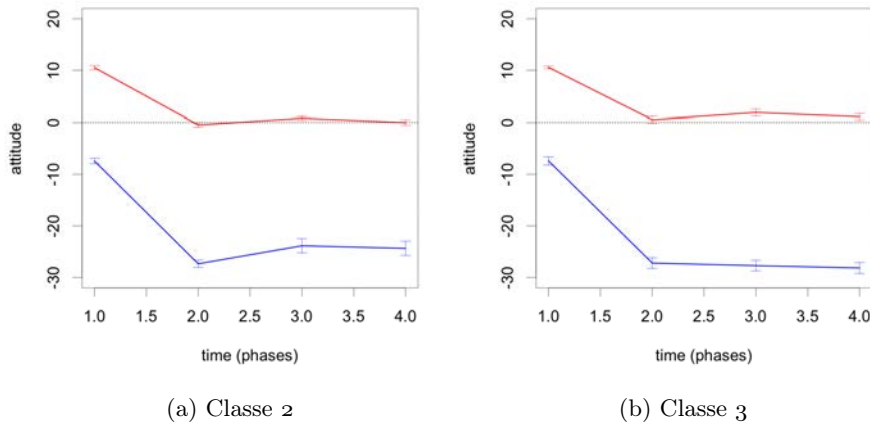


FIGURE 33 – Moyennes et écart-types d'attitudes des classes 2 et 3

1. dans la distribution initiale (P_1), les valeurs d'attitudes vis-à-vis des Bleus et des Rouges sont opposées ;
2. les tendances d'attitudes vis-à-vis des Rouges sont globalement similaires pour les deux classes ;
3. tandis que la classe 3 maintient une attitude positive durant toute la durée de la simulation, la classe 2 voit son attitude décroître entre la phase 1 et 2 jusqu'à atteindre des valeurs négatives et ce pour tous les individus de la classe, comme le montre l'écart-type. Par ailleurs, à la phase 4, une partie des individus de la classe 2 affiche des valeurs négatives vis-à-vis des Rouges ;
4. la moyenne d'attitude vis-à-vis des Bleus est décroissante monotone chez la classe 3, bien qu'elle augmente sur la phase 2 chez la classe 2 ;
5. à la fin de la simulation, les attitudes vis-à-vis des Bleus sont très négatives, mais celles envers les Rouges sont très proches de zéro.

L'observation (2) témoigne d'une certaine qualité du cluster. De plus, les observations (3) et (4) montrent que malgré les similarités entre les classes, il existe des différences cruciales entre elles, confirmant l'intérêt d'avoir classifié les individus selon deux catégories pour ce cluster. En outre, l'observation (4) est une conséquence de l'observation (3) : l'attitude envers les Bleus augmente durant la phase 2 car les individus de la classe 2 sont défavorables vis-à-vis des Rouges au cours de cette période. L'évaluation des actions néfastes aux Rouges est considérée comme bénéfique pour ces individus. Or, ce n'est pas le cas pour la classe 3 qui a une attitude positive pour les Rouges. D'où cette différence de tendance entre les deux classes.

D'après ces observations ainsi que d'après (1) et (5), nous pouvons conclure que les anti-Bleus ne sont pas nécessairement en faveur des Rouges. Au contraire, les attitudes envers les Rouges décroissent pour tous les individus au cours de la période simulée pour atteindre des valeurs neutres.

CLUSTER PRO-BLEUS La figure 34 présente les moyennes d'attitudes et écart-types des classes 8 et 9 appartenant au cluster des *pro-Bleus*. Nous observons que :

1. les tendances d'évolution des classes 8 et 9 sont similaires, seule une translation de la courbe bleue les différencie ;
2. la favorabilité envers les Bleus est renforcée par des attitudes croissantes vis-à-vis de cette Force ;
3. similairement, la défavorabilité envers les Rouges est accentuée ;
4. la moyenne d'attitude vis-à-vis des Rouges croît très légèrement durant la seconde phase.

D'après l'observation (1), dans le cas de ce cluster, les deux classes n'ont pas nécessairement besoin d'être étudiées séparément : elles sont suffisamment similaires pour ne faire qu'une seule classe.

Les observations (2) et (3) montrent que les individus de ce cluster s'extrémisent au cours de la simulation. En effet, l'attitude vis-à-vis des Bleus augmente considérablement sur la première phase car elle est la plus cinétique de la simulation (phase durant laquelle les actions ciblant les Talibans sont les plus nombreuses). Les attitudes envers les Rouges étant négatives, ces actions Bleues ne peuvent qu'être en leur faveur. Cependant, les actions bleues ciblant les Rouges étant en diminution à partir de la seconde phase marquée par le début du retrait des troupes françaises, les quelques actions bleues ne parviennent qu'à maintenir les attitudes bleues jusqu'à la fin de la période étudiée. La dynamique de la courbe rouge s'explique de la même manière : leurs actions sont considérées comme néfastes et font généralement baisser les attitudes rouges. A une exception près, relevée par l'observation (4). Cette croissance retranscrit la période d'hiver où l'activité talibane est réduite, seuls les prêches (bénéfiques) gardent une fréquence d'occurrences régulière, augmentant ainsi les attitudes vis-à-vis des Rouges de manière sensible.

Ces observations nous permettent d'affirmer que les pro-Bleus sont aussi anti-Rouges : on assiste dans ce cluster à une extrémisation des attitudes envers les Forces.

A l'issue des analyses de ces deux clusters (anti et pro-Bleus) nous pouvons émettre deux observations :

1. Les valeurs d'attitudes vis-à-vis des Rouges sont comprises entre -20 et 20, tandis que celles vis-à-vis des Bleus varient entre -30 et 60.
2. Les signes des attitudes initiales sont conservés dans ces clusters.

L'observation (1) s'explique du fait que les actions bleues telles que les opérations de brigade impactent de nombreux Talibans (de l'ordre d'une dizaine) tandis que les actions rouges telles que les IED n'affectent que quelques individus. Selon la formule de l'évaluation d'une action (voir l'équation 22 à la page 52), l'impact d'une action est fonction du nombre de bénéficiaires. L'observation (2) semble indiquer que les actions d'initialisation AIRP et AIBN jouent un rôle central dans la dynamique des attitudes. Les résultats de la simulation ont l'air d'être déterminés par les attributs de ces actions. Cependant, nous verrons dans l'analyse du prochain cluster que cela n'est pas le cas.

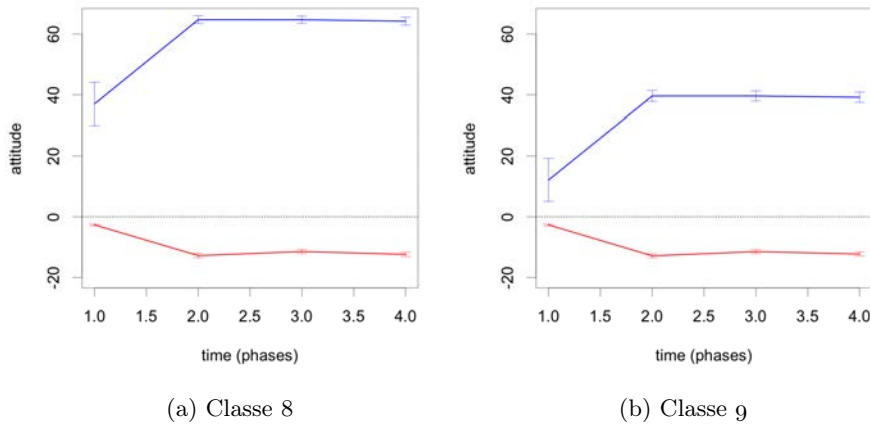


FIGURE 34 – Moyennes et écart-types d'attitudes des classes 8 et 9

CLUSTER NEUTRES La figure 35 présente les moyennes d'attitudes et écart-types des classes 1, 4, 5 et 7 appartenant au cluster des *neutres*. Nous observons que :

1. les distributions des attitudes à l'initialisation sont similaires pour les quatre classes avec celles vis-à-vis des Bleus, plus importantes que celles vis-à-vis des Rouges ;
2. les attitudes ne dépassent jamais ± 4 à partir de la phase 2 ;
3. bien que les classes 1, 4 et 5 soient similaires, la classe 7 a une tendance plus particulière : la moyenne d'attitude envers les Bleus redevient positive à partir de la troisième phase, après être devenue négative lors de la seconde phase.

