

Bruno Cancellieri, ChatGPT et Claude.ai

Psychologie du jugement

Comment le comportement social est régi par les jugements
que nous anticipons et de ceux que nous formulons

4 août 2026

Nous agissons en fonction de jugements que presque personne n'exprime. Avant de parler, de nous exposer, de choisir ce que nous allons montrer et ce que nous allons taire, notre esprit évalue la façon dont nous serons jugés, et réagit davantage à cette estimation qu'à la situation elle-même.

Cette estimation ne repose pas sur une lecture objective des autres. Elle se construit principalement à partir de la façon dont nous nous percevons nous-mêmes et dont nous percevons notre relation aux autres ; ce qu'ils laissent transparaître entre en ligne de compte, mais pèse moins que nous ne le croyons.

Nos jugements sont observables, et nous anticipons inconsciemment d'être jugés pour la manière dont nous jugeons : cette anticipation entre ainsi dans la formation des critères selon lesquels nous évaluons.

Le jugement qui en résulte est perçu et jugé par les autres même lorsqu'il n'est pas exprimé explicitement. C'est cette *récurtivité* qui est au cœur de cet essai : une évaluation qui devient elle-même objet d'évaluation, et qui, de ce fait, n'est jamais seulement une description de ce qui est évalué. D'autre part, les jugements adaptés à notre interlocuteur afin d'être approuvés ou de ne pas susciter de conflits ne restent pas une simple posture : en les exprimant à plusieurs reprises, nous finissons par croire à ce que nous avons dit.

Le mécanisme, en soi, n'est pas défectueux : il rend possibles la confiance, la coopération et la réputation, tout comme la compétition et le conflit. Le problème ne réside pas dans son fonctionnement, mais dans son calibrage — il est calibré pour un environnement où celui qui jugeait prenait des décisions, et où, aujourd'hui encore, on continue de réagir à des jugements qui n'auront aucune conséquence.

Cet essai n'enseigne pas à s'en débarrasser. Il propose un modèle pour le reconnaître lorsqu'il se met en marche, un critère pour distinguer les interventions susceptibles de mieux l'ajuster de celles qui l'alimentent, et deux questions à garder à l'esprit : « Cette personne peut-elle changer quelque chose dans ma vie ? » et « Que penserais-je si mon jugement n'était observable par personne ? » Elles n'empêchent pas le système d'agir : elles aident à observer un mécanisme qui, en règle générale, agit sans être remarqué.

Sommaire

Abstract	5
Note	6
Introduction	7
PREMIÈRE PARTIE — Le mécanisme	11
1 Évaluer, c’est attribuer une valeur	12
2 Quand l’évaluation devient réursive	17
3 À quoi sert le jugement ?	20
4 Le jugement présumé	24
5 Être évalué : d’où vient le jugement présumé ?	28
6 L’observateur que nous portons en nous	33
DEUXIÈME PARTIE — Les deux facettes	37
7 Juger, c’est s’exposer	38
8 Exprimer un jugement nous change	44
9 Un mécanisme calibré pour un monde qui n’existe plus	50
TROISIÈME PARTIE — Que faire de tout cela ?	54
10 Pourquoi les remèdes intuitifs aggravent la situation	55
11 Le principe du recalibrage	60
12 Deux questions fondamentales	67
13 Le déclin du recalibrage	72
14 Quand mettre fin à l’observation	75
Conclusion	79
Annexe I — Références	81
Annexe II — Ce qu’affirme le modèle	87
Annexe III — Questions pour approfondir	91
Annexe IV — Glossaire	94

« La société est un entrelacement et un jeu d'esprits. J'imagine ton esprit, et surtout ce que ton esprit pense du mien, ainsi que ce que ton esprit pense de ce que le mien pense du tien. Je dévoile mon esprit devant le tien, et j'attends de toi que tu dévoiles le tien devant le mien. »

Charles Horton Cooley, 1902

Abstract

« **Psychologie du jugement** », de Bruno Cancellieri, rédigé en collaboration avec ChatGPT et Claude.ai, propose un modèle du jugement articulé autour de sa récursivité : on évalue quelqu'un ou quelque chose, et cette évaluation peut à son tour être évaluée.

La thèse centrale porte sur la représentation que chacun se fait du jugement d'autrui — le jugement présumé — et soutient qu'elle se construit principalement à partir de la façon dont nous nous percevons nous-mêmes, tandis que le retour d'information provenant des autres n'a qu'un poids réduit et variable.

Deux paramètres indépendants régissent sa construction et son maintien : la présupposition, c'est-à-dire dans quelle mesure l'estime est déjà fixée avant même d'observer, et la dépréciation, c'est-à-dire dans quelle mesure le retour d'information qui la contredit est dévalorisé. Chacun de ces paramètres peut être élevé ou faible, et les quatre combinaisons correspondent à autant de profils : (1) l'anxiété sociale, où les deux sont élevés ; (2) le narcissisme, où seule la minimisation est élevée ; (3) une image de soi instable, où seule la présupposition est élevée et où la perception de la façon dont on est perçu oscille à chaque retour d'information ; et (4) le fonctionnement le mieux calibré, où les deux sont faibles.

L'essai aborde ensuite l'aspect actif : exprimer un jugement expose celui qui l'émet, et le jugement adapté à l'interlocuteur tend à devenir sincère. Le mécanisme, en soi, n'est pas défectueux — il rend possibles la confiance, la coopération et la réputation, tout comme la compétition et le conflit — mais il est calibré pour un environnement où celui qui jugeait prenait des décisions, alors qu'aujourd'hui, il réagit également à des jugements qui n'auront aucune conséquence.

La dernière partie montre pourquoi les remèdes spontanés, qu'il s'agisse de se contrôler ou de rechercher des confirmations, ont tendance à alimenter le mécanisme plutôt qu'à le corriger, et propose un critère pour évaluer toute intervention. Il en découle deux questions à garder à l'esprit : « Cette personne peut-elle changer quelque chose dans ma vie ? » et « Que penserais-je si mon jugement n'était observable par personne ? » — la première pour l'aspect passif, la seconde pour l'aspect actif. Elles n'empêchent pas le système de réagir : elles rendent observable un mécanisme qui opère à l'abri de notre propre inspection.

Cet essai ne revendique pas d'originalité théorique : l'architecture existe déjà dans la clinique de l'anxiété sociale et dans les modèles d'élaboration prédictive. La contribution déclarée est la généralisation du pathologique au fonctionnement ordinaire et le lien entre le volet passif et le volet actif.

Note

Cet article est le fruit d'un dialogue avec ChatGPT et Claude.ai, selon une méthode que j'ai également adoptée pour d'autres travaux : je propose un thème, je discute avec ces outils de la structure et du contenu, puis je leur fais rédiger et corriger les textes au cours d'un échange structuré, au cours duquel les modifications sont proposées, argumentées et discutées ensemble.

Le cadre théorique spécifique se développe au fil de ces dialogues, mais s'inscrit dans une formation psychologique et philosophique que j'ai acquise de manière autonome bien avant de recourir à ces outils.

La responsabilité des thèses défendues, ainsi que de leurs éventuelles faiblesses, m'incombe toutefois entièrement.

Bruno Cancellieri

Introduction

Ce que défend cet essai

Il convient de dissiper deux attentes, toutes deux raisonnables mais aucune des deux pertinente. Il ne s'agit pas d'un essai sur le jugement en tant que processus décisionnel — heuristiques, distorsions, estimations de probabilité —, ce qui constitue un autre domaine de recherche. Et ce n'est pas non plus un essai qui incite à juger moins.

Sur ce dernier point, il convient de noter que l'italien est moins univoque que le mot ne le laisse croire. « Ne pas juger » considère le jugement comme un défaut ; mais « avoir du jugement », « une personne sensée », et même la « dent de sagesse » — ainsi nommée parce qu'elle apparaît à l'âge où l'on suppose que le jugement s'est formé — le considèrent comme une vertu, dont l'absence est un reproche. La langue rend compte de ces deux aspects, et ce n'est pas par inadvertance : le jugement est une faculté nécessaire, et comme toute faculté, il peut être exercé bien ou mal.

Le jugement dont il est question dans ces pages est une opération ordinaire, en grande partie automatique, sans laquelle la vie sociale ne fonctionnerait pas. Il est décrit, non pas déploré — ni loué.

Chaque jour, nous portons des jugements. Nous disons qu'un film est ennuyeux, qu'une personne est digne de confiance, qu'une décision politique est erronée. Et en même temps, nous imaginons sans cesse comment les autres nous jugent. Une grande partie de la vie sociale se déroule au sein de cet entrelacement de jugements exprimés ou présumés.

Il convient de distinguer d'emblée deux termes qui, par la suite, auront des significations différentes. *Évaluer* et *juger* sont tous deux des actes d'attribution de valeur, et en ce sens, le jugement est une forme d'évaluation. Mais évaluer, c'est estimer la valeur que quelque chose a pour nous, en fonction de nos besoins, tandis que juger, c'est transformer cette estimation en un verdict qui influence une relation sociale. Lorsque l'objet du jugement est une personne, le verdict influence la manière dont nous la traiterons ou nos attentes quant à la façon dont elle nous traitera ; lorsqu'il s'agit d'une chose, il expose néanmoins celui qui l'a émis au jugement d'autrui. Le premier chapitre reprendra cette distinction de manière détaillée ; ici, il suffit de savoir que ces deux termes ne sont pas synonymes.

Nous sommes spontanément enclins à penser que ces deux processus sont indépendants. Nous croyons que, lorsque nous jugeons quelque chose, notre jugement dépend principalement des caractéristiques de ce que nous jugeons. Et nous croyons que, lorsque nous imaginons le jugement des autres à notre égard, notre représentation dépend avant tout de ce qu'ils font, disent ou montrent.

La thèse de cet essai est différente.

Dans les situations sociales, les jugements que nous formulons sont également influencés par les conséquences que ces jugements auront pour nous. Et la représentation que nous nous faisons du jugement des autres découle, dans une large mesure, non pas de l'observation directe des autres, mais de notre propre perception de nous-mêmes.

Il en résulte un système qui utilise comme information principalement ce qu'il a lui-même produit.

En résumé : le système effectue un traitement récursif sur une donnée non observée mais déduite, en la construisant principalement à partir de sa propre perception de soi.

Cette formulation résume le principe théorique de cet essai.

Par « *récursif* », on n'entend pas une réflexion consciente du type « je pense que tu penses que je pense ». La recherche montre que la récursivité spontanée est étonnamment limitée : dans la plupart des situations, elle ne dépasse pas un ou deux niveaux. Il s'agit d'un processus rapide, automatique et en grande partie inconscient.

Par « *information déduite* », on n'entend pas une information observée avec peu de précision. Le problème ne concerne pas la qualité de la mesure, mais son origine. Dans la plupart des interactions sociales, nous ne connaissons pas réellement le jugement que les autres portent sur nous. L'esprit comble cette lacune en construisant une représentation qui découle principalement de la façon dont nous nous percevons nous-mêmes et de nos expériences antérieures avec les autres en général.

Ce que cet essai transmet

Ce principe n'est pas né avec le présent essai.

Dans mes travaux précédents, notamment **Le jugement présumé** et **L'autogouvernance sociale**, j'avais déjà introduit les notions de jugement présumé, de superviseur social, de coût relationnel attendu, de désadaptation évolutive et leur lien avec le *traitement prédictif*. Le présent ouvrage ne propose donc pas une théorie formulée ex novo, mais représente la consolidation d'un parcours de recherche déjà entamé, réorganisé, corrigé et soumis à une vérification critique systématique.

La littérature scientifique contient elle aussi déjà une partie importante de l'architecture mentale décrite dans ces pages. Le modèle de Clark et Wells sur l'anxiété sociale décrit un système qui construit l'image de ce à quoi on apparaît aux autres principalement à partir d'informations internes plutôt que de l'observation. Les modèles de *traitement prédictif* formalisent une architecture très similaire en termes computationnels. D'autres axes de recherche ont étudié, séparément, les métaperceptions, la réputation, le jugement en tant que signal social, la falsification des préférences et la transformation des jugements par la communication.

La principale contribution de cet essai consiste à rassembler ces traditions au sein d'un cadre théorique unifié, en les intégrant à la démarche développée dans les travaux précédents.

La proposition la plus audacieuse

La proposition la plus importante, cependant, ne concerne pas cette synthèse.

Elle consiste en une prédiction.

Les modèles cliniques décrivent le mécanisme que nous venons d'illustrer comme caractéristique de l'anxiété sociale et d'autres troubles dans lesquels la représentation du jugement d'autrui est fortement dominée par la perception de soi. Le présent essai propose au contraire que ce même mécanisme relève du fonctionnement social ordinaire et que l'anxiété sociale en constitue le cas extrême.

C'est l'affirmation la plus risquée du modèle, et c'est pour cette raison qu'elle n'est pas présentée comme un fait établi, mais comme une prédiction empirique.

Certains résultats de la recherche sont compatibles avec cette hypothèse. Les études taxométriques suggèrent que l'anxiété sociale s'inscrit sur un continuum plutôt que d'être organisée en catégories nettement distinctes. Toutefois, cela concerne la répartition des symptômes, et non celle des mécanismes psychologiques. Un continuum phénoménologique ne prouve pas que les personnes présentant différents niveaux d'anxiété sociale utilisent le même processus avec une intensité différente. C'est précisément ce que propose le modèle et ce qui devra être vérifié.

Ce que l'essai apporte

Outre cette prédiction, l'essai introduit certaines extensions et corrections par rapport aux travaux antérieurs.

La première concerne ce qui détermine à quel moment le jugement des autres devient réellement important. Dans les travaux précédents, cela était formulé en termes de poids social de l'observateur et de probabilité d'interactions futures. Dans cet essai, cela est redéfini comme le pouvoir que l'autre exerce sur nous. La différence n'est pas d'ordre terminologique. Deux personnes peuvent être amenées à interagir quotidiennement sans pour autant que l'une d'elles ait la moindre raison de gérer sa réputation auprès de l'autre : la relation entre maître et esclave en est un exemple évident. La probabilité d'interaction ne suffit pas à expliquer cette asymétrie ; le pouvoir, lui, le fait.

La deuxième correction découle d'une prédiction du modèle qui a été réfutée lors de sa conception. L'analyse de la littérature sur le narcissisme a montré que la métaperception ne peut être décrite comme une simple conséquence de l'autoperception. Le modèle distingue donc deux paramètres indépendants : dans quelle mesure la métaperception est construite à partir de l'autoperception et quel poids est attribué aux informations susceptibles de la corriger.

La troisième consiste à introduire une prédiction sur la profondeur de la récursivité : le nombre de niveaux récursifs devrait augmenter avec l'anxiété sociale. Celle-ci n'est pas non plus présentée comme un résultat acquis, mais comme une hypothèse vérifiable.

Enfin, l'essai intègre les corrections apparues au cours de la vérification. Certaines hypothèses ont été abandonnées, d'autres revues à la baisse, d'autres encore reformulées.

Le champ d'application du modèle

Le modèle ne prétend pas expliquer l'ensemble de la vie sociale.

Son champ d'application est plus restreint. Il concerne les situations dans lesquelles une évaluation peut, à son tour, faire l'objet d'une évaluation. Dans ces cas-là, le jugement ne se contente pas de décrire ce qui est évalué, mais révèle également quelque chose sur celui qui juge. C'est cette structure récursive, et non le simple fait d'évaluer, qui constitue l'objet de cet essai.

Ses implications pourraient toutefois être plus vastes qu'il n'y paraît.

Si ce mécanisme influence les jugements que nous exprimons et ceux que nous attribuons aux autres, il pourrait également influencer ce que nous sommes prêts à prendre en considération. Avant même de nous demander si une idée est vraie ou fausse, l'esprit pourrait en évaluer implicitement

les conséquences relationnelles : qu'est-ce qui changerait, dans ma vie sociale, si je prenais cette idée au sérieux ? Cette hypothèse sera abordée dans la deuxième partie de cet essai.

Ce que cet essai ne promet pas

Comprendre ce mécanisme ne signifie pas apprendre à mieux contrôler son image sociale.

L'un des résultats les plus solides de la recherche sur l'anxiété sociale montre en effet qu'une attention excessive portée à soi-même contribue à entretenir le problème. L'objectif de cet essai est tout autre : aider le lecteur à reconnaître que le jugement présumé est souvent une construction de son propre esprit, et que les remèdes intuitifs — réfléchir davantage, se surveiller en permanence, rechercher des assurances — tendent à renforcer le mécanisme plutôt qu'à le corriger.

Comment est-il organisé ?

Cet essai s'articule en trois parties.

La première décrit le fonctionnement du système : qu'est-ce qu'une évaluation, que se passe-t-il lorsqu'elle devient réursive, pourquoi le jugement des autres compte et par quelle représentation il agit.

La deuxième montre comment le système opère sur deux fronts : lorsque nous imaginons être jugés, et lorsque c'est nous qui jugeons — jusqu'au point où nous adaptons un jugement à celui qui nous écoute et finissons par y croire vraiment. Elle se termine en expliquant pourquoi un mécanisme calibré pour de petits groupes continue de réagir dans un monde d'inconnus.

La troisième aborde le changement : pourquoi les remèdes spontanés échouent, dans quelles conditions le système peut se recalibrer, quels outils en découlent, pourquoi tant ce recalibrage que ces outils finissent par perdre leur efficacité, et quand il convient d'interrompre l'introspection.

Quatre annexes regroupent les références, ce que le modèle affirme, quelques questions à se poser face à un jugement déjà formé, et les termes utilisés au sens technique.

Deux questions fondamentales

Les pages qui suivent convergent finalement vers deux questions fondamentales.

- Cette personne peut-elle changer quelque chose dans ma vie ?
- Que penserais-je si mon jugement n'était observable par personne ?

Ce sont des questions que chacun peut se poser en quelques secondes, face à un jugement reçu ou au moment où il en formule un. Cet essai vise à expliquer pourquoi il vaut la peine de se les poser, et ce que l'on peut raisonnablement attendre de la réponse.

À ces deux questions s'en ajoutent d'autres, rassemblées dans l'annexe « *Questions pour approfondir* » : ce ne sont pas des outils à garder à portée de main au moment où le jugement se forme, mais des aides à consulter à tête reposée, lorsqu'il reste à décider quoi faire d'un jugement déjà formulé.

PREMIÈRE PARTIE — Le mécanisme

1 Évaluer, c'est attribuer une valeur

Chaque instant de notre vie mentale est traversé par des évaluations.

Nous entrons dans une pièce et la percevons comme accueillante ou hostile. Nous goûtons un aliment et le trouvons agréable ou désagréable. Nous lisons une information et la considérons comme intéressante, sans importance ou préoccupante. Un bruit soudain nous alarme, une musique nous détend, un visage nous inspire confiance ou méfiance.

Ces opérations sont si rapides et habituelles qu'elles passent presque inaperçues. Elles constituent pourtant l'une des fonctions fondamentales de l'esprit.

Avant même de décider comment agir, nous devons déterminer ce qui, pour nous, a de la valeur (positive ou négative).

Dans cet essai, nous appellerons « évaluation » précisément cette opération élémentaire : l'attribution d'une valeur positive ou négative, d'une certaine intensité, à une personne, un objet, un événement, une idée ou une situation.

Cette définition est volontairement large. Elle ne concerne pas encore le jugement social, la réputation ou les relations. Elle concerne une fonction cognitive générale, sans laquelle aucun comportement orienté vers un but ne serait possible.

Un esprit incapable d'évaluer ne saurait distinguer ce qui favorise ses objectifs de ce qui les entrave, ce qui mérite l'attention de ce qui peut être ignoré, ce qu'il convient de poursuivre de ce qu'il convient d'éviter.

Évaluer, c'est attribuer une signification pratique à l'expérience.

Toutes les évaluations ne sont pas conscientes. Au contraire, la plupart se font automatiquement : nous percevons plus facilement leur résultat que le processus qui les a générées. Nous sentons que quelque chose nous plaît ou nous dérange bien avant de pouvoir en expliquer la raison.

Un niveau déjà étudié, et un autre plus large

La psychologie étudie depuis longtemps ce niveau fondamental.

En particulier, les théories *de l'évaluation*, développées par des auteurs tels que Richard Lazarus et Klaus Scherer, décrivent la manière dont les événements sont évalués par rapport à nos objectifs, à nos besoins et à nos possibilités d'action, donnant ainsi naissance aux émotions.

Cet essai s'appuie explicitement sur cette tradition. Il n'entend ni la remplacer, ni proposer une théorie alternative de l'évaluation.

Au contraire, il part du principe que ces recherches ont clarifié un point essentiel : les émotions ne dépendent pas directement des événements, mais de la signification que l'esprit leur attribue à travers un processus d'évaluation.

Le niveau dont nous nous occupons ici est toutefois plus large.

L'évaluation a surtout été étudiée comme un processus par lequel une appréciation produit une réponse émotionnelle. Dans ces pages, l'évaluation est au contraire considérée comme une fonction cognitive générale, dont l'évaluation émotionnelle représente l'un des cas les plus importants, mais pas le seul.

Nous pouvons juger un argument convaincant sans éprouver d'émotion intense. Nous pouvons considérer une source comme fiable, estimer qu'un projet est prometteur ou juger une procédure peu efficace. Dans tous ces cas, l'esprit attribue une valeur, même lorsque la composante émotionnelle est minime ou difficilement perceptible.

Le niveau de base du modèle concerne donc toute forme d'attribution de valeur, qu'elle suscite ou non une émotion manifeste.

Ce choix est également important pour une autre raison, qu'il convient d'énoncer ouvertement.

Si nous identifions l'évaluation à la seule « appraisal » émotionnelle, le modèle proposé dans ces pages coïnciderait avec une théorie déjà existante, et sa contribution se réduirait à une terminologie différente.

Le point de départ est en revanche plus général : l'esprit évalue en permanence le monde en fonction de ce qu'il considère comme pertinent pour ses propres objectifs, et seule une partie de ces évaluations a été étudiée par la psychologie des émotions sous le nom d'« appraisal ».

Qu'est-ce qui détermine la valeur ?

En première approximation, la valeur est déterminée par nos besoins et nos désirs.

Ces concepts ne seront pas redéfinis ici. Ils ont été abordés en détail dans mon essai *Psychologie des besoins*, auquel le présent ouvrage renvoie. Il suffit de considérer que les besoins et les désirs constituent les critères à travers lesquels l'esprit attribue une valeur aux événements et oriente le comportement.

Une même situation peut donc faire l'objet d'évaluations profondément différentes selon les personnes, non pas parce que l'une perçoit correctement la réalité et l'autre non, mais parce qu'elles attribuent de la valeur à des choses différentes.

Pour celui qui a faim, un sandwich a une grande valeur ; pour celui qui vient de terminer un copieux déjeuner, ce même sandwich peut sembler insignifiant, voire repoussant.

Pour celui qui aspire à la tranquillité, une fête bruyante représente un inconvénient ; pour celui qui recherche des occasions de rencontre ou de divertissement, cette même fête représente une opportunité.

Les appréciations changent parce que les critères selon lesquels la valeur est attribuée changent.

Évaluer et juger

Comme annoncé dans l'introduction, ces deux termes ne sont pas synonymes. Il est temps de préciser cette distinction, qui sera maintenue tout au long de cet essai.

Par « *évaluation* », nous désignons l'attribution de valeur décrite jusqu'à présent : cette chose, pour moi, en ce moment, vaut beaucoup ou peu. Il s'agit d'une opération ouverte. Elle admet des degrés, elle peut être suspendue, elle peut être corrigée. « Je suis encore en train d'évaluer » est une phrase qui a du sens.

Par « *jugement* », nous désignons l'évaluation qui devient un verdict et qui influence la relation avec ce qui a été jugé. Il s'agit donc de deux éléments, qu'il convient d'aborder l'un après l'autre.

Le premier est le verdict. Un jugement n'est pas un point sur une échelle graduée : c'est une qualification. C'est bon, c'est injuste, c'est médiocre, c'est peu fiable.

La différence réside dans la manière dont nous utilisons ces deux mots. Nous pouvons évaluer un bien immobilier à deux cent mille euros, ou un devoir à huit sur dix. Nous ne pouvons pas juger un huit sur dix.

Et nous pouvons dire « je suis encore en train d'évaluer », alors que nous ne pouvons pas dire « je suis encore en train de juger ». Le jugement est un point d'arrivée, une décision finale, pas un parcours.

Le deuxième élément est la conséquence relationnelle, et c'est ce qui distingue *concrètement* le jugement de l'évaluation. Celui qui juge une personne ne se contente pas de constater sa valeur : il adopte, à son égard, une attitude plus ou moins favorable. Le verdict influence ce que nous ferons — si nous lui ferons confiance, si nous la rechercherons, si nous l'éviterons, si nous la choisirons lorsque nous aurons la possibilité de choisir. C'est la raison pour laquelle se sentir jugé pèse plus lourd que se sentir évalué. Ce n'est pas le fait que quelqu'un se soit fait une opinion sur nous qui est désagréable : c'est que cette opinion orientera son comportement à notre égard.

Et quand l'objet n'est pas une personne ? Juger une œuvre, une loi, une décision a également une conséquence relationnelle, mais celle-ci retombe sur celui qui juge. Qualifier une œuvre de médiocre revient à s'exposer, car ce jugement en dit long également sur celui qui l'émet. Nous verrons plus loin à quel point cela pèse.

La différence entre les deux termes ne réside donc pas dans l'objet — tous deux peuvent avoir pour objet des choses ou des personnes — mais dans leurs conséquences respectives.

Dans le langage courant, les deux termes se recoupent sans cesse, et rien n'empêche de dire qu'on juge un vin ou qu'on évalue un collègue. Ici, en revanche, les distinguer permettra de séparer l'opération générale du cas particulier qui nous intéresse.

Cette distinction n'explique toutefois pas encore pourquoi les évaluations sociales se comportent de manière si différente des autres. La propriété véritablement décisive est une autre, et c'est le chapitre suivant qui l'introduira.

Avec les personnes, l'évaluation devient un verdict

Jusqu'à présent, l'évaluation a été décrite comme une opération ouverte : une estimation qui peut être revue, nuancée, corrigée. C'est ainsi que cela fonctionne avec les objets. Avec les personnes, il en va autrement.

En observant le comportement de quelqu'un, nous en déduisons automatiquement des traits de caractère stables — il est désordonné, il est généreux, il n'est pas fiable — sans l'avoir décidé, sans

but explicite et sans nous rendre compte du processus. Ce phénomène s'appelle l'inférence spontanée de traits de caractère ; cette ligne de recherche remonte à James Uleman et ses collaborateurs, et quarante ans d'études menées avec différentes méthodes en ont confirmé l'existence.

C'est un fait qui mérite notre attention, car il décrit le passage de l'estimation au verdict. L'estimation d'un poids reste une estimation : elle se situe sur une échelle et peut être révisée. L'inférence concernant une personne devient un prédicat — « c'est comme ça » — et les prédicats n'ont pas de degrés.

Deux aspects rendent ce phénomène plus répandu qu'il n'y paraît.

Le premier est que l'inférence se forme même sur la base d'un nombre très restreint d'éléments : il n'est pas nécessaire de connaître quelqu'un, un comportement observé une seule fois suffit. C'est la raison pour laquelle « je n'ai pas d'éléments pour juger » est presque toujours faux. Cela ne décrit pas l'absence de jugement, mais un jugement formé sur peu d'éléments — et le fait que celui qui le dit soit de bonne foi fait partie du phénomène, au lieu de le contredire.

Le second concerne le résultat qui ressemble à une absence de jugement. Etablir qu'une personne est sans importance, qu'elle ne mérite pas qu'on lui prête attention, qu'on peut l'ignorer sans conséquences, n'est pas un manque de jugement : c'est un jugement, et l'un des plus lourds de conséquences, car il établit que l'autre n'a aucun pouvoir sur nous. Même décider qu'il n'est pas utile de recueillir des informations sur une personne présuppose une estimation de sa valeur attendue.

On ne peut s'empêcher de juger

Il découle de ce qui précède un raisonnement en trois étapes.

Premièrement : nous évaluons sans cesse des choses et des personnes, et nous ne pouvons pas nous en passer. C'est la fonction décrite au début de ce chapitre.

Deuxièmement : lorsque l'objet est une personne, l'évaluation a tendance à ne pas rester ouverte. Elle devient une attribution dispositionnelle. Cela peut se produire délibérément, mais cela se produit aussi sans intention et sans que l'on en ait conscience.

Le troisième : puisque, en société, nous sommes entourés de personnes qui nous perçoivent, et pour lesquelles les mêmes prémisses s'appliquent, il s'ensuit que

on ne peut s'empêcher de juger, ni d'être jugé.

L'affirmation est forte, et il convient de préciser d'emblée ce qui la contredirait. Si les inférences sur les personnes exigeaient une intention, ou si elles restaient situationnelles et modulables au lieu de se consolider en dispositions, elle serait fausse. Les recherches disponibles indiquent le contraire.

Ce que cela n'autorise pas

Une conclusion de ce type peut facilement être interprétée de deux manières erronées, opposées l'une à l'autre.

La première est le moralisme qu'elle réfute. Il existe une tradition, morale avant d'être psychologique, selon laquelle *juger* serait en soi un défaut — enfermer, étiqueter, condamner — et *ne pas juger* serait une vertu accessible par l'exercice. Si l'argument que je viens d'exposer est correct, cette vertu n'est pas accessible : le verdict se forme avant qu'une décision ne soit possible, et se proposer de ne pas le former revient à se proposer de ne pas reconnaître un visage familier.

La deuxième erreur est symétrique, et il convient de l'écarter avec la même netteté. Que le jugement se forme de toute façon ne signifie pas qu'on puisse l'exprimer librement. Former un verdict et le déclarer sont deux actes distincts : le premier est automatique, le second est un geste qui a ses propres conséquences, et un chapitre suivant montrera que celui qui l'accomplit s'expose. Affirmer qu'on ne peut s'empêcher de juger n'est donc pas un blanc-seing pour juger à haute voix ; et l'ancienne injonction de ne pas juger, interprétée comme concernant le fait de prononcer et de condamner plutôt que de penser, reste parfaitement sensée.

Cet essai décrit un mécanisme. Il ne recommande ni de s'y plier, ni de le combattre.

Vers le chapitre suivant

Jusqu'à présent, nous n'avons parlé que de la fonction fondamentale de l'évaluation.

L'introduction a déjà laissé entrevoir que le véritable objet de cet essai est un cas particulier de ce processus : celui où une évaluation ne se limite pas à décrire son objet, mais peut devenir elle-même l'objet d'évaluations supplémentaires.