L'observation (2) nous permet de confirmer que les individus de ce cluster ont bien des attitudes neutres en comparaison avec les valeurs d'attitudes des deux autres clusters (allant de -30 à $+65$).

Comme le témoigne l'observation (3), la classe 7 est singulière parmi ce cluster, dans le sens où les signes des deux moyennes d'attitudes à P4 sont opposés. En effet, dans les classes 1, 4 et 5, les signes des moyennes d'attitudes sont intervertis entre P1 et P4 (1) : les attitudes vis-à-vis des Bleus sont négatives à la fin de la simulation bien qu'elles aient été initialisées positivement. Cette différence trouve encore une fois son origine dans le subtil écart de valeur au point P2 : le signe négatif envers les Rouges chez les classes 1, 4 et 5 et positif chez la classe 7.

Les classes 1, 4 et 5 ont montré que la favorabilité ou la défavorabilité de la population vis-à-vis d'une Force à la fin de la simulation n'étaient pas données directement par le processus d'initialisation comme nous aurions pu le croire au vu des tendances des deux autres clusters.

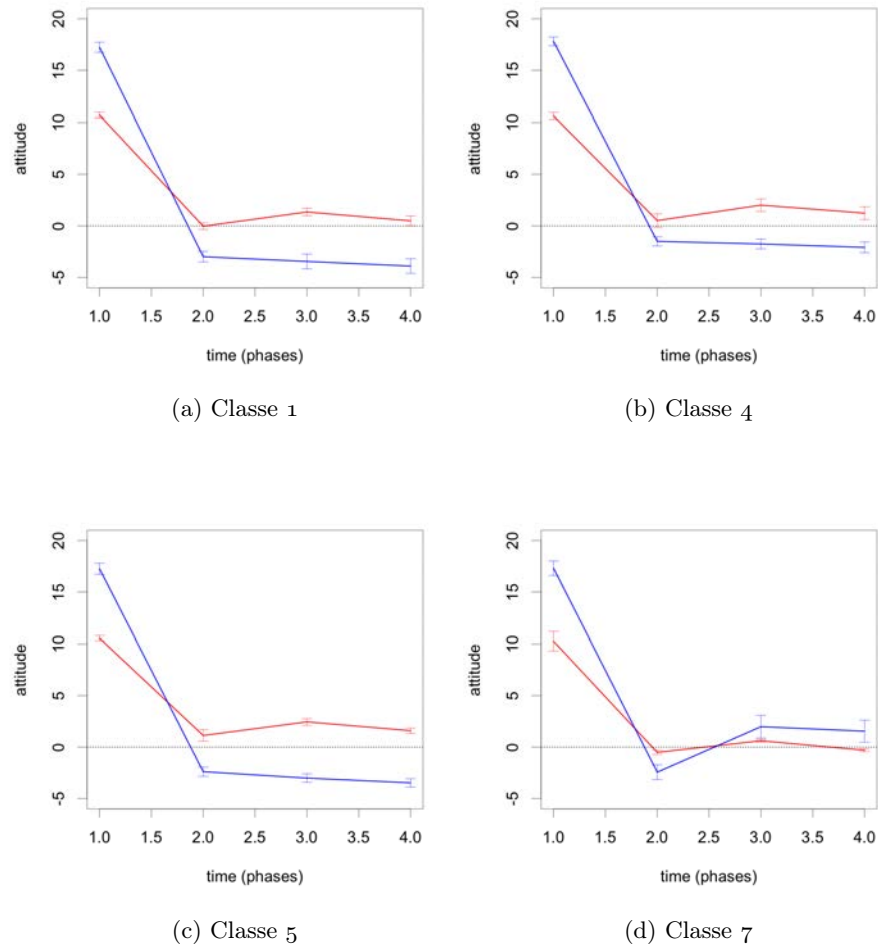


FIGURE 35 – Moyennes et écart-types d’attitudes des classes 2 et 3

Ces analyses nous montrent qu’en plus des résultats macroscopiques valides sur le cas d’étude, le modèle est également capable de reproduire des tendances d’évolutions d’attitudes variées et rationnelles au sein de la population. De plus, la classification des individus nous a permis de retrouver des catégories de profils que nous avons pu analyser de manière intuitive. Bien que nous ne disposons pas de données précises qui permettraient de valider ce niveau des résultats, nous pouvons conclure que les sorties du modèle sont riches.

5.5 APPLICATION DU MODÈLE : BILAN

Nous avons rassemblé des données de terrain sur les opérations militaires françaises de stabilisation en Afghanistan. Ces données portent d’une part sur la nature des actions de contre-insurrections, aussi bien du côté allié qu’ennemi, ainsi que leur déroulement au cours de l’intervention du CIAE en Kapisa entre 2011 et 2012. D’autre part, nous avons également obtenu, traité et analysé des données de sondages classifiées qui ont été collectées par la société Altai Consulting sur la zone géographique et la période correspondant aux opérations militaires étudiées.

Les données sur la progression des opérations recueillies à travers les interviews nous ont permis de reconstituer un scénario que nous avons implémenté dans le cadre du modèle Polias. En couplant ce modèle aux données de sondages, nous avons procédé à une calibration des paramètres en appliquant l'algorithme CMA-ES. Ce processus a permis de reproduire les données de sondages avec une erreur inférieure à 3 points d'écart.

Les trajectoires des attitudes des individus résultant de la simulation calibrée ont été classifiées selon deux niveaux. Un niveau fin obtenu à l'aide d'une carte de Kohonen composée de neuf cellules, permettant ainsi de garder une diversité de tendances d'attitudes. Un niveau plus macroscopique en appliquant une analyse de clusters hiérarchiques sur les cellules de la carte afin d'obtenir trois classes qui correspondent à des types de profils sémantiquement plus classiques : les pro-Bleus, les anti-Bleus et les neutres. Ces classes et clusters ont été analysés de manière à expliquer les tendances des attitudes.

Dans un futur proche, nous analyserons plus en profondeur la diversité des agents en étudiant notamment l'évolution des bases croyances des individus qui résultent de leurs perceptions des actions, ainsi que de leurs évaluations. En effet, les dates auxquelles les agents apprennent les actions semblent constituer un facteur déterminant dans l'évolution des attitudes. De plus, les agents ne sont pas au courant de toutes les actions, une classification des agents selon leurs connaissances semble alors être une piste intéressante à investiguer.