L'évaluation sociale n'est pas une faculté différente de l'évaluation ordinaire. Il s'agit de la même opération cognitive insérée dans un contexte social.

Lorsque nous attribuons une valeur à quelque chose par nous-mêmes, l'évaluation s'arrête à l'objet.

Lorsque nous l'exprimons en présence d'autrui, ou que nous savons qu'elle pourra être observée, la situation change. À partir de ce moment, ce n'est plus seulement l'objet de notre jugement qui est évalué : le fait même de l'avoir évalué devient une information pour les autres.

C'est à partir de cette transformation de la structure de l'évaluation que s'articule le modèle développé dans les chapitres suivants.

2 Quand l'évaluation devient réursive

Imaginons que nous dégustions un vin seuls, dans la cuisine, à la fin d'une journée. Nous le trouvons bon, médiocre ou excellent. L'évaluation naît, s'accomplit et s'épuise là. Personne ne la connaîtra jamais.

Imaginons maintenant ce même vin à table, avec cinq personnes, dont l'une l'a choisi.

Quelque chose a changé, et ce n'est pas le vin.

Ce qui a changé, c'est que notre évaluation est devenue observable. Et dès lors qu'elle devient observable, elle cesse de ne concerner que le vin : elle devient à son tour quelque chose que les autres peuvent juger. Si nous disons qu'il est excellent, certains jugeront que nous nous y connaissons en vins, ou que nous faisons preuve de complaisance envers celui qui l'a choisi. Si nous disons qu'il est médiocre, quelqu'un jugera que nous avons un palais raffiné, ou que nous sommes impolis, ou encore que nous prenons de grands airs.

Le problème n'est pas que nous nous préoccupions de ces possibilités. Souvent, nous ne nous en rendons même pas compte. Le problème est que la structure de la situation a changé : l'évaluation est entrée dans un circuit où elle est elle-même l'objet d'une évaluation, et souvent d'un jugement.

Nous qualifions de réursive une évaluation qui se trouve dans cette situation.

C'est là la propriété que le chapitre précédent avait laissée en suspens. Ce qui distingue les évaluations sociales des autres, ce n'est pas leur objet — on peut évaluer des personnes sans que rien de tel ne se produise, et évaluer des objets dans des situations où cela se produit — et ce n'est pas non plus le type de critère invoqué. C'est leur structure systémique.

Ce que la réursivité exclut

Une propriété définitoire n'est utile que si elle exclut quelque chose. Il convient donc de préciser d'emblée quelles évaluations ne relèvent pas du champ d'application de cet essai.

N'y figurent pas les évaluations que personne ne peut connaître et dont rien ne dépend : le jugement silencieux porté sur un paysage traversé en train, la préférence jamais exprimée entre deux nuances de gris, l'estimation instinctive de la température de l'eau.

N'en font pas partie les réactions immédiates à des stimuli physiques — douleur, faim, alerte — qui ne nécessitent aucune anticipation relationnelle.

Et n'y figurent pas non plus, si ce n'est de manière marginale, les appréciations formulées dans des conditions d'anonymat réel et définitif.

Sur ce dernier point, une précision s'impose toutefois, qui reviendra à plusieurs reprises dans cet essai. La frontière n'est pas nette, car l'esprit ne semble pas bien reconnaître l'anonymat. Dans les jeux économiques menés en double aveugle, où personne ne peut remonter jusqu'à celui qui a pris telle ou telle décision, la générosité des participants s'effondre — ce qui confirme à quel point la possibilité d'être observé pèse — mais elle ne tombe pas à zéro. Quelque chose continue de fonctionner comme si quelqu'un observait.

Le champ d'application de cet essai est donc plus restreint que celui des évaluations en général, mais ses contours sont flous. Cette imprécision n'est pas un défaut de la description : c'est une propriété du phénomène, et dans le chapitre consacré à l'observateur intériorisé, elle deviendra elle-même objet d'analyse.

Une structure déjà décrite ailleurs

La récursivité n'est pas une idée nouvelle. Sa formulation la plus célèbre remonte à John Maynard Keynes, qui l'a illustrée en 1936 à l'aide de l'image d'un concours de beauté organisé par un journal. Les lecteurs doivent désigner les six plus beaux visages parmi une centaine de photographies, et le gagnant est celui qui se rapproche le plus du choix moyen des participants. Le concurrent avisé, observait Keynes, ne choisit pas les visages qu'il trouve les plus beaux : il choisit ceux qu'il prévoit que les autres choisiront. Et comme les autres raisonnent eux aussi ainsi, le choix de chacun dépend des prévisions que chacun fait sur les prévisions des autres. Keynes décrivait la Bourse, mais il décrivait en réalité une structure bien plus générale.

Cette même structure se retrouve ailleurs sous d'autres noms. En sociologie, Niklas Luhmann l'a appelée « double contingence » : chacun adapte son comportement aux attentes qu'il attribue à l'autre, qui fait de même. En psychiatrie, Ronald Laing en a décrit les spirales — « je sais que tu sais que je sais » — comme une réalité clinique quotidienne. En théorie des jeux, cela correspond aux croyances d'ordre supérieur. Et la tradition qui va de George Herbert Mead à Erving Goffman a fondé sur ce principe une grande partie de sa compréhension de la vie sociale.

Cet essai ne prétend donc pas découvrir la récursivité. Comme déjà indiqué dans l'introduction, l'architecture globale décrite ici existe sous forme clinique dans le modèle de Clark et Wells et sous forme computationnelle dans les modèles de traitement prédictif. Ce que propose cet essai, c'est de la considérer comme un fonctionnement ordinaire plutôt que comme un mécanisme pathologique, et de l'étendre systématiquement au domaine du jugement, que ces traditions ne traitent pas.

Quelle est la profondeur de la récursivité ?

Décrite avec des mots, la récursivité peut évoquer l'image d'une spirale vertigineuse : je pense à ce que tu penses de ce que je pense de ce que tu penses, et ainsi de suite à l'infini.

L'image est évocatrice, mais elle est fausse. Le jeu du concours de beauté a été reproduit en laboratoire sous une forme simplifiée : on demande aux participants de choisir un nombre compris entre zéro et cent, et celui qui se rapproche le plus d'une fraction de la moyenne de tous les choix remporte la victoire. La solution logique du jeu nécessiterait un nombre indéfini d'itérations. Les personnes réelles, en revanche, s'arrêtent presque toujours après un ou deux tours. Rosemarie Nagel a montré que le comportement observé s'explique bien par un modèle à plusieurs niveaux : la plupart des joueurs réfléchissent à ce que feront les autres, certains à ce que les autres prévoient qu'ils feront, et presque personne ne va plus loin.

La récursivité réelle, en somme, est faible. Ce n'est pas une spirale : il s'agit d'un ou deux échelons, parcourus rapidement et presque jamais de manière consciente.

De ce constat découle une hypothèse avancée par l'essai et qui n'a pas encore été vérifiée : la profondeur de la récursivité ne serait pas la même pour tout le monde, et augmenterait à mesure que l'anxiété sociale s'accroît. Ceux qui redoutent fortement le jugement des autres devraient, selon ce

modèle, gravir davantage d'échelons — non seulement se demander s'ils auront l'air mal à l'aise, mais aussi si l'on verra qu'ils essaient de ne pas l'être.

La conséquence principale

Si une évaluation peut être évaluée, alors l'exprimer n'est jamais un acte neutre.

Il s'ensuit que, dans les situations sociales, nos évaluations ne dépendent pas seulement de ce que nous évaluons. Elles dépendent aussi, dans une mesure variable, de ce que cette évaluation impliquera pour nous.

Il convient de bien nuancer cette affirmation, car dans sa forme absolue, elle serait fausse. Il n'est pas vrai que nous n'évaluons jamais les choses en elles-mêmes. Ce qui est vrai, c'est quelque chose de plus précis et de plus vérifiable : la distorsion augmente proportionnellement à l'observabilité de l'évaluation et à l'importance de ses conséquences. Un jugement exprimé devant quelqu'un qui peut décider de notre carrière est plus exposé qu'un jugement exprimé entre amis, qui est lui-même plus exposé qu'un jugement confié à un journal intime.

Et, comme on l'a vu, elle ne tombe pas tout à fait à zéro, même lorsque personne ne regarde.

Cette formulation nuancée présente un avantage par rapport à la version absolue : elle indique où l'effet est à prévoir et où il ne l'est pas, et peut donc être réfutée. Si, en manipulant expérimentalement l'observabilité et les conséquences d'un jugement, on n'observait pas de variations systématiques dans son expression, le modèle serait erroné.

Ce qui manque encore

La récursivité explique pourquoi les évaluations sociales sont conditionnées. Elle n'explique pas encore pourquoi elles sont si souvent erronées.

En effet, pour adapter une évaluation à la façon dont elle sera jugée, il faut disposer d'une estimation de la façon dont elle sera jugée. Et cette estimation doit bien venir de quelque part.

On serait naturellement tenté de supposer qu'elle provient de l'observation : on observe les autres, on recueille des indices, on corrige. C'est là que le modèle aboutit à son résultat le plus contre-intuitif. Car cette estimation, dans une large mesure, ne provient pas de l'observation des autres.

Elle vient de nous.

3 À quoi sert le jugement ?

Jusqu'à présent, nous avons décrit une propriété des évaluations sociales : leur récursivité. Une évaluation, lorsqu'elle devient observable, peut à son tour être évaluée. De cette structure découle la possibilité que nos évaluations soient modifiées en prévision de la façon dont elles seront accueillies.

Il reste cependant une question plus profonde.

Pourquoi le jugement des autres nous importe-t-il tant ?

La réponse la plus immédiate semblerait être : parce que nous recherchons l'approbation. Ou bien parce que nous voulons être appréciés, éviter le rejet, acquérir du prestige ou conserver une bonne réputation.

Toutes ces réponses contiennent une part de vérité, mais elles ne décrivent le phénomène qu'à un niveau encore superficiel. L'approbation, le prestige, le conformisme, la honte et la recherche de statut sont autant de manifestations d'un problème plus général. Pour les comprendre, il faut se demander quelle fonction remplit le jugement au sein des relations humaines.

La valeur sociale

Toute relation implique, implicitement ou explicitement, un choix.

Nous choisissons avec qui collaborer, à qui faire confiance, qui écouter, qui fréquenter, de qui apprendre, avec qui travailler, qui embaucher, qui inviter, qui aider, qui éviter.

Le jugement contribue à orienter ces choix.

Lorsque nous évaluons une personne, nous ne nous contentons en effet pas de la décrire : nous produisons une information susceptible d'influencer la manière dont les autres entreront en relation avec elle. Être jugé favorablement signifie augmenter la probabilité d'être choisi ; être jugé défavorablement signifie la réduire.

Le jugement d'autrui ne constitue donc pas une fin en soi. C'est un indicateur dont dépend notre valeur sociale : à quel point nous sommes désirables en tant que partenaire d'interaction.

Par « partenaire d'interaction », on n'entend pas uniquement un partenaire affectif. Le même principe s'applique à un ami, un collègue, un collaborateur, un salarié, un employeur, un allié ou un modèle à imiter. Les situations changent, mais ce qui reste constant, c'est le fait que quelqu'un puisse nous préférer à quelqu'un d'autre.

Il s'agit de la même opération que celle décrite dans le premier chapitre, mais avec les rôles inversés. Là, la valeur était ce qu'une chose vaut pour celui qui l'évalue, en fonction de ses besoins et de ses désirs ; ici, l'objet évalué, c'est nous, et le critère réside dans les besoins de celui qui nous évalue.

Il s'ensuit que la valeur d'une personne n'est pas une quantité unique, et qu'elle ne lui appartient pas. Elle est relative à celui qui évalue et à ce dont il a besoin : on peut avoir beaucoup de valeur

pour quelqu'un et peu pour quelqu'un d'autre sans que l'un ou l'autre ne se trompe, et sans que la personne ait changé.

Or, le système ne le présente pas ainsi. Le jugement présumé se présente comme un verdict unique et global — « je ne vaudrais pas grand-chose » — et non comme l'estimation d'un évaluateur particulier en fonction de ce dont il a besoin.

Et nous ne sommes pas seulement évalués. Nous évaluons à notre tour, et cela entre également en ligne de compte : la valeur de nos jugements, le fait de savoir si nous sommes des personnes dont il convient de solliciter l'avis ou auprès desquelles il convient d'apprendre, fait à son tour l'objet d'une évaluation par autrui. Celui qui juge ne se tient pas en dehors du marché sur lequel il est jugé — c'est la raison pour laquelle la deuxième partie de cet essai devra également aborder le volet actif.

Être choisi

De nombreux phénomènes qui semblent distincts se ramènent à ce même mécanisme. La coopération, en effet, consiste à choisir quelqu'un avec qui faire quelque chose ensemble ; l'amitié, à se choisir mutuellement comme interlocuteurs privilégiés ; la sélection du personnel, à préférer un candidat à un autre.

Le choix d'un partenaire affectif, où la réciprocité est déclarée, est encore plus évident. Sous des noms différents, c'est donc toujours la même opération qui se répète : quelqu'un choisit, et quelqu'un est choisi.

La compétition ne fait pas non plus exception. Il s'agit simplement du cas où plusieurs personnes cherchent à être choisies pour une ressource limitée. On ne rivalise pas contre quelqu'un pour le simple plaisir de rivaliser : on rivalise pour être préféré.

Être choisi est donc la principale voie par laquelle le jugement produit des conséquences. Ce n'est pas la seule : celui qui nous juge peut aussi nous aider ou nous mettre des bâtons dans les roues, nous récompenser ou nous punir, parler de nous à d'autres. Mais c'est celle qui revient le plus souvent, et c'est autour d'elle que le système s'est vraisemblablement constitué.

L'approbation, le prestige et la réputation acquièrent ainsi une signification plus précise. Ce ne sont pas des objectifs finaux, mais des moyens par lesquels augmente ou diminue la probabilité que les autres nous choisissent.

Ce que le jugement rend possible

Avant d'aborder les limites, il convient de s'attarder sur ce que ce mécanisme produit lorsqu'il fonctionne — car il fonctionne normalement, et c'est cet aspect que l'on a tendance à oublier.

La coopération entre individus non apparentés est rare dans la nature et remarquable chez notre espèce. Elle exige que chacun puisse évaluer, avant de s'engager, dans quelle mesure l'autre est digne de confiance, et que chacun sache qu'il est à son tour estimé. Sans un système capable d'anticiper les conséquences sociales de ses propres actions, cette estime n'existerait pas.

De cette même capacité dépendent la confiance envers des personnes que nous ne connaissons pas, le respect des règles quand personne ne surveille, l'apprentissage social — qui suppose de

reconnaître ceux qu'il convient d'imiter — et la réputation en tant qu'institution, c'est-à-dire en tant qu'information qui circule et survit à celui qui l'a produite.

L'autocensure dont nous parlerons en tant que problème est cette même capacité qui nous empêche de dire, lors d'un enterrement, ce que l'on pense du défunt.

Il en découle une conséquence qu'il vaut la peine d'énoncer, car elle va à l'encontre d'une habitude répandue.

Condamner le fait de juger en bloc revient à condamner le mécanisme qui rend la coexistence possible. Il ne s'agit pas d'une défense morale du jugement : c'est le constat que ce qui, dans le langage courant, est déploré comme un « jugement » — classer les personnes, prévoir comment elles se comporteront, s'adapter en conséquence — est précisément l'opération qui permet de faire confiance à quelqu'un, de choisir un collaborateur, de reconnaître qui mérite d'être écouté.

Les phénomènes que les chapitres suivants traiteront comme des problèmes ne sont donc pas le fonctionnement de ce système. Ce sont ses réglages erronés.

Quel est le poids d'un jugement ?

Cette conclusion serait toutefois encore trop générale. Si le jugement ne servait qu'à déterminer la valeur sociale, tout jugement devrait avoir plus ou moins la même importance — et l'expérience montre que ce n'est pas le cas : un même commentaire peut nous laisser totalement indifférents ou nous préoccuper pendant des jours, selon la personne qui l'exprime. La différence ne réside pas dans le fait que cette personne nous plaise ou nous déplaise, mais dans son pouvoir d'influencer quelque chose qui nous concerne.

Un jugement émis par un inconnu que l'on n'a rencontré qu'une seule fois a, en règle générale, beaucoup moins de poids que celui émis par un collègue avec lequel on travaille tous les jours, par un supérieur qui décide d'une promotion ou par une personne dont dépend une relation importante.

Ce qui rend un jugement pertinent, ce n'est donc pas sa simple existence, mais le pouvoir que celui qui le formule exerce sur nous : la mesure dans laquelle ce que cette personne fera, ou ne fera pas, peut changer quelque chose dans notre vie — nous choisir ou nous écarter, nous aider ou nous mettre des bâtons dans les roues, nous récompenser ou nous punir, parler de nous aux autres de manière positive ou négative.

Ce pouvoir ne crée pas l'effet du jugement : il en détermine l'intensité. Un jugement en produit néanmoins un, même lorsqu'il émane de quelqu'un qui ne peut rien changer à notre vie — mais son poids dépend justement de cela.

La motivation à gérer le jugement d'autrui grandit à mesure que grandit le pouvoir de l'autre sur nous ou sur quelque chose qui a de la valeur à nos yeux.

Pourquoi le pouvoir qu'une personne exerce sur nous en dit plus long que la probabilité de se rencontrer

Dans l'un de mes travaux précédents, **Le jugement présumé**, j'avais formulé cette variable différemment, en accordant une importance majeure à la probabilité d'interactions futures.

L'idée était intuitive : plus une relation a de chances de se poursuivre, plus l'intérêt pour le jugement de l'autre sera grand.

Cette formulation, cependant, n'explique pas tous les cas.

Prenons l'exemple d'une relation profondément asymétrique, comme celle entre un maître et un esclave.

Les interactions futures sont pratiquement certaines et très fréquentes. Pourtant, le maître n'a, en général, aucune raison de soigner sa réputation auprès de l'esclave. Il a en revanche intérêt à la préserver auprès des autres maîtres, dont dépendent les alliances, le prestige et la coopération.

La fréquence des interactions ne suffit donc pas à expliquer l'importance du jugement.

Le pouvoir qu'ils exercent sur nous, lui, suffit.

Une question à garder à l'esprit

Avant de poursuivre, il convient de dégager une conséquence de ce qui précède. Cela ne sert pas encore à modifier le comportement : cela sert à rendre observable une variable du modèle.

Si le poids d'un jugement dépend du pouvoir que celui qui le formule exerce sur nous, alors face à un jugement qui nous trouble, une question vérifiable se pose :

Cette personne peut-elle changer quelque chose dans ma vie ?

Ce n'est pas une question consolatrice, et elle ne demande pas de minimiser. Elle demande une vérification. Dans de nombreuses situations, la réponse est oui, et alors l'inquiétude est justifiée. Dans bien d'autres — le commentaire d'un inconnu, l'avis de quelqu'un que nous ne reverrons pas, l'opinion de quelqu'un qui n'a aucune influence sur notre vie —, la réponse est non, et le trouble persiste malgré tout, car le système le produit inconsciemment et automatiquement.

Quant à savoir pourquoi le système ne procède pas à cette vérification, c'est une question que nous ne pourrions aborder que plus tard.

Vers le jugement présumé

Nous disposons désormais de tous les éléments pour aborder le problème suivant. Nos évaluations sont modifiées parce que le jugement des autres peut influencer nos chances d'être choisis et/ou d'influencer notre vie ; mais pour adapter notre comportement, il faut d'abord estimer quel sera ce jugement. D'où provient cette estimation ? L'hypothèse la plus intuitive est qu'elle se construit en observant attentivement les autres et en recueillant des informations sur la manière dont ils nous perçoivent.

Le chapitre suivant montrera que, dans la plupart des cas, le processus fonctionne différemment. La représentation du jugement d'autrui ne coïncide ni avec le jugement réel ni avec l'auto-évaluation : elle constitue une troisième variable, le jugement présumé, et c'est celle-ci qui régit une grande partie de notre comportement.

4 Le jugement présumé

Dans les chapitres précédents, nous avons vu pourquoi le jugement des autres peut modifier notre comportement.

Une évaluation observable peut elle-même être évaluée, et son poids dépend du pouvoir que l'observateur exerce sur nous. Si le jugement des autres peut influencer notre vie, il devient intéressant d'essayer de le prévoir.

C'est là qu'apparaît le concept central de cet essai.

Pour nous orienter dans les relations sociales, nous ne réagissons pas directement au jugement réel des autres, que nous ignorons dans la plupart des cas. Nous réagissons à une représentation mentale de celui-ci.

Nous appellerons cette représentation « jugement présumé ».

Ce concept n'est pas nouveau dans cet essai : je l'avais introduit dans un essai précédent, **Le jugement présumé**, en le définissant comme la représentation anticipée — généralement implicite et inconsciente — du jugement que les autres pourraient porter sur notre comportement, actuel ou futur. Ces pages restent la formulation de référence ; ici, le concept est repris, replacé dans un cadre plus large et, sur un point, corrigé.

Il convient de préciser la définition initiale, car sans cela, le modèle semble beaucoup plus restrictif qu'il ne l'est en réalité. Le terme « jugement » ne désigne pas ici une opinion exprimée verbalement. Il désigne l'ensemble de la réaction sociale attendue : approbation ou désapprobation, admiration ou déception, acceptation ou rejet, hostilité ou bienveillance, disposition à coopérer ou perte de confiance. La plupart des jugements qui nous concernent ne sont jamais formulés, mais n'en ont pas moins un effet.

Une troisième variable

Il est facile de penser qu'il n'existe que deux possibilités : le jugement réel, c'est-à-dire ce que les autres pensent effectivement de nous, et l'auto-évaluation, c'est-à-dire ce que nous pensons de nous-mêmes. Dans le comportement quotidien, cependant, une troisième variable intervient.

Nous pouvons nous considérer comme compétents tout en étant convaincus que les autres nous sous-estiment, ou avoir une mauvaise opinion de nous-mêmes tout en croyant être très estimés. Nous pouvons être réellement appréciés sans nous en douter, ou nous sentir admirés alors que les autres nous jugent avec froideur.

Aucune de ces situations ne peut être décrite uniquement par le jugement réel et l'auto-évaluation. Il faut une troisième représentation : la manière dont nous croyons être évalués. Et c'est celle-ci qui oriente une grande partie de nos décisions.

Le jugement qui guide le comportement

Imaginons que nous devions prendre la parole lors d'une réunion. Si nous étions certains que les autres apprécieront ce que nous dirons, nous parlerions avec aisance ; si nous étions certains qu'ils nous jugeront incompétents, nous resterions probablement silencieux. Dans la quasi-totalité des cas,

cependant, nous ne disposons d'aucune de ces deux certitudes : nous ne disposons que d'une prévision. Et c'est cette prévision, et non le jugement réel, qui détermine si nous devons intervenir ou non.

Il en va de même lorsque nous décidons comment nous habiller, dans quelle mesure nous devons nous montrer compétents, s'il faut raconter une expérience personnelle, contredire quelqu'un, demander de l'aide ou exprimer une opinion impopulaire.

Le comportement n'est pas régi par le jugement des autres, mais par le jugement que nous attribuons aux autres. Le système réagit à la représentation, et non directement à la réalité.

Pourquoi le jugement réel ne suffit pas

Cela peut sembler évident, mais cela a une conséquence importante. Le jugement réel, dans la plupart des cas, nous est tout simplement inconnu : les autres nous disent rarement avec précision ce qu'ils pensent de nous, et lorsqu'ils le font, ils peuvent être sincères ou non, ne connaître qu'une partie de nous, ou changer d'avis avec le temps. L'esprit ne peut donc pas attendre une information qui n'arrive presque jamais. Pour fonctionner, il doit en construire une.

Le jugement présumé naît précisément de cette nécessité. C'est une réponse au problème de l'incertitude : sans estimation de la façon dont nous serons probablement évalués, toute décision sociale exigerait une information que nous ne possédons pas.

Quant à savoir s'il s'agit d'une bonne réponse, c'est une autre question, qu'il convient de distinguer. Un système peut être à la fois indispensable et mal calibré. C'est précisément ce dont traite le chapitre suivant.

Le jugement présumé ne coïncide pas avec le jugement réel

Le jugement présumé doit être distingué du jugement réel. Parfois, les deux coïncident, d'autres fois ils divergent profondément : une personne peut être très appréciée et se sentir constamment inadéquate, une autre peut susciter de l'irritation sans s'en rendre compte. Toutes deux adaptent leur comportement en fonction du jugement présumé, et non du jugement réel.

Il s'ensuit que modifier le jugement des autres ne suffit pas à modifier le comportement. Si le jugement réel change mais que le jugement présumé reste inchangé, le comportement a tendance à rester le même — une conséquence qui deviendra décisive par la suite.

Une variable observable

Le jugement présumé n'est pas une entité théorique introduite pour expliquer n'importe quel phénomène : il peut être observé indirectement. Nous pouvons demander à une personne comment elle pense être perçue, comparer sa réponse au jugement exprimé par les autres, et comparer ces deux évaluations à l'auto-évaluation de la personne elle-même.

Ces trois grandeurs ne coïncident que partiellement, et c'est leur divergence qui fait du jugement présumé une variable distincte : s'il coïncidait toujours avec l'auto-évaluation, il serait inutile de le distinguer ; s'il coïncidait toujours avec le jugement réel, il suffirait de mesurer ce dernier.

La recherche sur les métaperceptions a mesuré ces trois grandeurs et montre qu'elles ne se recoupent que partiellement. Les résultats seront examinés dans le chapitre suivant, car c'est de la manière dont elles se recoupent que dépend le point le plus contre-intuitif du modèle.

Les critères selon lesquels nous jugeons les autres

Jusqu'à présent, nous avons considéré le jugement présumé comme ce qui régit notre comportement : ce que nous disons, ce que nous montrons, ce que nous taisons.

Mais le deuxième chapitre a établi qu'une évaluation, lorsqu'elle est observable, peut à son tour être évaluée. Et nos évaluations comptent parmi les comportements les plus observables que nous produisons.

Il en découle une conséquence que l'essai reprendra dans la deuxième partie.

Si nous prévoyons d'être jugés également pour la manière dont nous jugeons, alors le jugement présumé ne régit pas seulement ce que nous faisons : il contribue à former les critères mêmes avec lesquels nous évaluons les choses et les personnes.

Un exemple permettra de concrétiser cette idée. Celui qui s'attend à ce que l'admiration pour un certain auteur soit interprétée comme un signe de naïveté aura tendance à trouver cet auteur moins convaincant qu'il ne le trouverait si cette anticipation n'intervenait pas — non pas parce qu'il a décidé de feindre, mais parce que l'anticipation précède la formation du jugement.

Le jugement présumé agit donc dans deux sens : sur la manière dont nous nous comportons face à ceux qui nous évaluent, et sur la manière dont nous évaluons à notre tour. Ce n'est pas seulement un mécanisme d'adaptation, qui régit ce que nous montrons : c'est aussi un mécanisme de construction, qui intervient dans la formation des évaluations elles-mêmes. La deuxième partie de cet essai accordera à cette seconde fonction la même attention qu'à la première.

Le lien vers le chapitre suivant

Nous avons désormais identifié la variable qui sert de médiateur entre le jugement des autres et notre comportement.

Il reste cependant une question décisive. Comment ce jugement présumé se construit-il ?

La réponse la plus intuitive est qu'il découle avant tout de l'observation des autres.

Nous observons leurs réactions, nous recueillons des informations, nous corrigeons progressivement notre représentation.

C'est une explication plausible, et c'est celle que le bon sens suggère spontanément.

Le chapitre suivant montrera que les choses semblent se passer autrement.

La représentation que nous nous faisons de la façon dont nous pensons être jugés provient, dans une large mesure, non pas de l'observation des autres, mais d'une source bien plus proche : la façon dont nous nous percevons nous-mêmes.

C'est là le résultat le plus contre-intuitif de tout ce modèle.

5 Être évalué : d'où vient le jugement présumé ?

Dans le chapitre précédent, nous avons vu que le comportement n'est pas guidé par le jugement réel des autres, mais par le jugement présumé : la représentation de la façon dont nous pensons être évalués.

Une question décisive demeure toutefois.

D'où vient cette représentation ?

L'intuition la plus spontanée est qu'elle découle principalement de l'observation des autres. Nous les observons, nous interprétons leurs réactions, nous recueillons des confirmations et des démentis et, petit à petit, nous construisons une image assez fidèle de la façon dont ils nous perçoivent.

La recherche suggère cependant une réponse différente.

Le jugement présumé découle, dans une large mesure, non pas du jugement réel des autres, mais de l'image que nous avons de nous-mêmes.

Une relation de cause à effet inattendue

On a longtemps pensé que le rapport était inverse : si les autres nous traitent avec estime, nous finissons par nous estimer nous-mêmes ; s'ils nous critiquent, nous finissons par nous dévaloriser. Ce processus existe bel et bien et est largement documenté, notamment dans le développement personnel et les relations importantes. Mais lorsqu'un adulte doit évaluer la façon dont il sera jugé dans une situation concrète, la tendance semble être inverse : il s'évalue d'abord lui-même, puis utilise cette évaluation pour déduire comment les autres le percevront probablement.

Ce n'est pas qu'elle cesse d'observer les autres : elle les observe bel et bien. Mais ce qu'elle en retire est en grande partie déterminé à l'avance, et un visage ambigu sera interprété dans le sens auquel la personne s'attendait déjà.

Le jugement des autres, plus qu'il n'est interprété, est présumé

La recherche sur les métaperceptions compare trois dimensions distinctes : la façon dont une personne s'évalue elle-même, la façon dont elle pense être évaluée par les autres, et la façon dont les autres l'évaluent réellement.

Dans la synthèse qui a donné naissance à ce domaine d'études, David Kenny et Bella DePaulo (1993) sont parvenus à une conclusion qui semblait à l'époque contre-intuitive : les personnes déduisent la façon dont les autres les perçoivent non pas tant à partir des retours qu'elles reçoivent, mais plutôt à partir de leur propre perception d'elles-mêmes. Kenny fait état, entre la façon dont une personne se perçoit et celle dont elle croit être perçue, d'une corrélation moyenne d'environ 0,87 — une valeur très élevée, qui indique un quasi-chevauchement entre les deux représentations.

La comparaison avec le jugement réellement exprimé par les autres révèle en revanche des correspondances bien plus faibles.

En d'autres termes, lorsque nous essayons d'imaginer comment nous sommes perçus, nous nous servons avant tout de nous-mêmes comme source d'information.