CONCLUSION ET PERSPECTIVES

6.1 BILAN

Dans cette thèse, nous avons étudié la problématique de la simulation de la formation et de la dynamique des attitudes d'une population. Pour ce faire, nous sommes partis en premier lieu d'une synthèse des travaux en psychologie sociale portant sur le concept de l'attitude. Ainsi, nous avons pu identifier les notions et concepts fondamentaux qui nous semblent nécessaires de modéliser dans le cadre de notre étude. Cette première étape s'inscrit dans l'approche psychométrique que nous défendons. Elle propose de se référer à la discipline originelle des concepts cibles à représenter. En complément à ce tour d'horizon sur les travaux en psychologie, nous avons étudié les différentes approches de la simulation sociale abordant la question de la dynamique d'attitudes/opinions. Dans cette seconde partie de l'état de l'art, nous avons décrit en détail quelques modèles ayant marqué la recherche sur cette problématique. Nous avons également souligné les manques et les limites de ces modèles qui restent en suspens.

Sur la base de ces deux synthèses portant sur les travaux connexes à la problématique étudiée dans cette thèse, nous avons proposé un modèle multi-agent. Le modèle cognitif des individus que nous proposons s'articule autour d'une théorie centrale issue de la psychologie sociale : l'attitude en tant qu'un ensemble de relations pondérées entre des évaluations et un objet social de [Fazio \(2007\)](#). Les agents perçoivent les actions entreprises par les acteurs sociaux. Les croyances portant sur ces actions sont interprétées individuellement, mémorisées ou non selon un certain degré de rétention. L'évaluation de ces connaissances forme et impacte les attitudes. Chaque partie du modèle se fonde sur des théories issues des sciences humaines et sociales : la révision de croyance reprend un mécanisme proposé par [Thiriot et Kant \(2007\)](#), l'évaluation d'une information est basée sur les travaux de [Fishbein et Ajzen \(1975\)](#), le calcul de l'intérêt narratif et de rétention d'une information reprend la Théorie de la Simplicité proposée par [Dessalles \(2007\)](#). Les interprétations informatiques, l'agencement de ces théories ainsi que la dynamique du modèle constituent les contributions principales de nos recherches.

Nous avons ensuite procédé à une analyse fonctionnelle du modèle proposé afin d'étudier l'impact de ses paramètres sur son comportement au niveau global. Cette analyse a notamment permis d'identifier les phénomènes descriptibles dans le cadre de nos simulations. Nous avons observé que le modèle est capable de reproduire une certaine dynamique de diffusion d'informations répertoriée par [Rogers \(2010\)](#), une dynamique d'attitudes qui tient compte des relations entre les groupes sociaux, l'impact de la topologie du réseau social sur les dynamiques d'attitudes, une résistance des individus à la désinformation ainsi que le phénomène de conformité étudié par [Crutchfield \(1955\)](#) ou encore [Kelman \(1958\)](#).

Enfin, nous avons appliqué le modèle au contexte militaire des menaces asymétriques dans lequel les attitudes de la population représentent une des priorités selon le nouveau paradigme de la guerre. Dans cette démarche, nous avons récolté des données sur le cas de l'intervention militaire française en Kapisa entre 2011

et 2012. D'une part, nous avons collecté des données sur les séquences d'actions militaires alliées et ennemies fournies notamment par des officiers du CIAE en charge d'opérations dans cette région au cours de cette période. D'autre part, nous avons analysé des résultats de sondages d'opinions évaluant la favorabilité de la population vis-à-vis des deux Forces en présence. Sur la base de ces données, nous avons calibré notre modèle en appliquant l'algorithme CMA-ES afin d'identifier les valeurs de paramètres permettant de rapprocher les résultats de simulation aux données de sondages. A l'issue de ce processus de calibration, les résultats de simulation diffèrent des données de sondage à raison d'une erreur moyenne inférieure à trois points d'écart. Enfin, nous avons étudié les résultats de simulation en classifiant les individus selon leurs dynamiques d'attitudes en appliquant les cartes de Kohonen ainsi que la classification hiérarchique. Cette étude a permis d'identifier trois groupes d'individus ayant chacun une tendance d'évolution d'attitude spécifique : les anti-Rouge, les anti-Bleu et les neutres. Ces trois groupes se distinguent principalement par la différence d'évaluation des actions négatives de la Force Bleue sur les Rouges, en fonction de la favorabilité des individus vis-à-vis des Rouges.

Bien que nous n'ayons pas appliqué le modèle à d'autres contextes, nous sommes convaincus que le modèle ne se limite pas à des applications militaires, compte tenu de la généralité des concepts manipulés. Par exemple, les acteurs sociaux pourraient représenter des partis politiques en campagnes électorales ciblant différents groupes sociaux caractérisés par des affinités idéologiques et politiques.

6.2 ETHIQUE

Nous sommes conscients que l'application de notre modèle s'expose à des questions d'éthique. En effet, à terme, nos travaux pourraient permettre de simuler des scénarii de manière artificielle afin d'anticiper leurs conséquences sur les attitudes d'une population donnée. Sur la base d'une comparaison des différentes éventualités simulées, une minorité d'individus protégeant ses propres intérêts aurait le pouvoir de manipuler l'opinion publique.

Ce risque n'est pas spécifique à nos travaux, il se retrouve dans l'ensemble de ceux sur lesquels nous nous basons. Par exemple, les travaux en psychologie sociale sur l'attitude peuvent être perçus comme des outils de manipulation. C'est le cas des recherches sur les impacts subliminaux de l'exposition sur l'attitude (voir la section 2.1.6 à la page 20). Les travaux sur l'exposition ont en effet été utilisés dans les milieux médiatiques et politiques (ex. : la publicité, les messages subliminaux dans les journaux télévisés). Cependant, l'objectif principal et immuable de cette discipline consiste à comprendre les structures et les mécanismes psychologiques sous-jacents à l'activité humaine.

Notre position quant à l'application de nos travaux peut se définir ainsi :

- L'agencement informatique de différentes théories sur l'attitude issues de la psychologie sociale, tel que nous l'avons fait dans cette thèse, peut permettre de mieux comprendre les interactions entre ces théories et, ainsi, de contribuer à éclairer les dangers des méthodes de manipulations qu'elles mettent en évidence.
- L'application de notre simulateur dans le cadre militaire a pour but d'encourager l'usage d'actions non-cinétiques, telles que la communication ou

la reconstruction, en substitution aux tactiques cinétiques conventionnelles impliquant des pertes humaines. En améliorant la formation des officiers spécialisés dans les actions d'influences, nous participons à l'optimisation de l'approche globale qui place la population au centre des intérêts militaires.

6.3 PERSPECTIVES

Nous n'avons abordé dans cette thèse qu'une partie de la problématique de la simulation de l'activité humaine. Plusieurs perspectives de recherche découlent de nos travaux.

CONTEXTES Notre modèle est mono-contextuel : les attitudes simulées ne portent que sur un unique contexte socio-culturel. Par exemple, dans notre cas d'application, nous avons simulé les attitudes vis-à-vis des Forces dans le contexte sécuritaire. Les actions simulées et les sondages d'opinions portaient sur la sécurité et le sentiment de sécurité de la population. Or, nous pouvons imaginer que l'attitude envers une Force, ou tout autre acteur social, ne relève pas d'un seul contexte. Par exemple, nous pouvons avoir une attitude très positive vis-à-vis d'un collègue dans le contexte professionnel car nous l'estimons compétent dans son métier. Cela n'empêche en rien d'avoir une attitude négative envers la même personne dans le contexte sportif compte tenu de son incapacité à jouer au football lors d'un match inter-entreprises. Dans le cas afghan, les Forces pourraient être évaluées dans les contextes économiques et religieux en complément du contexte sécuritaire. Une manière simple de simuler ces différentes attitudes contextuelles serait de procéder à une simulation par scénario regroupant des actions relatives à un contexte. Cependant, nous pensons qu'il existe des relations d'influences entre les différents contextes : un impact négatif sur l'attitude vis-à-vis de notre collègue dans le milieu social peut avoir une répercussion négative sur l'attitude professionnelle. Ce problème pourrait être représenté en utilisant des modèles ontologiques définissant les contextes, leurs relations et les appartenances des actions.