Cela ne signifie pas que nous ignorons complètement ce que les autres font ou disent. Cela signifie que ces informations ont moins de poids que ne le laisse croire l'introspection.

Le général et le particulier

Cette même tradition de recherche a mis en évidence une asymétrie qui mérite notre attention, car elle explique un phénomène que l'expérience quotidienne rend difficile à cerner.

Savoir comment « le monde en général » nous perçoit — si nous sommes perçus comme extravertis, compétents, fiables — est quelque chose qui nous réussit assez bien.

Savoir comment une personne en particulier nous perçoit, et en quoi sa perception diffère de celle des autres, est beaucoup plus difficile.

La précision globale est donc acceptable ; celle au niveau de la dyade est faible.

Le système semble ainsi fonctionner dans l'abstrait et échouer dans le concret, c'est-à-dire précisément là où il serait utile. L'estimation que nous formulons sur la façon dont une personne donnée nous jugera, dans une situation donnée, est la moins fiable de toutes — et c'est celle sur laquelle nous décidons de parler, de nous taire ou de nous exposer.

L'écart de sympathie

Une manifestation particulièrement instructive de cet écart a été décrite par Erica Boothby et ses collègues (2018), qui l'ont baptisée « liking gap ».

Après une conversation avec une personne qu'ils venaient de rencontrer, les participants avaient tendance à sous-estimer à quel point leur interlocuteur les avait appréciés. Il ne s'agissait pas d'une erreur ponctuelle, mais d'un écart systématique et unidirectionnel : les autres nous apprécient, en règle générale, plus que nous ne l'imaginons.

Ce phénomène s'est révélé solide et persistant, et apparaît dans différents contextes.

Si le jugement présumé correspondait à une interprétation du jugement d'autrui, un tel écart ne devrait pas exister, ou devrait se répartir de manière aléatoire dans les deux sens. Le fait qu'il soit orienté dans un seul sens est cohérent avec l'hypothèse selon laquelle cette représentation est construite ailleurs.

Les deux sources du jugement présumé

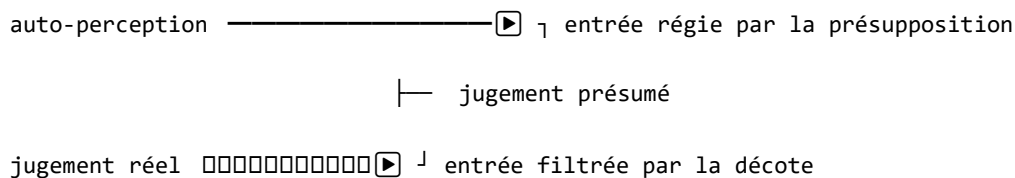
Les études sur ce qu'on appelle la « méta-intuition », menées notamment par Erika Carlson et ses collaborateurs, montrent que les personnes sont capables de distinguer de manière valable la façon dont elles se perçoivent de la façon dont elles sont perçues, et qu'elles n'atteignent pas un certain degré de précision simplement parce qu'elles projettent sur elles-mêmes l'opinion des autres. Elles puisent dans des sources d'information supplémentaires et fiables. Elles parviennent en outre à reconnaître, dans certaines limites, quand leur propre estimation est plus précise et quand elle l'est moins — une capacité qui s'avère liée à la qualité de la relation avec la personne en question.

Une revue systématique récente, réalisée par Harper et ses collègues, montre d'ailleurs que la méta-précision varie sensiblement en fonction du contexte relationnel, du trait considéré et de la méthode

de mesure utilisée. Les affirmations générales selon lesquelles « nous ne savons pas comment les autres nous perçoivent » ne sont donc pas tenables.

Le problème n'est donc pas que le jugement réel reste en dehors du système. C'est qu'il y entre avec un poids réduit et variable — et pour mesurer à quel point il est réduit, il convient d'examiner d'où proviennent les données utilisées par le système.

Le jugement présumé ne provient pas d'une seule source. Il reçoit des informations provenant de deux directions, avec des pondérations très différentes.



Les deux flèches n'ont pas la même épaisseur : le retour externe existe, mais pèse beaucoup moins lourd que l'autoperception. Les deux termes sont définis un peu plus loin.

Il convient à ce stade de changer de représentation. Jusqu'ici, le jugement présumé est apparu comme une variable parmi d'autres. Désormais, il est plus utile de le considérer comme un nœud : il reçoit des informations de deux sources, et c'est de lui que dépendent les deux classes de conséquences décrites dans le chapitre précédent, à savoir le comportement et les critères selon lesquels nous jugeons.

Ces deux sources sont l'auto-perception et le retour provenant des autres, et la construction du jugement présumé dépend du poids relatif attribué à chacune d'elles. Il s'ensuit que les différences individuelles peuvent avoir deux origines distinctes : une personne peut construire son jugement présumé presque exclusivement à partir de la façon dont elle se perçoit elle-même, ou bien elle peut également utiliser le retour social tout en lui attribuant un poids très faible. Ces deux cas produisent des comportements similaires mais ont des causes différentes.

Deux paramètres

Pour décrire ces deux processus, nous utiliserons deux paramètres.

Le premier s'appellera « présupposition » : il indique dans quelle mesure le jugement présumé est déjà établi avant même d'observer, c'est-à-dire déterminé par la perception de soi plutôt que par les informations disponibles dans la situation.

Cela ne concerne pas le fait de regarder ou non, mais la part de ce que l'on croit voir qui était déjà établie. Une présupposition élevée ne décrit pas une personne distraite : elle décrit une personne qui observe et trouve ce qu'elle s'attendait à trouver.

Le second sera appelé « ajustement ». Il indique dans quelle mesure le retour d'information provenant des autres est relativisé, réinterprété ou simplement ignoré lorsqu'il contredit le jugement présumé.

Ces deux paramètres décrivent différents aspects d'un même système.

La présupposition concerne la manière dont le jugement présumé est construit.

La remise s'applique à la manière dont elle est maintenue.

Pourquoi faut-il deux paramètres ?

À première vue, un seul paramètre pourrait sembler suffisant.

Ce n'est pas le cas : il existe deux configurations qu'un seul paramètre ne parvient pas à distinguer, et l'une d'entre elles a contraint à corriger le modèle. L'exemple suivant sert précisément à illustrer cela.

Dans sa première formulation, le modèle prévoyait que le jugement présumé soit une fonction de l'auto-perception. Une auto-évaluation très positive aurait dû entraîner une représentation tout aussi positive du jugement des autres : même mécanisme, signe inversé.

Une étude menée par Erika Carlson, Simine Vazire et Thomas Oltmanns (2011) montre que les choses ne se passent pas ainsi. Les personnes présentant des traits narcissiques marqués savent que les autres les perçoivent moins favorablement qu'elles ne se perçoivent elles-mêmes ; leurs métaperceptions s'avèrent moins faussées que leurs perceptions de soi. Elles reçoivent donc un signal relativement précis — et le rejettent. Selon les auteurs, l'un des moyens par lesquels elles conservent une image de soi aussi favorable consiste à considérer que les autres ne sont pas capables de reconnaître leur valeur.

Il convient de préciser que la littérature n'est pas unanime sur ce point : parallèlement à la position que nous venons de décrire, il existe une orientation selon laquelle les narcissiques surestimeraient également le jugement d'autrui, incapables qu'ils sont d'adopter le point de vue de ceux qui les observent. Le résultat utilisé ici est celui qui bénéficie du plus grand soutien, mais ce n'est pas le seul disponible.

Dans le cas de l'anxiété sociale, le tableau diffère de ces deux cas : la personne a tendance à la fois à construire le jugement présumé principalement à partir de sa perception de soi, et à dévaloriser les informations positives qu'elle reçoit.

Un seul paramètre ne peut décrire deux configurations aussi différentes. Le modèle à une variable était erroné, et il a été remplacé.

Quatre combinaisons

Si l'on considère conjointement la présupposition et la réduction, on obtient quatre combinaisons théoriquement possibles.

Une présupposition élevée associée à une dévalorisation élevée correspond au profil typique de l'anxiété sociale.

Une présupposition faible mais une dévalorisation élevée décrit le fonctionnement observé dans le narcissisme.

Les deux autres combinaisons constituent des prédictions du modèle et doivent être considérées comme telles.

Une présupposition élevée mais une remise réduite devrait produire une image de soi instable, facilement modifiable par les retours extérieurs : la perception de la façon dont on est perçu oscillerait à chaque rencontre.

Une présupposition faible et une décote réduite représentent en revanche le fonctionnement le mieux calibré, dans lequel le jugement présumé est construit à partir à la fois de l'autoperception et du retour d'information réellement disponible. Ce n'est pas un trouble : c'est le cas où le système fait ce pour quoi il semble avoir été conçu.

Ces deux dernières combinaisons n'ont pas encore été vérifiées. Elles ont été formulées avant de consulter la littérature, précisément pour ne pas les adapter aux résultats, et restent des prévisions qui pourraient s'avérer fausses. Si la présupposition et la décote s'avéraient étroitement corrélées dans la population — c'est-à-dire si les deux combinaisons mixtes étaient pratiquement vides —, alors il n'y aurait qu'un seul paramètre, et la distinction proposée dans ce chapitre aurait été un artifice.

.

Une conséquence pratique

La distinction entre présupposition et actualisation modifie également la manière d'interpréter le changement.

Si le problème réside principalement dans la présupposition, il faut augmenter la quantité et la qualité des informations provenant de l'extérieur.

Si, en revanche, le problème concerne la remise, l'information est déjà disponible : elle est reçue et dévalorisée. En ajouter davantage ne sert à rien, et dans certains cas, cela aggrave la situation.

Il s'agit de deux problèmes distincts, qui nécessitent des interventions différentes. Nous reviendrons sur ce point lorsque nous aborderons le recalibrage.

Le lien vers le chapitre suivant

Jusqu'à présent, nous avons considéré le jugement présumé comme le produit d'informations provenant de l'intérieur et de l'extérieur.

Il existe cependant une autre particularité du système.

Même lorsque personne ne nous observe, nous continuons souvent à nous comporter comme si quelqu'un pouvait nous juger.

Nous éprouvons de la honte même lorsque nous sommes seuls.

Nous tenons des promesses que personne ne pourrait vérifier.

Nous nous sentons coupables même lorsque personne ne saura jamais ce que nous avons fait.

Pour comprendre ce phénomène, nous devons introduire une dernière figure : l'observateur intériorisé.

6 L'observateur que nous portons en nous

Dans le chapitre précédent, nous avons vu que le jugement présumé se construit principalement à partir de la perception de soi et seulement en partie à partir du retour d'information provenant des autres.

Cette conclusion semble toutefois entrer en contradiction avec une expérience courante.

Pourquoi continuons-nous à nous sentir jugés même lorsque personne ne nous observe ?

Nous éprouvons de la honte pour un geste accompli en toute solitude.

Nous tenons une promesse que personne ne pourrait vérifier.

Nous nous sentons coupables d'un comportement qui restera à jamais inconnu.

Si le jugement des autres est le moteur du système, pourquoi continue-t-il d'agir même en leur absence ?

La réponse la plus intuitive est qu'au fil de la vie, le regard extérieur s'intériorise progressivement.

Une idée ancienne, et qu'est-ce qui la distingue ici ?

Cette idée n'est pas nouvelle, et elle a des antécédents illustres qu'il convient toutefois de prendre pour ce qu'ils sont : des antécédents partiels, non équivalents.

Le spectateur impartial d'Adam Smith, dans la *Théorie des sentiments moraux* (1759), est avant tout un dispositif normatif : un observateur idéal à l'aune duquel nous nous évaluons moralement.

L'« autre généralisé » de George Herbert Mead est une construction sociologique : l'intériorisation des attentes de la communauté.

L'observateur dont il est question ici est, quant à lui, un mécanisme cognitif prédictif. La différence n'est pas une question de degré. Le spectateur impartial et l'autre généralisé ont une cible unique et stable — un idéal moral, la communauté — tandis que l'observateur décrit ici change de cible : il simule le partenaire, le supérieur, un inconnu, un groupe particulier, et passe de l'un à l'autre d'un instant à l'autre. Ce n'est pas la voix d'une norme : c'est l'estimation d'une réaction.

Le point décisif, en tout état de cause, n'est pas de déterminer si un observateur intériorisé existe. Les preuves suggèrent qu'il existe. Le problème est d'en comprendre le rôle, et surtout son poids.

Un observateur réel, mais incomplet

Le comportement ne devient pas totalement égoïste lorsque personne ne peut savoir ce que nous faisons.

Le paradigme le plus couramment utilisé pour le vérifier est le jeu du dictateur en conditions de double aveugle, introduit par Elizabeth Hoffman, Kevin McCabe et Vernon Smith dans les années 1990 : la procédure est conçue de telle sorte que même l'expérimentateur ne puisse remonter

jusqu'au choix de chaque participant. Dans ces conditions, la part allouée à l'autre diminue nettement par rapport aux situations où l'anonymat n'est qu'apparent.

Elle diminue, cependant, sans pour autant tomber à zéro.

Une partie des participants continue de donner quelque chose. En dehors du laboratoire, de nombreuses personnes continuent de tenir leurs promesses, de dire la vérité ou d'aider les autres même lorsqu'aucun avantage en termes de réputation n'est possible.

Il existe donc quelque chose qui continue d'exercer une pression même en l'absence d'observateurs réels.

Il est naturel d'interpréter cela comme une forme intériorisée du jugement social.

Pourquoi il ne suffit pas de parler de conscience

Cette conclusion risque toutefois d'expliquer trop de choses.

Chaque fois qu'un comportement continue de se manifester en l'absence d'observateurs, nous pourrions nous contenter de dire que c'est l'observateur intériorisé qui intervient.

Mais une explication qui peut s'adapter à n'importe quel résultat n'exclut rien. C'est le mécanisme par lequel certaines théories psychologiques du passé ont perdu leur crédibilité : non pas parce qu'elles étaient fausses, mais parce qu'aucun résultat ne pouvait les contredire.

Pour être utile, le concept doit produire au moins une prédiction susceptible de s'avérer fausse.

Dans le modèle proposé dans cet essai, la prédiction est simple.

L'observateur intériorisé ne remplace pas l'observateur externe. Il en représente une version systématiquement atténuée.

Une prédiction du modèle

Si cette interprétation est correcte, le comportement privé ne devrait jamais coïncider parfaitement avec le comportement public.

Lorsque le jugement réel des autres devient impossible, le comportement prosocial devrait diminuer, mais pas disparaître complètement.

C'est ce que l'on observe dans le jeu du dictateur évoqué plus haut, et que les expériences sur l'anonymat confirment assez régulièrement : la coopération, la générosité et l'honnêteté diminuent, mais aucune des trois ne disparaît.

Cette prédiction est importante car elle empêche d'utiliser l'observateur intériorisé comme explication universelle.

Si le comportement privé s'avérait impossible à distinguer du comportement public, le modèle serait réfuté.

Si, au contraire, tout comportement coopératif disparaissait dans l'anonymat, il serait tout aussi réfuté.

Le modèle prévoit une situation intermédiaire, et c'est une prédiction que l'on peut vérifier.

Une heuristique, et non un substitut de l'observateur externe

Il reste à déterminer la nature de cette atténuation.

Le système se comporte comme si la réputation était en jeu même lorsqu'elle ne l'est pas. Non pas parce qu'il calcule délibérément les conséquences futures, mais plutôt parce qu'il applique un raccourci : dans le doute, se comporter comme si quelqu'un observait.