AGENTS HÉTÉROGÈNES Chaque agent dans notre modèle est unique selon ses bases de connaissances, ses liens dans le réseau social, ainsi que ses affiliations aux groupes sociaux. Cependant, leurs paramètres cognitifs sont totalement homogènes : toutes les valeurs de paramètres sont identiques à travers la population. Il serait intéressant de représenter différents profils cognitifs dans une même population afin d'étudier séparément chaque groupe. Par exemple, nous pourrions reprendre les profils définis dans la section 4.1.1. Mais ces groupes ont été définis de manière arbitraire à des fins d'analyses de sensibilité sur des scénarii abstraits. Afin de pouvoir appliquer le modèle avec des groupes hétérogènes, il est nécessaire de recueillir des données sur les différents profils cognitifs existants dans la population à modéliser.

REPRÉSENTATIONS SOCIALES Les travaux de Fazio (2007) sur la structure de l'attitude et ceux de Moscovici et Abric (1984) sur les représentations sociales présentent de nombreuses similitudes. En effet, ces travaux se basent sur une représentation symbolique centrée sur un objet social relié à d'autres concepts

connexes avec des pondérations plus ou moins importantes sous la forme d'un graphe. Tous deux n'imposent aucune contrainte quant à la nature des concepts manipulés. Ils doivent toutefois être évaluables dans la théorie de Fazio car l'attitude est une synthèse d'évaluations. La pondération des liens est mesurée selon l'accessibilité mémoire pour Fazio. Chez Moscovici, cette pondération représente le nombre de personnes ayant évoqué le concept ainsi que l'ordre dans lequel ils ont été évoqués. Ces deux travaux se distinguent principalement par la différence d'échelle : Fazio travaille au niveau de l'individu tandis que Moscovici se place au plan des groupes sociaux. Ainsi, l'écart entre nos travaux et ceux des représentations sociales est infime. Il serait intéressant de rallier ces deux travaux sur la base de notre simulateur en annexant un modèle de représentations sociales afin d'étudier leurs interactions et leur complémentarité avec le concept de l'attitude.

COMPORTEMENTS L'objectif *in fine* d'un simulateur d'attitude consiste à reproduire des comportements. Les individus simulés devraient avoir la possibilité d'agir sur l'environnement et les autres individus par le biais d'actions. En d'autres termes, la frontière entre les individus et les acteurs sociaux devrait être moins nette. Par exemple, nous pourrions simuler le comportement à la participation à des émeutes, le vote ou encore l'adoption de nouveaux produits. Cette transition de l'attitude vers le comportement représente une des thématiques phares en psychologie sociale. En effet, Kraus (1995) a conclu l'existence d'une corrélation de 38% entre l'attitude et le comportement en se basant sur une méta-analyse portant sur 88 études. Cependant, de nombreux biais entravent cette relation, par exemple le taux d'implication du sujet (ex. : Chaiken et Eagly (1989)), les normes sociales (ex. : Fishbein et Ajzen (1975)), la couche volitionnelle (incompatibilité entre la volonté inconsciente et le comportement), les obstacles environnementaux (les obstacles physiques à l'accomplissement de l'action). Il existe aussi des influences des comportements sur les attitudes telles que la dissonance cognitive succédant un comportement discordant avec l'attitude du sujet (Festinger (1957)). Une perspective à long terme serait alors d'établir un modèle combinant une composante de dynamique d'attitude et une autre sur les comportements d'une population en se référant à ces théories en psychologies sociale et comportementale.

BIBLIOGRAPHIE

- Abelson, R., D. Kinder, M. Peters et S. Fiske. 1982, «Affective and semantic components in political person perception.», *Journal of personality and social psychology*, vol. 42, n° 4, p. 619. URL <http://psycnet.apa.org/journals/psp/42/4/619/>.
- Ajzen, I. 1991, «The theory of planned behavior», *Organizational behavior and human decision processes*, vol. 50, n° 2, p. 179–211. URL <http://www.cas.hse.ru/data/816/479/1225/Oct%2019%20Cited%20231%20Manage%20THE%20THEORY%20OF%20PLANNED%20BEHAVIOR.pdf>.
- Ajzen, I. et M. Fishbein. 1977, «Attitude-behavior relations : A theoretical analysis and review of empirical research.», *Psychological Bulletin ; Psychological Bulletin*, vol. 84, n° 5, p. 888. URL <http://psycnet.apa.org/journals/bul/84/5/888/>.
- Ajzen, I. et M. Fishbein. 2005, «The influence of attitudes on behavior», dans *The handbook of attitudes*, édité par D. Albarrac, B. T. Johnson et M. P. Zanna, Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US, ISBN 0-8058-4492-9 (Hardcover) ; 0-8058-4493-7 (Paperback), p. 173–221.
- Allport, G. W. 1935, «Attitudes.», URL <http://psycnet.apa.org/psycinfo/1935-19907-010>.
- Alt, J. K., A. Leroy, D. Hudak et S. Lieberman. 2009, «The cultural geography model : Evaluating the impact of tactical operational outcomes on a civilian population in an irregular warfare environment», *The Journal of Defense Modeling and Simulation : Applications, Methodology, Technology*, vol. 6, n° 4, p. 185–199. URL <http://dms.sagepub.com/content/6/4/185.short>.
- Anderson, J. R. et L. J. Schooler. 1991, «Reflections of the environment in memory», *Psychological science*, vol. 2, n° 6, p. 396–408. URL <http://pss.sagepub.com/content/2/6/396.short>.
- Bandura, A. 1969, «Social-learning theory of identificatory processes», *Handbook of socialization theory and research*, vol. 213, p. 262. URL <http://web.iaincirebon.ac.id/ebook/indrya/Bandura/Bandura1969%20Handbook.pdf>.
- Beauvois, J.-L. et J.-C. Deschamps. 1996, *La psychologie sociale* . . ., Presses Univ. de Grenoble, Grenoble, ISBN 2706106697 9782706106699.
- Bem, D. J. 1973, «SELF-PERCEPTION THEORY₁», URL <http://www.dbem.ws/SP%20Theory.pdf>.
- Berlyne, D. E. 1970, «Novelty, complexity, and hedonic value», *Attention, Perception, & Psychophysics*, vol. 8, n° 5, p. 279–286. URL <http://www.springerlink.com/index/C027232G41U30377.pdf>.

- Bornstein, R. F. 1989, «Exposure and affect : Overview and meta-analysis of research, 1968–1987.», *Psychological bulletin*, vol. 106, n° 2, p. 265. URL <http://psycnet.apa.org/journals/bul/106/2/265/>.
- Breckler, S. J. 1984, «Empirical validation of affect, behavior, and cognition as distinct components of attitude.», *Journal of personality and social psychology*, vol. 47, n° 6, p. 1191. URL <http://psycnet.apa.org/journals/psp/47/6/1191/>.
- Brock, T. C. 1967, «Communication discrepancy and intent to persuade as determinants of counterargument production», *Journal of Experimental Social Psychology*, vol. 3, n° 3, p. 296–309. URL <http://www.sciencedirect.com/science/article/pii/0022103167900315>.
- Burnkrant, R. E. et A. G. Sawyer. 1983, «Effects of involvement and message content on information-processing intensity», *Information processing research in advertising*, vol. 12, n° 2, p. 46–64.
- Cacioppo, J., B. Marshall-Goodell, L. Tassinary et R. Petty. 1992, «Rudimentary determinants of attitudes : Classical conditioning is more effective when prior knowledge about the attitude stimulus is low than high», *Journal of Experimental Social Psychology*, vol. 28, n° 3, p. 207–233. URL <http://www.sciencedirect.com/science/article/pii/002210319290053M>.
- Cacioppo, J. et R. Petty. 1982, «The need for cognition.», *Journal of Personality and Social Psychology; Journal of Personality and Social Psychology*, vol. 42, n° 1, p. 116. URL <http://psycnet.apa.org/journals/psp/42/1/116/>.
- Castellano, C., S. Fortunato et V. Loreto. 2009, «Statistical physics of social dynamics», *Reviews of modern physics*, vol. 81, n° 2, p. 591. URL http://rmp.aps.org/abstract/RMP/v81/i2/p591_1.
- Chaiken, S. 1980, «Heuristic versus systematic information processing and the use of source versus message cues in persuasion.», *Journal of personality and social psychology*, vol. 39, n° 5, p. 752. URL <http://psycnet.apa.org/journals/psp/39/5/752/>.
- Chaiken, S. et A. H. Eagly. 1989, «Heuristic and systematic information processing within and», *Unintended thought*, vol. 212. URL <http://books.google.com/books?hl=fr&lr=&id=HT6ddclz6EAC&oi=fnd&pg=PA212&dq=Heuristic-Systematic+chaiken+89&ots=dc6THRn7rg&sig=7rHWRo3zyamLDwfD0YA-QP91zVQ>.
- Chater, N. et P. Vitányi. 2003, «Simplicity : a unifying principle in cognitive science?», *Trends in cognitive sciences*, vol. 7, n° 1, p. 19–22.
- Chattoe-Brown, E. 2014, «Using agent based modelling to integrate data on attitude change», *Sociological Research Online*, vol. 19, n° 1, p. 16, ISSN 1360-7804.
- Crano, W. D. et J. Sivacek. 1984, «The influence of incentive-aroused ambivalence on overjustification effects in attitude change», *Journal of Experimental Social Psychology*, vol. 20, n° 2, p. 137–158. URL <http://www.sciencedirect.com/science/article/pii/0022103184900179>.

- Crutchfield, R. S. 1955, «Conformity and character.», *American Psychologist*, vol. 10, n° 5, p. 191. URL <http://psycnet.apa.org/journals/amp/10/5/191/>.
- Darwin, C., P. Ekman et P. Prodger. 1998, *The expression of the emotions in man and animals*, Oxford University Press, USA. URL https://books.google.fr/books?hl=fr&lr=&id=TFRtLZSHMcYC&oi=fnd&pg=PR9&dq=ekman&ots=Y2hBfnWu0a&sig=LD0lEcy9dFw_ZT2ViqMIaXUnMKI.
- De Houwer, J., S. Thomas et F. Baeyens. 2001, «Association learning of likes and dislikes : A review of 25 years of research on human evaluative conditioning.», *Psychological Bulletin; Psychological Bulletin*, vol. 127, n° 6, p. 853. URL <http://psycnet.apa.org/journals/bul/127/6/853/>.
- Deci, E. L. et R. M. Ryan. 1980, «The empirical exploration of intrinsic motivational processes», *Advances in experimental social psychology*, vol. 13, n° 2, p. 39–80. URL http://books.google.fr/books?hl=fr&lr=&id=l2pdTKSEnAIC&oi=fnd&pg=PA39&dq=deci+ryan+1980&ots=mrF9k-0oVZ&sig=ss8V_BNMMuw_AfSj-Xnqv5Du0p4.
- Deffuant, G., D. Neau, F. Amblard et G. Weisbuch. 2000, «Mixing beliefs among interacting agents», *Advances in Complex Systems*, vol. 3, n° 01n04, p. 87–98. URL <http://www.worldscientific.com/doi/abs/10.1142/S0219525900000078>.
- Dessalles, J. L. 2007, «Le rôle de l'impact émotionnel dans la communication des événements», URL <http://hal.archives-ouvertes.fr/hal-00188867/>.
- Dimulescu, A. et J.-L. Dessalles. 2009, «Understanding narrative interest : Some evidence on the role of unexpectedness», dans *Proceedings of the 11th Annual Conference of the Cognitive Science Society*, p. 1734–1739. URL <http://csjarchive.cogsci.rpi.edu/proceedings/2009/papers/367/paper367.pdf>.
- Drogoul, A., D. Vanbergue et T. Meurisse. 2003, «Simulation orientée agent : où sont les agents», *Actes des Journées de Rochebrune, Rencontres interdisciplinaires sur les systèmes complexes naturels et artificiels*. URL <ftp://ftp-sop.inria.fr/members/Didier.Parigot/PDFarticle/francesco/AlexisDrogoul2003.pdf>.
- Eagly, A. H. et S. Chaiken. 1993, *The psychology of attitudes.*, Harcourt Brace Jovanovich College Publishers. URL <http://psycnet.apa.org/psycinfo/1992-98849-000>.
- Eagly, A. H., S. Chaiken et W. Wood. 1981, «An attribution analysis of persuasion», *New directions in attribution research*, vol. 3, p. 37–62.
- Eagly, A. H., A. Mladinic et S. Otto. 1994, «Cognitive and affective bases of attitudes toward social groups and social policies», *Journal of Experimental Social Psychology*, vol. 30, n° 2, p. 113–137. URL <http://www.sciencedirect.com/science/article/pii/S0022103184710067>.