Pendant la majeure partie de leur histoire, les êtres humains ont vécu en petits groupes, relativement stables et dépourvus de véritable anonymat. Dans un tel environnement, ce raccourci était presque toujours correct, car presque toujours quelqu'un observait effectivement, ou l'aurait su peu après.

Les situations d'anonymat garanti, fréquentes aujourd'hui, constituent pour ce système une condition anormale. Le raccourci continue néanmoins de fonctionner.

Ce qui est naturel et ce qui est acquis dans l'intériorisation du jugement d'autrui

À ce stade, il convient de distinguer deux choses qui sont souvent confondues, et que ce chapitre risquerait autrement de superposer.

Le raccourci que nous venons de décrire est une disposition propre à l'espèce : il ne nécessite pas qu'une personne en particulier ait été jugée d'une manière particulière. Il s'appliquerait, en principe, même à ceux dont l'histoire sociale serait très différente de la nôtre.

La partie acquise est en revanche un réglage individuel : qui nous a jugés, selon quels critères et avec quelle sévérité, contribue à déterminer ce que notre observateur intérieur remarque et à quel point il est exigeant.

Les deux composantes coexistent et ne sont pas en concurrence. La première explique pourquoi ce mécanisme existe chez tout le monde ; la seconde, pourquoi il varie autant d'une personne à l'autre.

Elles sont également, du moins en principe, distinguables : la composante spécifique devrait présenter moins de variabilité entre les individus et les cultures que la composante acquise. Si elle ne présentait aucune différence, cette distinction serait inutile.

L'intériorisation n'explique pas tout

Il reste un fait significatif.

L'observateur intériorisé ne produit pas la même force que le jugement réel.

Cela aide à comprendre pourquoi de nombreuses personnes respectent les normes avec plus de rigueur lorsqu'elles sont observées, et pourquoi la réputation conserve un pouvoir régulateur qu'aucune conviction privée ne semble égaler.

Une distinction utile, et étayée par des données empiriques, est celle entre la honte et la culpabilité étudiée par June Tangney. La honte est centrée sur l'image de soi face aux autres, et s'accorde bien avec une interprétation axée sur la réputation. La culpabilité est centrée sur le préjudice causé, elle se nourrit d'empathie et peut être intense même dans le plus grand secret. Tout ce qui nous retient lorsque personne ne regarde n'est donc pas uniquement attribuable à un observateur intérieur : une partie relève d'un autre registre.

Le système ne fonctionne pas comme s'il existait deux observateurs équivalents.

Il existe un observateur réel et un observateur intériorisé. Le second découle du premier, mais ne le remplace pas.

Un lien avec le chapitre précédent

Le chapitre précédent montrait que le jugement présumé se construit avant tout de l'intérieur.

Cela pourrait laisser penser à un système presque entièrement autoréférentiel.

Le présent chapitre apporte une précision.

Cet « intérieur » n'est pas né de rien. Il s'agit en partie d'une disposition propre à l'espèce, et en partie du résultat d'une longue histoire d'interactions réelles.

L'observateur intériorisé n'est donc pas une alternative à la société : il est, du moins pour sa composante acquise, la société qui continue d'agir lorsque les autres ne sont pas présents.

Le lien vers le chapitre suivant

Jusqu'ici, nous avons examiné le jugement présumé sous son aspect passif : la manière dont nous imaginons être évalués.

Mais le système est récursif.

Nous ne nous contentons pas d'anticiper le jugement des autres. Nous jugeons sans cesse des personnes, des idées, des œuvres et des comportements.

Et ces jugements, à leur tour, sont observés et évalués.

Pour comprendre l'autre moitié du modèle, nous devons donc changer de perspective.

Il ne s'agit plus de la façon dont nous pensons être jugés, mais de la manière dont notre façon de juger devient elle-même un indice sur nous-mêmes.

DEUXIÈME PARTIE — Les deux facettes

7 Juger, c'est s'exposer

Jusqu'à présent, nous avons suivi le jugement social dans sa facette passive : la manière dont nous imaginons être évalués par les autres. À partir de maintenant, nous devons changer de perspective.

Nous n'introduisons pas un nouveau principe. C'est le même mécanisme décrit dans les chapitres précédents qui révèle désormais sa deuxième facette.

Nous ne sommes pas seulement l'objet d'un jugement.

Nous sommes aussi, en permanence, ceux qui jugent.

Nous approuvons et condamnons, nous faisons l'éloge et ridiculisons, nous admirons et méprisons. Cela pourrait sembler être un processus complètement différent. En réalité, c'est la continuation du même système. Si le comportement des autres nous permet de déduire comment ils vont probablement nous juger, notre comportement permet également aux autres de déduire comment nous allons probablement juger.

Le jugement social est récursif.

Chaque critère en dit long sur celui qui l'utilise

Lorsque nous affirmons qu'un comportement est juste, faux, admirable ou honteux, il semble que notre jugement ne concerne que l'objet de l'évaluation.

En réalité, il en dit aussi long sur celui qui l'exprime.

Tout critère rendu public suggère que nous le considérons suffisamment valable pour l'utiliser afin d'évaluer les autres.

C'est précisément pour cette raison que ce critère devient potentiellement applicable à nous-mêmes également.

C'est le mécanisme évoqué dans le premier chapitre. Un jugement n'est pas sans effet : il oriente le comportement de celui qui l'a formulé vers ce qu'il a jugé et/ou vers les autres qui portent un jugement sur le même objet. Et c'est précisément pour cette raison qu'il devient lisible — il n'est pas nécessaire qu'il soit déclaré, car la manière dont nous traitons quelqu'un laisse des traces que les autres savent interpréter. Celui qui juge, explicitement ou implicitement, s'expose donc à être jugé pour son propre jugement.

Non pas parce qu'il existe une règle morale qui l'impose.

Mais parce que le critère lui-même constitue une information sur la personne qui le propose.

Les jugements sont des signaux

Cette observation modifie la manière d'interpréter de nombreux comportements sociaux.

Une condamnation publique ne se contente pas d'indiquer qu'un certain comportement est considéré comme répréhensible.

Elle indique également, au moins en partie, que celui qui la prononce souhaite être associé au critère exprimé.

La composante d'autoprésentation du jugement n'est pas constante.

Elle s'accroît lorsque le jugement est public, lorsque les observateurs sont nombreux ou influents et lorsque le sujet concerne des caractéristiques importantes pour la réputation.

Elle diminue dans les situations privées, mais ne disparaît pas complètement. Comme nous l'avons vu dans le chapitre précédent, nous continuons souvent à nous comporter comme si quelqu'un pouvait nous observer, même lorsque personne n'est présent.

Les aspects passif et actif sont donc régis par le même principe.

Quand la condamnation remplace le don

Une série de recherches menées par Jillian Jordan et David Rand (2019) illustre ce mécanisme de manière particulièrement claire.

Les expériences se déroulent dans des conditions anonymes et non répétitives, dans lesquelles aucune réputation ne peut être construite. Dans certaines versions, le participant disposait d'un moyen direct de se montrer fiable et altruiste — céder une partie de son argent à un tiers ; dans d'autres versions, cette possibilité n'était pas prévue.

Lorsqu'elle n'était pas prévue, l'indignation envers ceux qui s'étaient comportés de manière égoïste s'avérait plus forte.

Ces deux comportements se substituent l'un à l'autre, et c'est cette substitution qui révèle ce qu'ils ont en commun : donner et condamner disent aux autres la même chose sur ceux qui les pratiquent. Une fois un canal fermé, l'autre subsiste.

La condamnation ne sert donc pas seulement à sanctionner un comportement. Elle sert aussi à communiquer quelque chose sur celui qui condamne — ce qui est la thèse de ce chapitre, ici observée en laboratoire plutôt que démontrée.

Il convient d'ajouter, afin que le résultat ne promette pas plus qu'il n'apporte, que l'anonymat de ces expériences est procédural et non phénoménologique : l'expérimentateur reste un observateur potentiel, et il n'est pas possible de vérifier directement dans quelle mesure les participants croient réellement ne pas être observés.

Pourquoi détestons-nous les hypocrites ?

Cette interprétation trouve une confirmation particulièrement forte dans une étude de Jillian Jordan, Roseanna Sommers, Paul Bloom et David Rand (2017) sur l'hypocrisie.

L'idée intuitive est que les hypocrites sont méprisés parce qu'ils manquent de cohérence.

Les expériences racontent une autre histoire, et ce à travers deux comparaisons distinctes.

La première concerne la gravité attribuée au comportement. Celui qui condamne une transgression qu'il commet ensuite est jugé plus sévèrement que celui qui ment ouvertement sur sa propre conduite morale, en déclarant ne pas commettre ce qu'il commet en réalité. Le mensonge direct s'avère moins grave que la condamnation hypocrite.

La deuxième comparaison est celle qui met en évidence le mécanisme. Les chercheurs ont créé le profil de l'« hypocrite honnête » : une personne qui condamne un comportement tout en admettant l'adopter elle-même. L'incohérence entre ce qu'elle dit et ce qu'elle fait est identique à celle de l'hypocrite classique. Pourtant, elle n'est pas jugée négativement.

La différence entre les deux cas ne réside pas dans l'incohérence, qui est la même.

C'est le signal.

L'hypocrite utilise la condamnation pour obtenir le bénéfice de réputation associé à un critère qu'il ne respecte pas. L'hypocrite honnête, en déclarant sa propre incohérence, renonce à ce bénéfice par le fait même de condamner.

Ce qui est sanctionné, ce n'est pas la contradiction. C'est la tentative d'obtenir les avantages d'un signal fallacieux.

Il est difficile d'imaginer une confirmation plus directe de la thèse de ce chapitre : si la condamnation n'était pas aussi une forme d'auto-présentation, l'hypocrite honnête et l'hypocrite ordinaire devraient s'avérer tout aussi déplaisants.

Distinction et appartenance

Pierre Bourdieu a montré que les goûts, les préférences et les jugements ne se contentent pas de classer les objets auxquels ils se rapportent.

Ils classifient également ceux qui les expriment.

Les préférences esthétiques, culturelles et morales deviennent des indicateurs d'appartenance, de compétence et de position sociale.

Il en va de même pour les critères moraux.

Lorsque nous portons un jugement, nous ne nous contentons pas de décrire le monde. Nous contribuons également à nous décrire nous-mêmes.

Pourquoi simulons-nous nos jugements ?

Si chaque critère public a des conséquences sur la réputation, on comprend mieux un phénomène étudié par Timur Kuran.

Les gens ne se contentent pas de falsifier leurs opinions.

Elles peuvent aussi feindre leur indignation, leur approbation et même leur silence.

Un critère exprimé publiquement en dit long sur l'identité sociale de celui qui l'émet.

C'est pourquoi le système exerce une pression constante pour que l'expression publique de ses jugements s'aligne sur les attentes de l'entourage, même lorsque ces jugements ne coïncident pas avec ceux que l'on défend réellement.

Le pacte que personne ne conclut

Il y a une conséquence de ce qui précède que le modèle prévoit, et que chacun reconnaît sans jamais l'avoir formulée.

Si exprimer un jugement expose celui qui l'exprime, alors deux personnes qui se trouvent face à face ont toutes deux plus à perdre qu'à gagner à se juger ouvertement. Il en résulte une réticence réciproque : je ne dis pas ce que je pense de toi, et tu ne dis pas ce que tu penses de moi. Ce n'est pas un accord — personne ne le conclut, personne ne s'en souvient, et l'appeler « pacte » relève de la métaphore. C'est un équilibre : il tient parce qu'il arrange les deux parties, et il se rompt dès lors qu'il cesse d'arranger l'une d'entre elles.

Erving Goffman a décrit cette structure sous le nom de « face-work ». Chacun, dans l'interaction, défend une certaine image de soi ; les autres l'acceptent, et en échange obtiennent que la leur soit acceptée. Goffman appelle cette entente « working consensus » et insiste sur un point que le modèle reprend : elle n'est pas négociée, et c'est précisément pour cela qu'elle est solide. Penelope Brown et Stephen Levinson en ont ensuite fait le fondement de la théorie de la courtoisie, selon laquelle une grande partie des formes linguistiques que nous utilisons s'explique par une protection réciproque contre l'exposition.

Le modèle ajoute une prédiction. Si la réticence dépend de l'exposition, elle devrait s'atténuer là où l'exposition diminue : dans l'anonymat, dans les relations fortement asymétriques, et lorsque l'objet du jugement est un tiers absent. C'est ce que l'on observe. On parle beaucoup plus librement de ceux qui ne sont pas présents, et ceux qui détiennent le pouvoir se permettent des jugements que ceux qui n'en ont pas gardent pour eux.

Le critère que l'on n'adopte pas

Ce principe s'applique également à l'inverse, et le cas le plus visible est celui des coutumes partagées — modes, goûts, façons de parler, positions qu'il est de bon ton d'adopter à un moment donné. Une coutume, au fond, est un critère rendu visible par l'usage.

Si un critère exprimé en dit long sur celui qui l'exprime, alors un critère répandu que l'on n'adopte *pas* en dit également long. Celui qui ne suit pas ce que les autres suivent — une mode, un goût, une façon de parler, une position qu'il est de rigueur d'adopter à ce moment-là — n'est pas perçu comme indifférent. Il est perçu comme quelqu'un qui a jugé, et mal jugé, ceux qui suivent cette tendance. L'abstention n'est pas perçue comme une absence de jugement, conformément à ce qu'a établi le premier chapitre : décider que quelque chose ne mérite pas d'attention est déjà un jugement.

Ce phénomène a été mesuré. Julia Minson et Benoît Monin ont montré que les consommateurs de viande évaluent les végétariens d'autant plus négativement que l'attente d'être jugés moralement par eux est mise en avant. La dévalorisation ne porte pas sur le choix alimentaire : elle porte sur le jugement qui lui est attribué. Les auteurs l'appellent « *do-gooder derogation* », c'est-à-dire la dévalorisation de celui qui fait ce qu'il faut, et en identifient la fonction : désamorcer à l'avance un reproche attendu.

Cela permet de comprendre un fait qui, autrement, serait curieux : ceux qui adoptent un critère plus exigeant que celui en vigueur sont davantage sanctionnés que ceux qui en adoptent un plus laxiste, bien qu'ils n'aient fait de mal à personne. Dans le cadre de ce modèle, cela n'a rien de curieux. Un critère déclaré valide évalue également ceux qui ne l'appliquent pas, et ceux qui se sentent évalués réagissent à un jugement que personne n'a exprimé : le cas le plus pur de jugement présumé régissant le comportement.

Il en découle également que la pression à se conformer ne naît pas seulement du désir de ressembler à quelqu'un. Elle naît du coût de ne pas le faire — qui n'est pas le coût d'être différent, mais celui d'être perçu comme quelqu'un qui juge.

Ostentation morale

Il existe une catégorie de comportements dans lesquels la mise en avant de sa propre vertu semble prévaloir sur la contribution effective à la cause défendue. Dans le débat public, ces comportements sont souvent désignés par l'expression journalistique « *virtue signalling* », en circulation depuis le milieu des années 2010 et presque toujours utilisée dans un sens accusatoire.

Le concept correspondant en philosophie morale a été élaboré et analysé par Justin Tosi et Brandon Warmke sous le nom de « *moral grandstanding* » : l'utilisation du discours moral public pour impressionner les autres quant à ses propres qualités morales.