- Esses, V. M., G. Haddock et M. P. Zanna. 1993, «Values, stereotypes, and emotions as determinants of intergroup attitudes.», URL <http://doi.apa.org/psycinfo/1993-97388-006>.
- Fazio, R. H. 1995, «Attitudes as object-evaluation associations : Determinants, consequences, and correlates of attitude accessibility.», URL <http://psycnet.apa.org/psycinfo/1995-98997-010>.
- Fazio, R. H. 2007, «Attitudes as object-evaluation associations of varying strength», *Social Cognition*, vol. 25, n° 5, p. 603. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2677817/>.
- Fazio, R. H., J. R. Jackson, B. C. Dunton et C. J. Williams. 1995, «Variability in automatic activation as an unobtrusive measure of racial attitudes : a bona fide pipeline?», *Journal of personality and social psychology*, vol. 69, n° 6, p. 1013. URL <http://psycnet.apa.org/journals/psp/69/6/1013/>.
- Fechner, G. T. 1832, «Repertorium der experimentalphysik : enthaltend eine vollständige zusammenstellung der neuern fortschritte dieser wissenschaft : als supplement zu neuern lehr-und wörterbüchern der physik», URL https://books.google.fr/books?hl=fr&lr=&id=Ww85AAAAcAAJ&oi=fnd&pg=PA1&dq=Fechner&ots=FtiotS8_aM&sig=dFPHKLv5cV0WfcUac2LnaCCy98o.
- Festinger, L. 1957, *A theory of cognitive dissonance*, vol. 2, Stanford University Press. URL http://books.google.fr/books?hl=fr&lr=&id=voeQ-8CASacC&oi=fnd&pg=PA1&dq=a+theory+of+cognitive+dissonance+festinger&ots=9w86Mxuhyx&sig=D1XIqvqWgsUm2yJl_z3SCFD2oYM.
- Finlay, K. A., D. Trafimow et D. Jones. 1997, «Predicting health behaviors from attitudes and subjective norms : Between-subjects and within-subjects analyses», *Journal of Applied Social Psychology*, vol. 27, n° 22, p. 2015–2031. URL <http://onlinelibrary.wiley.com/doi/10.1111/j.1559-1816.1997.tb01638.x/abstract>.
- Fishbein, M. et I. Ajzen. 1975, *Belief, Attitude, Intention and Behavior : An Introduction to Theory and Research*, Addison-Wesley Pub (Sd), ISBN 0201020890.
- Fishbein, M. et I. Ajzen. 1981, «Acceptance, yielding, and impact : Cognitive processes in persuasion», *Cognitive responses in persuasion*, p. 339–359.
- Fishbein, M. et I. Ajzen. 2005, «The influence of attitudes on behavior», *The handbook of attitudes*, p. 173–222. URL http://www.bm.ust.hk/~mark790b/readings/T5bR2_ajzen2005.pdf.
- Fraser, R. S. et B. G. Stacey. «A psychophysical investigation of the influence of attitude on the judgement of social stimuli», vol. 12, n° 4, p. 337–352. URL <http://onlinelibrary.wiley.com/doi/10.1111/j.2044-8260.1973.tb00080.x/pdf>.
- Gargiulo, F., S. Ternes, S. Huet et G. Deffuant. 2010, «An iterative approach for generating statistically realistic populations of households», *PloS one*, vol. 5, n° 1, p. e8828. URL <http://dx.plos.org/10.1371/journal.pone.0008828>.

- Geller, A. et S. J. Alam. 2010, «A socio-political and-cultural model of the war in afghanistan», *International Studies Review*, vol. 12, n° 1, p. 8–30. URL <http://onlinelibrary.wiley.com/doi/10.1111/j.1468-2486.2009.00910.x/full>.
- Gerull, F. C. et R. M. Rapee. 2002, «Mother knows best : effects of maternal modelling on the acquisition of fear and avoidance behaviour in toddlers», *Behaviour research and therapy*, vol. 40, n° 3, p. 279–287. URL <http://www.sciencedirect.com/science/article/pii/S0005796701000134>.
- Glauber, R. J. 1963, «Time-dependent statistics of the ising model», *Journal of mathematical physics*, vol. 4, n° 2, p. 294–307. URL <http://scitation.aip.org/content/aip/journal/jmp/4/2/10.1063/1.1703954>.
- Greenwald, A. 1968, «Cognitive learning, cognitive response to persuasion, and attitude change», *Psychological foundations of attitudes*, p. 147–170. URL <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.19.7121&rep=rep1&type=pdf>.
- Greenwald, A. G., D. E. McGhee et J. L. K. Schwartz. 1998, «Measuring individual differences in implicit cognition : the implicit association test.», *Journal of personality and social psychology*, vol. 74, n° 6, p. 1464. URL <http://psycnet.apa.org/psycinfo/1998-02892-004>.
- Greenwald, C. S. 1977, *Group power : Lobbying and public policy*, Praeger Publishers.
- Haddock, G. et M. P. Zanna. 1997, «Impact of negative advertising on evaluations of political candidates : The 1993 canadian federal election», *Basic and Applied Social Psychology*, vol. 19, n° 2, p. 205–223. URL http://www.tandfonline.com/doi/abs/10.1207/s15324834basp1902_4.
- Hansen, N., S. Müller et P. Koumoutsakos. 2003, «Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES)», *Evolutionary Computation*, vol. 11, n° 1, p. 1–18. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6790790.
- Hart, H. L. A. 2008, *Punishment and responsibility : Essays in the philosophy of law*, Oxford University Press. URL https://books.google.fr/books?hl=fr&lr=&id=aXRCAGAAQBAJ&oi=fnd&pg=PP36&dq=hart+responsibility&ots=-hoWa1cxE6&sig=Bmy26WD8D3Zm_sGpa9duL5BPDJI.
- Hass, R. G. 1981, «Effects of source characteristics on cognitive responses and persuasion», *Cognitive responses in persuasion*, vol. 2.
- Hegselmann, R. et U. Krause. 2002, «Opinion dynamics and bounded confidence models, analysis, and simulation», *Journal of Artificial Societies and Social Simulation*, vol. 5, n° 3. URL http://engineering.ucsb.edu/~mirtabatabaei/Opinion%20Dynamics%20and%20Bounded%20Confidence%20Models_Krause.pdf.
- Heider, F. 1982, *The psychology of interpersonal relations*, Lawrence Erlbaum. URL <http://books.google.fr/books?hl=fr&lr=&id=Zh6TDmayL0AC&oi=fnd&pg=PA1&dq=heider+the+psychology+of+interpersonal&ots=x1TPe443dN&sig=kiA-MFxsDX3NN47JJYfWU9xtv0E>.

- Hewgill, M. A. et G. R. Miller. 1965, «Source credibility and response to fear-arousing communications», URL <http://www.tandfonline.com/doi/pdf/10.1080/03637756509375436>.
- Hogg, M. A. et D. Abrams. 1990, «Social motivation, self-esteem and social identity», *Social identity theory : Constructive and critical advances*, vol. 28, p. 47.
- Hovland, C. I., I. L. Janis et H. H. Kelley. 1953, «Communication and persuasion; psychological studies of opinion change.», URL <http://doi.apa.org/psycinfo/1954-05952-000>.
- Hovland, C. I., A. A. Lumsdaine et F. D. Sheffield. 1949, «Experiments on mass communication.(studies in social psychology in world war II, vol. 3.)», URL <http://psycnet.apa.org/psycinfo/1950-00627-000>.
- Inkeles, A. 1969, «Participant citizenship in six developing countries», *American Political Science Review*, vol. 63, n° 4, p. 1120–1141. URL http://journals.cambridge.org/abstract_S0003055400263235.
- Insko, C. A. et K. W. Butzine. 1967, «Rapport, awareness, and verbal reinforcement of attitude.», *Journal of personality and social psychology*, vol. 6, n° 2, p. 225. URL <http://psycnet.apa.org/journals/psp/6/2/225/>.
- Jones, E. E. et V. A. Harris. 1967, «The attribution of attitudes», *Journal of Experimental Social Psychology*, vol. 3, n° 1, p. 1–24. URL <http://www.sciencedirect.com/science/article/pii/0022103167900340>.
- Kahneman, D., D. Kahneman et A. Tversky. 2003, «Experienced utility and objective happiness : A moment-based approach», *The psychology of economic decisions*, vol. 1, p. 187–208. URL http://books.google.fr/books?hl=fr&lr=&id=f0I31h_G6UkC&oi=fnd&pg=PA187&dq=xperienced+utility+and+objective+happiness:+a+moment-based+approach&ots=-ARqQbEr4&sig=_EXvUu0Rt9Rdzm_K_LcPurfkTS0.
- Katz, D. 1960, «The functional approach to the study of attitudes», *Public opinion quarterly*, vol. 24, n° 2, p. 163–204. URL <http://poq.oxfordjournals.org/content/24/2/163.short>.
- Kelley, H. 1973, «The processes of causal attribution.», *American psychologist*, vol. 28, n° 2. URL <http://psycnet.apa.org/journals/amp/28/2/107/>.
- Kelman, H. C. 1958, «Compliance, identification, and internalization : Three processes of attitude change», *Journal of conflict resolution*, p. 51–60. URL <http://www.jstor.org/stable/172844>.
- Kelman, H. C. 1961, «Processes of opinion change», *Public opinion quarterly*, vol. 25, n° 1, p. 57–78. URL <http://poq.oxfordjournals.org/content/25/1/57.short>.
- Kish, L. 1965, «Survey sampling», URL <http://www.citeulike.org/group/108/article/553273>.
- Kohonen, T. 1998, «The self-organizing map», *Neurocomputing*, vol. 21, n° 1, p. 1–6. URL <http://www.sciencedirect.com/science/article/pii/S0925231298000307>.