Ce phénomène existe bel et bien.

Dans le cadre du modèle développé dans cet essai, cependant, il représente un cas particulier d'un processus beaucoup plus général.

Tout jugement public comporte également une composante de réputation, dont l'intensité varie selon le contexte.

L'ostentation morale n'introduit donc pas un mécanisme nouveau. C'est le cas où cette composante devient suffisamment évidente pour être remarquée — et, de ce fait, le cas où le signal cesse de fonctionner.

Une prévision encore ouverte

Les recherches évoquées dans ce chapitre montrent déjà que la sanction de l'hypocrisie dépend de la falsification du signal et non d'une simple incohérence. Sur ce point, nous ne formulons pas de prédiction : nous rapportons un résultat.

Il reste toutefois une conséquence du modèle qui n'a pas encore été vérifiée.

Si le jugement public fonctionne également comme un signal de réputation, il est raisonnable de s'attendre à ce que le coût de la violation augmente à mesure que s'élargit le champ d'application dans lequel le critère a été déclaré valable.

Celui qui présente un principe comme valable pour tout le monde devrait être jugé plus sévèrement, lorsqu'il le viole, que celui qui le présente comme une préférence personnelle ou comme une norme limitée à certaines circonstances.

Le lien vers le chapitre suivant

À ce stade, le lecteur pourrait tirer une conclusion désagréable.

Si nos jugements publics fonctionnent comme des signaux, il semblerait que la vie sociale consiste en une vaste mise en scène, et que derrière chaque critère affiché se cache un calcul.

Le chapitre suivant montrera que les choses sont encore plus étranges que cela.

Nous ne jouons pas la comédie, ou du moins pas principalement.

Lorsque nous adaptons un jugement à notre auditoire, ce jugement a tendance à ne pas rester une simple posture. La convenance ne produit pas seulement de l'hypocrisie : elle modifie souvent aussi ce que nous finissons par croire sincèrement.

C'est cette transition qui rend le modèle plus inquiétant, et en même temps moins cynique, que ce que ce chapitre laisse supposer.

8 Exprimer un jugement nous change

Le chapitre précédent pourrait laisser au lecteur une impression désagréable.

Si nos jugements publics fonctionnent également comme des signaux de réputation, il semblerait qu'une grande partie de la vie sociale soit une mise en scène consciente. Nous dirions ce qu'il convient de dire, nous afficherions les opinions les plus avantageuses et nous condamnerions ce qui produit le meilleur effet sur les autres.

Cette impression est compréhensible, et en partie fondée. Les êtres humains mentent, simulent et font semblant ; l'hypocrisie existe, et le chapitre précédent en a décrit une forme particulière.

Mais si la plupart de nos jugements étaient le résultat d'un calcul délibéré, nous devrions sans cesse nous rappeler quelles sont nos opinions authentiques et lesquelles nous affichons. Cela ne semble pas être le cas.

Dans la plupart des cas, nous ne jouons pas un rôle.

Nous finissons par croire ce que nous avons dit.

Le message adapté à l'auditeur

Ce phénomène a été étudié en laboratoire à l'aide d'un paradigme simple.

On fournit à un participant une description ambiguë d'une personne : certains comportements se prêtent à une interprétation favorable, d'autres à une interprétation défavorable. On lui demande ensuite de décrire cette personne à un interlocuteur dont il connaît l'opinion. Comme on pouvait s'y attendre, le participant adapte sa description : il la rend plus bienveillante si l'interlocuteur est bien disposé, plus critique s'il est hostile.

Jusque-là, rien de surprenant. C'est de la courtoisie, ou une adaptation communicative.

Le résultat intéressant vient ensuite.

Lorsque, dans un deuxième temps, on demande au participant ce qu'il pense réellement de cette personne — et ce dont il se souvient de la description initiale —, son jugement et son souvenir se sont décalés dans le sens du message qu'il avait formulé.

Non seulement il s'est exprimé différemment, mais il a également commencé à penser différemment.

La communication, en d'autres termes, ne se limite pas à transmettre des convictions déjà formées. Elle contribue à les construire. Chaque fois que nous adaptons un jugement à notre interlocuteur, nous modifions également légèrement la représentation que nous en gardons nous-mêmes.

Cet effet a été baptisé « *saying-is-believing* » (dire, c'est croire) et a été décrit par Tory Higgins et ses collaborateurs dès la fin des années 1970. Des recherches ultérieures l'ont mis en relation avec le besoin de construire avec l'interlocuteur une représentation partagée de la réalité : la modulation n'est pas une adaptation superficielle, mais la manière dont une compréhension commune s'établit — et ce qui est établi ensemble tend alors à être ce dont on se souvient et ce que l'on croit.

Trois théories apparentées

Ce phénomène est parfois confondu avec trois autres, plus connus, dont il convient de le distinguer.

La théorie de la dissonance cognitive de Leon Festinger décrit le malaise provoqué par la contradiction entre ce que nous faisons et ce que nous pensons, ainsi que la tendance à le réduire en modifiant l'un ou l'autre.

La théorie de l'auto-perception de Daryl Bem propose que, dans de nombreuses circonstances, nous déduisons nos attitudes en observant notre propre comportement, presque comme le ferait un observateur extérieur.

La théorie de l'équilibre de Fritz Heider décrit la pression visant à maintenir la cohérence des relations au sein d'une triade : si j'estime quelqu'un et que je découvre qu'il évalue différemment de moi une chose qui nous concerne tous les deux, l'un des deux jugements a tendance à céder.

Ces trois théories ont un point commun : elles décrivent une pression qui s'exerce a posteriori, lorsqu'une divergence s'est déjà manifestée.

L'effet décrit ici intervient plus tôt, et porte sur un autre type de situation. Il ne concerne pas le comportement en général, mais la communication adressée à quelqu'un : ce qui modifie la conviction, ce n'est pas le fait d'avoir agi, mais celui d'avoir formulé un jugement à l'intention d'un destinataire dont on connaissait la position. La divergence ne s'est pas encore produite — elle est évitée.

C'est le cas social, et c'est précisément le cas dont traite cet essai.

De l'adaptation à la conviction

Ce mécanisme aide à comprendre de nombreux changements apparemment spontanés.

Une personne entre dans un nouvel environnement — un lieu de travail, un groupe, une communauté professionnelle.

Au début, elle adopte le langage, les critères et les jugements courants, surtout pour permettre l'échange : ce sont les attentes de son interlocuteur, et les formuler autrement nécessiterait sans cesse des explications et créerait des frictions.

Au bout de quelques années, elle peut avoir l'impression d'avoir toujours pensé ainsi.

Ce changement n'est pas nécessairement le résultat d'une simulation réussie. Dans la plupart des cas, il n'y a jamais eu de simulation : l'écart entre ce qui était dit et ce qui était pensé s'est progressivement réduit, sans que personne ne s'en aperçoive.

C'est aussi la raison pour laquelle une conversion de ce type est presque impossible à reconnaître de l'intérieur. Il n'y a pas de moment précis où l'on a décidé de changer d'avis.

Le cas des modes

Ce processus s'observe à grande échelle dans les modes — au sens large du terme : non seulement l'habillement, mais aussi les goûts musicaux, les lectures, les positions qu'il convient d'adopter, les mots que l'on utilise et ceux que l'on n'utilise plus.

Georg Simmel en a donné, au début du XX^e siècle, une description qui tient toujours : la mode vit de la tension entre deux pulsions opposées, imiter ceux que l'on souhaite rattraper et se distinguer de ceux que l'on a laissés derrière soi. C'est une description de la dynamique collective. Le mécanisme individuel est celui de ce chapitre.

Celui qui se conforme à une coutume répandue ne reste pas dans la position de celui qui feint. L'adaptation à l'auditeur modifie ce que l'on finit par croire, et l'auditeur, ici, est un public vaste et anonyme dont on ne connaît que les attentes moyennes. Il s'ensuit que l'adhésion à une mode n'est presque jamais de l'hypocrisie : celui qui trouve beau ce qu'il faut trouver beau à ce moment-là le trouve vraiment beau.

Et on s'en rend compte aussi parce que le changement d'une mode apparaît, rétrospectivement, incompréhensible à ceux qui l'avaient suivie. Nous n'avons pas cessé de faire semblant. Nous avons cessé de croire, en même temps que le public qui nous écoutait.

Le même défaut des deux côtés

À ce stade, on peut percevoir la symétrie qui sous-tend toute la deuxième partie de cet essai.

Dans le cinquième chapitre, nous avons vu que nous ne savons pas précisément comment les autres nous perçoivent : l'image que nous en avons est en grande partie construite à partir de nous-mêmes.

Nous voyons maintenant que nous ne savons pas non plus précisément pourquoi nous pensons ce que nous pensons : une partie de nos jugements s'est formée en s'adaptant à ceux qui nous écoutaient, et ne porte aucune trace de cette origine.

Les deux camps présentent le même défaut, et ce pour la même raison.

Dans les deux cas, le mécanisme qui devrait s'en rendre compte est le même que celui qui en est affecté. Il n'existe pas, au sein du système, de point d'observation externe à partir duquel vérifier ni comment nous sommes perçus, ni pourquoi nous croyons ce que nous croyons.

Telle est la formulation la plus précise de la thèse énoncée dans l'introduction : un processus récursif qui opère sur une donnée non observée mais déduite.

Si cette interprétation est correcte, cependant, il en découle une autre conséquence. Elle ne concerne plus les jugements que nous formulons, mais ceux qui ne parviennent même pas à se former.

Ce que nous sommes prêts à prendre en considération

Jusqu'à présent, nous avons parlé de jugements formulés : quelque chose que nous disons, et que nous finissons ensuite par croire.

Mais si le système anticipe les conséquences relationnelles de ce que nous exprimons, il n'y a aucune raison pour qu'il ne s'active qu'au moment de l'expression. Il pourrait intervenir plus tôt, lorsqu'une idée se présente et doit être examinée.

Avant même de nous demander « est-ce vrai ? », l'esprit pourrait avoir déjà répondu à une autre question : « qu'est-ce que cela impliquerait, pour ma vie sociale, de prendre cette idée au sérieux ? »

Cela me rapprocherait-il ou m'éloignerait-il des personnes qui comptent pour moi ? Comment le fait que je l'envisage, ne serait-ce que brièvement, serait-il perçu aux yeux de ceux qui me connaissent ?

Cela expliquerait pourquoi certaines idées semblent irrecevables avant même d'avoir été examinées, rejetées non pas par un argument mais par une résistance qui précède l'argument. Et pourquoi une thèse peut être ignorée non pas parce qu'elle est jugée fausse, mais parce que la prendre au sérieux modifierait la position de celui qui le ferait.

Nous pouvons également relire sous cet angle l'habitus de Pierre Bourdieu, déjà évoqué dans le chapitre précédent : des dispositions incorporées qui reflètent et reproduisent la position sociale de ceux qui les possèdent. Là, cela servait à expliquer que nos jugements nous classent ; ici, cela suggère quelque chose de plus, à savoir qu'ils peuvent sélectionner à l'avance ce que nous en venons à considérer. La différence par rapport à Bourdieu est qu'il ne s'agirait pas seulement d'une disposition ancrée, mais d'un filtre prédictif actif à chaque instant, qui agit sur les idées nouvelles comme il agit sur les comportements.

Il faut préciser clairement que cette extension n'est pas démontrée. C'est une conséquence plausible du modèle, et en tant que telle, elle peut être fausse.

Une précision s'impose toutefois, sans laquelle cette hypothèse deviendrait une arme. Si elle est correcte, elle ne concerne pas seulement les idées d'autrui. Elle concerne également ces pages, et la manière dont le lecteur les évalue en ce moment même — et elle concerne ceux qui les ont écrites, qui ne disposent d'aucun point d'observation privilégié sur leur propre travail.

Convergence et divergence

Il découle de ce qui précède un élément qui concerne les relations entre les personnes, et pas seulement les jugements de chacune.

Deux personnes qui se côtoient assez fréquemment adaptent leurs jugements l'une à l'autre, et d'après ce que nous avons vu dans ce chapitre, cette adaptation ne reste pas une simple posture : elle devient une conviction.

La convergence des jugements n'est donc pas seulement une condition de la coopération. Elle en est aussi un produit, et elle continue de se produire tant que la coopération dure.

Ce n'est pas un effet secondaire. Celui qui juge comme nous est prévisible, et c'est cette prévisibilité qui permet de s'engager avant de savoir comment les choses vont se terminer — cette capacité que le troisième chapitre attribuait au jugement.

La divergence, à l'inverse, n'est pas perçue comme une différence. Le chapitre précédent a établi que le fait de ne pas adopter un critère largement répandu est interprété comme un jugement sur ceux qui l'adoptent, et que celui qui se sent jugé réagit à un jugement que personne n'a exprimé. Il s'ensuit que les jugements divergents ont un coût relationnel même lorsqu'il n'y a rien à se partager : pour susciter de l'hostilité, nul besoin d'un conflit d'intérêts, la divergence suffit.

Et ce lien fonctionne également dans le sens inverse. Adopter un critère de l'adversaire reviendrait à le valider ; si l'extension proposée ci-dessus est correcte — le coût relationnel intervient avant même que l'idée ne soit examinée —, prendre au sérieux ce que l'autre soutient devient coûteux,

quel que soit le bien-fondé de son argument. Le conflit engendre la divergence, tout comme la divergence engendre le conflit.

Il en découle une prédiction vérifiable. Chez ceux qui coopèrent, les jugements devraient converger au-delà de ce que justifient les informations partagées ; chez ceux qui sont en conflit, ils devraient diverger au-delà de ce que justifient les informations. Ce n'est pas la convergence qui est prédite — pour cela, il suffirait que deux personnes disposent des mêmes données — mais son excès par rapport aux preuves disponibles.

Ce domaine est étudié depuis longtemps, sous des appellations telles que « polarisation de groupe » et « raisonnement motivé », et ces pages n'ont pas passé en revue la littérature à ce sujet. Ce qui précède est une conséquence du modèle, et non une contribution à ces discussions ; il est possible que ces idées y figurent déjà, sous une forme plus aboutie.

Pourquoi le système fonctionne-t-il ainsi ?

À première vue, cela ressemble à un défaut de conception. Un système qui modifie ses convictions en fonction de son interlocuteur semble peu fiable.

On peut toutefois formuler une hypothèse sur les raisons pour lesquelles cela est avantageux, à condition qu'elle reste clairement présentée comme telle.

Maintenir en parallèle deux versions de chaque jugement — celle privée et celle affichée — a un coût cognitif et comporte un risque : la divergence peut apparaître, et lorsqu'elle apparaît, l'atteinte à la réputation est celle de l'hypocrite décrite dans le chapitre précédent, c'est-à-dire la plus grave. Réduire progressivement l'écart entre les deux versions élimine à la fois le coût et le risque.

La conviction, une fois établie, ne nécessite pas de surveillance. La simulation, en revanche, oui.

Il convient de préciser que cette explication est plausible mais non prouvée : le fait qu'un trait soit aujourd'hui avantageux ne prouve pas qu'il ait été sélectionné pour cet avantage.

Une limite importante

Rien de ce qui précède n'implique que toutes les convictions naissent de l'adaptation sociale.

Beaucoup découlent de l'expérience directe, du raisonnement, de l'apprentissage ou de dispositions individuelles.

L'affirmation du modèle est plus restreinte : lorsqu'un jugement est formulé à plusieurs reprises à l'intention d'interlocuteurs dont on connaît la position, la probabilité qu'il s'oriente de manière stable dans cette direction augmente.