- Kolmogorov, A. 1956, «On the shannon theory of information transmission in the case of continuous signals», *IRE Transactions on Information Theory*, vol. 2, n° 4, doi :10.1109/TIT.1956.1056823, p. 102–108, ISSN 0096-1000.
- Kottonau, J. et C. Pahl-Wostl. 2004, «Simulating political attitudes and voting behavior», URL <http://jasss.soc.surrey.ac.uk/7/4/6.html>.
- Kraus, S. J. 1995, «Attitudes and the prediction of behavior : A meta-analysis of the empirical literature», *Personality and Social Psychology Bulletin*, vol. 21, n° 1, p. 58–75. URL <http://psp.sagepub.com/content/21/1/58.short>.
- Kruglanski, A. W. et E. P. Thompson. 1999, «Persuasion by a single route : A view from the unimodel», *Psychological Inquiry*, vol. 10, n° 2, p. 83–109. URL <http://www.tandfonline.com/doi/abs/10.1207/S15327965PL100201>.
- Lasswell, H. D. 1948, «The structure and function of communication in society», *The communication of ideas*, vol. 37, p. 215–228. URL <http://www.dhpescu.org/media/elip/The%20structure%20and%20function%20of.pdf>.
- Latane, B. 1981, «The psychology of social impact.», *American psychologist*, vol. 36, n° 4, p. 343. URL <http://psycnet.apa.org/journals/amp/36/4/343/>.
- Leyens, J.-P., P. M. Paladino, R. Rodriguez-Torres, J. Vaes, S. Demoulin, A. Rodriguez-Perez et R. Gaunt. 2000, «The emotional side of prejudice : The attribution of secondary emotions to ingroups and outgroups», *Personality and Social Psychology Review*, vol. 4, n° 2, p. 186–197. URL <http://psr.sagepub.com/content/4/2/186.short>.
- Likert, R. 1932, «A technique for the measurement of attitudes.», *Archives of psychology*. URL <http://psycnet.apa.org/psycinfo/1933-01885-001>.
- Lord, C. G., L. Ross et M. R. Lepper. 1979, «Biased assimilation and attitude polarization : the effects of prior theories on subsequently considered evidence.», *Journal of personality and social psychology*, vol. 37, n° 11, p. 2098. URL <http://psycnet.apa.org/psycinfo/1981-05421-001>.
- Maio, G. et G. Haddock. 2010, *The Psychology of Attitudes and Attitude Change*, SAGE Publications Ltd, ISBN 978-1-4129-2975-2.
- Maio, G. R. et G. Haddock. 2004, «Theories of attitude : Creating a witches' brew», URL <http://orca.cf.ac.uk/34741/>.
- McGuire, W. J. 1968, «Personality and attitude change : An information-processing theory», *Psychological foundations of attitudes*, p. 171–196. URL <http://uts.cc.utexas.edu/~tecas/syllabi2/adv382jfall2002/readings/McGuire.pdf>.
- Milgram, S. 1967, «The small world problem», *Psychology today*, vol. 2, n° 1, p. 60–67. URL http://measure.igpp.ucla.edu/GK12-SEE-LA/Lesson_Files_09/Tina_Wey/TW_social_networks_Milgram_1967_small_world_problem.pdf.

- Moscovici, S. et J.-C. Abrieu. 1984, *Psychologie sociale*, Presses universitaires de France. URL <http://www.lavoisier.fr/livre/notice.asp?ouvrage=2608260>.
- Norman, R. 1975, «Affective-cognitive consistency, attitudes, conformity, and behavior», *Journal of Personality and Social Psychology*, vol. 32, n° 1, p. 83. URL <http://psycnet.apa.org/journals/psp/32/1/83/>.
- Nowak, A., J. Szamrej et B. Latane. 1990, «From private attitude to public opinion : A dynamic theory of social impact», *Psychological Review*, vol. 97, n° 3, p. 362. URL <http://psycnet.apa.org/journals/rev/97/3/362/>.
- Olson, M. A. et R. H. Fazio. 2001, «Implicit attitude formation through classical conditioning», *Psychological Science*, vol. 12, n° 5, p. 413–417. URL <http://pss.sagepub.com/content/12/5/413.short>.
- Olson, M. A. et R. H. Fazio. 2003, «Relations between implicit measures of prejudice what are we measuring?», *Psychological Science*, vol. 14, n° 6, p. 636–639. URL <http://pss.sagepub.com/content/14/6/636.short>.
- Ortony, A., G. L. Clore et A. Collins. 1990, *The cognitive structure of emotions*, Cambridge university press. URL <https://books.google.fr/books?hl=fr&lr=&id=dA3JEEAp6TsC&oi=fnd&pg=PR9&dq=Ortony,+A.,+Clore,+G.L.,+Collins,+A.:+The+Cognitive+Structure+of+Emotions.+&ots=HVMDZsaUj8&sig=5gxmcs5ZNaBHU0z3drS6G7SqDc>.
- Paulhus, D. L. 1998, «Interpersonal and intrapsychic adaptiveness of trait self-enhancement : A mixed blessing?», *Journal of personality and social psychology*, vol. 74, n° 5, p. 1197. URL <http://psycnet.apa.org/psycinfo/1998-01923-007>.
- Perlman, D. et S. Oskamp. 1971, «The effects of picture content and exposure frequency on evaluations of negroes and whites», *Journal of Experimental Social Psychology*, vol. 7, n° 5, p. 503–514. URL <http://www.sciencedirect.com/science/article/pii/0022103171900126>.
- Petty, R. et J. Cacioppo. 1986, «The elaboration likelihood model of persuasion», *Advances in experimental social psychology*, vol. 19, n° 1, p. 123–205. URL <http://books.google.com/books?hl=fr&lr=&id=UPL-PadJSRQC&oi=fnd&pg=PA123&dq=Elaboration+Likelihood+Model+&ots=sdr3ij4GLM&sig=0sU0v5KjopN6ySxtMdpa0zeBm0o>.
- Petty, R. E., P. Briñol et K. G. DeMarree. 2007, «The meta-cognitive model (MCM) of attitudes : Implications for attitude measurement, change, and strength», *Social Cognition*, vol. 25, n° 5, p. 657–686. URL <http://guilfordjournals.com/doi/abs/10.1521/soco.2007.25.5.657>.
- Petty, R. E., G. L. Wells et T. C. Brock. 1976, «Distraction can enhance or reduce yielding to propaganda : Thought disruption versus effort justification», *Journal of Personality and Social Psychology*, vol. 34, n° 5, p. 874. URL <http://psycnet.apa.org/journals/psp/34/5/874/>.

- Plutchik, R. 1980, «A general psychoevolutionary theory of emotion», *Theories of emotion*, vol. 1. URL https://books.google.fr/books?hl=fr&lr=&id=Ie1GBQAAQBAJ&oi=fnd&pg=PA3&dq=Robert+Plutchik&ots=Mx3_ribCBc&sig=LDZxt242rd5LDt03ch6TzPAWR1I.
- Pratkanis, A. R. et A. G. Greenwald. 1989, «A sociocognitive model of attitude structure and function», *Advances in experimental social psychology*, vol. 22, p. 245–285. URL <http://www.sciencedirect.com/science/article/pii/S006526010860310X>.
- Rogers, E. M. 2010, *Diffusion of innovations*, Simon and Schuster. URL <https://books.google.fr/books?hl=fr&lr=&id=v1ii4QsB7jIC&oi=fnd&pg=PR15&dq=Everett+M.+Rogers,+Diffusion+of+Innovations&ots=DKVwrPXtcR&sig=jQZ6dm5EzvJrfeVC8v-Cx53kNXs>.
- Rosenberg, M. J. et C. I. Hovland. 1960, «Cognitive, affective, and behavioral components of attitudes», *Attitude organization and change : An analysis of consistency among attitude components*, vol. 3, p. 1–14.
- Sakoda, J. M. 1971, «The checkerboard model of social interaction», *The Journal of Mathematical Sociology*, vol. 1, n° 1, p. 119–132. URL <http://www.tandfonline.com/doi/abs/10.1080/0022250X.1971.9989791>.
- Schwartz, I. S. et D. M. Baer. 1991, «Social validity assessments : Is current practice state of the art ?», *Journal of Applied Behavior Analysis*, vol. 24, n° 2, p. 189–204. URL <http://onlinelibrary.wiley.com/doi/10.1901/jaba.1991.24-189/abstract>.
- Shannon, C. E. et W. Weaver. 1949, «The mathematical theory of information», URL <http://www.citeulike.org/group/160/article/753060>.
- Sherif, M. et C. Hovland. *Social judgment theory*, New Haven, CT : Yale University Press.
- Siero, F. W. et B. J. Doosje. «Attitude change following persuasive communication : Integrating social judgment theory and the elaboration likelihood model», vol. 23, n° 5, doi :10.1002/ejsp.2420230510, p. 541–554, ISSN 1099-0992. URL <http://onlinelibrary.wiley.com/doi/10.1002/ejsp.2420230510/abstract>.
- Silverman, B. G., D. Pietrocola, N. Weyer, R. Weaver, N. Esomar, R. Might et D. Chandrasekaran. 2009, «Nonkin village : An embeddable training game generator for learning cultural terrain and sustainable counter-insurgent operations», dans *Agents for Games and Simulations*, Springer, p. 135–154. URL http://link.springer.com/chapter/10.1007/978-3-642-11198-3_10.
- Simon, H. A. 1955, «A behavioral model of rational choice», *The quarterly journal of economics*, vol. 69, n° 1, p. 99–118. URL <http://qje.oxfordjournals.org/content/69/1/99.short>.
- Smith, M. B. 1956, *Opinions and personality*, Wiley.
- Smith, R. 2007, *The utility of force : the art of war in the modern world*, Knopf, New York, ISBN 0307265625 9780307265623.

- Snyder, M. 1974, «Self-monitoring of expressive behavior.», *Journal of personality and social psychology*, vol. 30, n° 4, p. 526. URL <http://psycnet.apa.org/journals/psp/30/4/526/>.
- Squazzoni, F. 2008, «The micro-macro link in social simulation», *Sociologica*, vol. 2, n° 1, p. 0–0. URL <http://www.rivisteweb.it/doi/10.2383/26578>.
- Steele, C. M. 1988, «The psychology of self-affirmation : Sustaining the integrity of the self», *Advances in experimental social psychology*, vol. 21, p. 261–302. URL http://books.google.com/books?hl=fr&lr=&id=oNZ18G2D_qsC&oi=fnd&pg=PA261&dq=steele&ots=q-wc-vikou&sig=WS-9F6n9ThNkoHuJPqZ9ZIwz5_k.
- Stevenson, A. 2010, *Oxford Dictionary of English*, OUP Oxford, ISBN 978-0-19-957112-3.
- Swift, J. et J.-J. Courtine. 2007, *L Art du mensonge politique - Le Mentir vrai*, édition revue et augmentée éd., Jérôme Millon, ISBN 978-2-84137-205-8.
- Thiriot, S. 2009, *Vers une modélisation plus réaliste de la diffusion d'innovations l'aide de la simulation multi-agents*, thèse de doctorat, Université Pierre et Marie Curie-Paris VI. URL <http://halshs.archives-ouvertes.fr/tel-00519016/>.
- Thiriot, S. et J.-D. Kant. 2007, «Représenter les croyances par des réseaux associatifs pour simuler la diffusion d'innovations», *ESSA*, vol. 7, p. 2. URL http://www-desir.lip6.fr/publications/pub_937_1_thiriot_kant_jfsma_07.pdf.
- Thiriot, S. et J.-D. Kant. 2008, «Generate country-scale networks of interaction from scattered statistics», dans *The fifth conference of the European social simulation association, Brescia, Italy*, vol. 240. URL http://samuelthiriot.res-ear.ch/IMG/pdf/thiriot_2008_3.pdf.
- Thurstone, L. L. et E. J. Chave. 1929, «The measurement of attitude.», URL <http://psycnet.apa.org/psycinfo/1929-04902-000>.
- Urbig, D. et R. Malitz. 2007, «Drifting to more extreme but balanced attitudes : Multidimensional attitudes and selective exposure», *ESSA, Toulouse*. URL <http://www.diemo.de/documents/UrbigMalitz-ESSA2007.pdf>.
- Vitanyi, P. M. B. et M. Li. 2002, «Simplicity, information, kolmogorov complexity, and prediction», *Simplicity, Inference and Modelling*, p. 135–155.
- Wang, X. F. et G. Chen. 2003, «Complex networks : small-world, scale-free and beyond», *Circuits and Systems Magazine, IEEE*, vol. 3, n° 1, p. 6–20. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1228503.
- Weiss, H. M. 1977, «Subordinate imitation of supervisor behavior : The role of modeling in organizational socialization», *Organizational Behavior and Human Performance*, vol. 19, n° 1, p. 89–105. URL <http://www.sciencedirect.com/science/article/pii/0030507377900563>.

- Wicker, A. W. 1969, «Attitudes versus actions : The relationship of verbal and overt behavioral responses to attitude objects», *Journal of social issues*, vol. 25, n° 4, p. 41–78. URL <http://onlinelibrary.wiley.com/doi/10.1111/j.1540-4560.1969.tb00619.x/abstract>.
- Winkielman, P. et J. Cacioppo. 2001, «Mind at ease puts a smile on the face : psychophysiological evidence that processing facilitation elicits positive affect.», *Journal of personality and social psychology*, vol. 81, n° 6, p. 989. URL <http://psycnet.apa.org/journals/psp/81/6/989/>.
- Yager, R. R. 1988, «On ordered weighted averaging aggregation operators in multicriteria decisionmaking», *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 18, n° 1, p. 183–190. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=87068.
- Zajonc, R. B. 1968, «Attitudinal effects of mere exposure.», *Journal of personality and social psychology*, vol. 9, n° 2p2, p. 1. URL <http://psycnet.apa.org/journals/psp/9/2p2/1/>.
- Zanna, M. P. et J. K. Rempel. 1988, *Attitudes : A new look at an old concept. The social psychology of attitudes (sid. -)*, New York : Cambridge University Press. I.

Ce document a été réalisé avec L^AT_EX.

Final Version as of 2 décembre 2015 (`classicthesis` version 0.1).