L'intensité de cet effet, et les conditions dans lesquelles il l'emporte sur d'autres processus, restent une question empirique.

Le lien vers le chapitre suivant

Les deux parties précédentes ont décrit un système calibré avec une grande précision pour une tâche précise : anticiper le jugement de ceux qui nous observent et s'adapter en conséquence.

Il subsiste toutefois une difficulté.

Cette tâche avait du sens dans un environnement très différent de celui dans lequel nous vivons aujourd'hui. Le système s'est formé au sein de petits groupes, où celui qui observait était aussi celui qui décidait, et où chaque jugement circulait dans son intégralité.

Aujourd'hui, nous sommes exposés à des observateurs qui ne décideront jamais de quoi que ce soit nous concernant, et à des jugements qui n'atteindront jamais personne.

Le chapitre suivant examine ce qu'il advient d'un mécanisme conçu pour un monde qui n'existe plus.

9 Un mécanisme calibré pour un monde qui n'existe plus

Dans les chapitres précédents, nous avons décrit un système étonnamment sophistiqué. D'une part, il reconstitue la façon dont les autres nous perçoivent probablement ; d'autre part, il modèle nos jugements en fonction des conséquences qu'ils auront. Ces deux aspects convergent vers une même tâche : prévoir la valeur sociale de nos actions et les ajuster en conséquence.

Si ce système avait été conçu pour le monde dans lequel nous vivons aujourd'hui, il devrait toutefois se comporter de manière très différente.

Chaque jour, nous rencontrons des personnes que nous ne reverrons jamais. Nous recevons des évaluations anonymes. Nous lisons les commentaires d'inconnus. Nous sommes observés par des individus qui n'ont aucun pouvoir de décision sur ce qui nous concerne. Dans ces conditions, le jugement de nombreux observateurs devrait être presque sans importance.

Et pourtant, ce n'est pas le cas.

Un simple regard, un rire, un commentaire sur Internet ou une critique d'un inconnu peuvent suffire à provoquer de la honte, de la gêne ou de l'irritation. Le système réagit comme si ce jugement avait des conséquences importantes, même lorsque nous savons parfaitement que, selon toute vraisemblance, il n'en aura aucune.

Le problème n'est pas d'expliquer pourquoi le jugement a de l'importance. Les chapitres précédents l'ont déjà fait.

Le problème est d'expliquer pourquoi il continue à compter alors qu'il ne devrait pas.

Une question restée en suspens

Cette difficulté était déjà apparue, de manière implicite, dans le chapitre consacré au poids du jugement.

Nous y avons établi que le poids d'un jugement dépend du pouvoir que celui qui le formule exerce sur nous, et nous en avons tiré une question : cette personne peut-elle changer quelque chose dans ma vie ? Nous avons également observé que, dans la plupart des cas, la réponse est non — et que le trouble persiste malgré tout, car le système qui le produit ne procède pas lui-même à cette vérification.

La question du « pourquoi » était restée en suspens.

Il est temps d'y répondre, car cette réponse explique à la fois l'inadéquation du mécanisme et la limite de l'outil pratique que l'essai proposera.

Le jugement n'est pas la valeur finale

La réponse proposée par cet essai est que le jugement ne constitue pas le véritable objectif du système.

La valeur finale réside dans les conséquences relationnelles.

Être choisi ou exclu. Obtenir la coopération ou essuyer un refus. Conserver ou perdre des opportunités futures.

Le jugement n'est que le signal qui, dans la plupart des situations où le système s'est formé, permettait de prévoir ces conséquences.

Le système n'accorde donc aucune valeur au jugement en soi. Il l'utilise comme un indicateur rapide des conséquences qui s'ensuivent habituellement.

Tout comme le thermomètre n'est pas la fièvre mais un moyen rapide de l'estimer, de même le jugement n'est pas le préjudice social mais le signal qui, en règle générale, l'anticipe.

C'est le rôle que l'on appelle, dans un système de prévision, *un prédicteur* : un signal rapide à détecter, qui remplace un résultat lent à se manifester ou difficile à mesurer. Le jugement d'autrui, pour ce système, est un prédicteur des conséquences relationnelles.

Pendant la majeure partie de l'histoire de l'espèce, ce raccourci a probablement bien fonctionné.

Dans les petits groupes où les êtres humains ont vécu pendant la quasi-totalité de leur passé, presque tout le monde se connaissait, presque tout le monde se serait revu et presque tout jugement aurait eu des conséquences concrètes. Celui qui critiquait quelqu'un aujourd'hui pouvait décider demain de l'aider, de l'éviter, de le choisir comme allié ou de l'exclure de toute coopération.

Dans un tel environnement, le jugement et les conséquences coïncidaient presque toujours. Il était donc pratique d'utiliser le premier comme indicateur des secondes.

Il convient de préciser que cette reconstitution est une inférence plausible sur les conditions dans lesquelles ce mécanisme s'est formé, et non un fait observé. Le fait qu'un trait puisse aujourd'hui s'expliquer en ces termes ne prouve pas qu'il ait été sélectionné pour cette raison.

Quand le prédicteur continue de fonctionner seul

Le problème survient lorsque l'indicateur survit aux conditions qui le rendaient fiable.

Aujourd'hui, nous sommes constamment évalués par des personnes qui n'auront jamais l'occasion d'influencer notre vie. Le lien entre jugement et conséquences s'est profondément affaibli.

Le mécanisme, cependant, continue d'utiliser le jugement comme s'il s'agissait encore d'un bon prédicteur.

C'est pourquoi nous pouvons rougir devant un touriste que nous ne reverrons jamais, nous inquiéter du commentaire d'un internaute anonyme ou repenser pendant des jours à une gaffe commise devant des personnes dont nous ne nous souvenons même pas du visage.

D'un point de vue rationnel, ce comportement semble disproportionné. Du point de vue du système, il ne l'est pas.

Le système n'évalue pas ce cas particulier. Il réagit à un signal qui, pendant la majeure partie de son histoire, s'est révélé statistiquement fiable.

Les émotions précèdent la vérification.

Qu'en est-il du pouvoir que l'autre exerce sur nous ?

À ce stade, le lecteur attentif pourrait relever une contradiction.

Le chapitre sur la valeur sociale affirmait que le poids d'un jugement dépend du pouvoir que celui qui le formule exerce sur nous. Ce chapitre soutient que nous réagissons également aux jugements de personnes qui n'ont aucun pouvoir sur nous. Ces deux affirmations semblent incompatibles.

Elles ne le sont pas, car elles concernent deux moments différents d'un même processus.

Le pouvoir réel détermine l'importance réelle d'un jugement : il en fixe les conséquences et, à long terme, oriente effectivement le comportement — nous accordons davantage d'attention à notre image face à ceux qui peuvent changer quelque chose dans notre vie, et le système en tient compte.

Le prédicteur explique en revanche pourquoi la réaction immédiate ne procède pas à cette vérification. Au moment où le jugement est perçu, le système ne calcule pas le pouvoir de celui qui l'exprime : il réagit au signal.

Le pouvoir réel décrit donc ce à quoi le système est calibré ; le prédicteur explique pourquoi il ne le vérifie pas au cas par cas.

C'est aussi la raison pour laquelle la question posée au chapitre trois ne dissipe pas le trouble. Vérifier que cette personne ne peut rien décider à notre sujet corrige l'évaluation, mais pas la réaction — qui a déjà eu lieu.

Un monde nouveau, un mécanisme ancien

Cette interprétation permet de comprendre pourquoi le mécanisme semble souvent excessif.

Il n'est pas excessif par rapport à l'environnement dans lequel il s'est formé. Il est excessif par rapport à l'environnement dans lequel il continue de fonctionner.

Nous vivons entourés d'une multitude d'observateurs qu'aucun être humain n'aurait pu rencontrer pendant la quasi-totalité de l'histoire de l'espèce. Les réseaux sociaux, les moyens de communication, les grandes villes et la mobilité ont multiplié les regards sans augmenter, dans la même mesure, les conséquences que ces regards produisent.

Le système continue pourtant à traiter bon nombre d'entre eux comme s'ils appartenaient encore à une petite communauté stable. Il ne fait pas spontanément la distinction entre le jugement d'un ami, celui d'un futur employeur et celui d'un passant croisé pendant quelques secondes.

Non pas parce qu'il se trompe dans son calcul, mais parce qu'il ne l'effectue pas. Le système applique une heuristique — il traite le jugement comme un indicateur des conséquences sociales, car il l'a été pendant la majeure partie de son histoire — et les heuristiques fonctionnent ainsi, en remplaçant un problème difficile par un autre plus simple. La vitesse augmente ; la précision diminue lorsque l'environnement change.

La distinction entre un ami et un passant peut être établie par la réflexion, mais elle intervient après la réponse automatique, et non avant.

Le pont vers la troisième partie

Jusqu'ici, nous avons cherché à comprendre le fonctionnement du système. Nous en connaissons désormais la structure, les points forts et les limites.

Une question demeure cependant.

Si le jugement présumé est en grande partie construit de l'intérieur, si nos propres jugements finissent par façonner ce que nous croyons, et si le système continue de réagir à des signaux qui, dans le monde actuel, ont souvent perdu leur signification d'origine, est-il possible d'en modifier au moins en partie le fonctionnement ?

La troisième partie aborde cette question. Elle ne propose pas d'éliminer le mécanisme, mais de comprendre quelles interventions parviennent réellement à le recalibrer et lesquelles, au contraire, finissent par le renforcer.

TROISIÈME PARTIE — Que faire de tout cela ?

10 Pourquoi les remèdes intuitifs aggravent la situation

Après avoir décrit le fonctionnement du système, une question s'impose inévitablement.

Si le jugement présumé découle en grande partie de la perception de soi plutôt que de l'observation des autres, et si le système continue de réagir à des signaux qui, dans le monde actuel, ont souvent perdu leur signification d'origine, comment peut-on le corriger ?

La réponse spontanée semble évidente.

Mieux observer ce qui se passe. Réfléchir davantage. Se contrôler avec plus d'attention. Demander aux autres comment ils nous perçoivent réellement.

Ce sont toutes des stratégies raisonnables.

Et c'est précisément là que réside le problème.

Presque toutes détériorent le fonctionnement du système qu'elles voudraient corriger.

Le paradoxe de l'auto-observation

Lorsque nous voulons comprendre un phénomène, la solution naturelle consiste à l'observer. Si nous voulons savoir comment fonctionne un moteur, nous l'ouvrons. Si nous voulons comprendre le comportement d'un animal, nous l'observons.

Il est naturel de penser que cela doit fonctionner de la même manière avec l'esprit. S'observer davantage. Mieux surveiller ce que l'on ressent. Prêter davantage attention à ses propres réactions.

Le problème est que, dans le cas du jugement présumé, l'observateur coïncide avec ce qui devrait être observé.

Il n'existe pas de point de vue extérieur. Toute tentative de contrôler le système modifie le fonctionnement de ce dernier.

Cette difficulté porte d'ailleurs un nom. Les recherches sur l'attention portée à soi-même ont mis en évidence ce qu'on appelle le paradoxe de l'introspection : ceux qui s'observent davantage ont tendance à avoir à la fois une meilleure connaissance d'eux-mêmes et un plus grand malaise psychologique. Ces deux éléments vont de pair, et il n'est pas possible d'obtenir le premier en évitant le second simplement en s'observant encore davantage.

Plus nous nous concentrons sur nos sensations lors d'une interaction sociale, moins nous recueillons d'informations sur l'interaction elle-même.

Cela ne concerne pas uniquement les personnes souffrant d'anxiété

Avant de poursuivre, il convient de dissiper un malentendu que le reste du chapitre risquerait autrement de susciter.

On pourrait penser que ce qui suit concerne les personnes particulièrement anxieuses, et que celles qui ne le sont pas peuvent s'en considérer exemptes.

Ce n'est pas le cas, et la recherche le démontre bien avant que ce phénomène ne soit étudié dans un contexte clinique.

Shelley Duval et Robert Wicklund, avec leur théorie de la conscience de soi objective (1972), ont montré que lorsque l'attention se porte sur soi-même en tant qu'objet — devant un miroir, une caméra, un public —, l'écart entre ce que l'on est et ce que l'on estime devoir être devient saillant, et le comportement s'oriente alors vers la réduction de cet écart : vers ce que cette personne considère comme juste, et non vers ce qui est juste. L'autofocalisation ne rend pas meilleur : elle rend plus cohérent avec soi-même.

Cet effet ne concerne pas une catégorie de personnes : il concerne une situation dans laquelle n'importe qui peut se trouver.

Allen Fenigstein et ses collaborateurs (1975) ont ensuite distingué la conscience de soi en public comme un trait stable : la disposition, variable d'un individu à l'autre, à se représenter comme l'objet du regard d'autrui. Il ne s'agit pas d'une ligne de démarcation entre les personnes saines et celles qui souffrent de troubles, mais d'un continuum sur lequel chacun se situe.

Ces deux recherches, prises ensemble, apportent un éclairage précis sur le modèle présenté dans cet essai : les conditions situationnelles et les différences individuelles qui renforcent l'activité du système de jugement présumé ont déjà été observées en laboratoire, au sein de populations ordinaires.

Ce qui apparaît sous une forme extrême dans l'anxiété sociale est donc la même chose qui opère, de manière plus faible, chez ceux qui ne sont pas anxieux. C'est l'hypothèse la plus risquée de cet essai : elle soutient que le système du jugement présumé n'appartient pas à l'anxiété sociale mais à la psychologie ordinaire, et que l'anxiété sociale en est une forme particulièrement intense. Si l'on découvrait que les personnes non anxieuses utilisent des processus différents plutôt que les mêmes processus atténués, cette généralisation ne tiendrait plus.

Quand l'attention se porte au mauvais endroit

Dans le modèle de l'anxiété sociale de David Clark et Adrian Wells (1995), ce phénomène est appelé « attention autofocalisée ».

En entrant dans une situation sociale, la personne cesse progressivement d'observer son environnement et concentre son attention sur elle-même.

Elle contrôle le ton de sa voix. Elle contrôle ses rougeurs. Elle contrôle sa posture. Elle contrôle ses paroles. Elle se demande même si elle ne contrôle pas trop.

L'effet est paradoxal. Plus elle cherche des indices sur la façon dont les autres l'évaluent, moins elle recueille d'informations sur leur comportement réel.

Les informations provenant de l'extérieur diminuent. Celles provenant de l'intérieur augmentent.

Les sensations corporelles, l'anxiété et l'image que la personne a d'elle-même finissent ainsi par se substituer à l'observation des autres, et sont considérées comme des indices de la façon dont elle apparaît.

À ce stade, un lien apparaît avec la première partie de l'essai.

La « présupposition élevée » décrite au chapitre cinq et l'« attention autofocalisée » décrite par Clark et Wells peuvent être interprétées comme deux descriptions d'un même phénomène issues de traditions qui ne se citent pas mutuellement.

Dans le premier cas, le problème est décrit comme une propriété de la manière dont le jugement présumé est construit. Dans le second, comme un déplacement de l'attention qui empêche de recueillir les informations nécessaires pour le corriger.

Ces deux descriptions ne s'excluent pas mutuellement. Elles se complètent.

Deux façons différentes d'échouer

Il convient désormais de recourir à la distinction introduite au chapitre cinq, car elle permet de voir que les remèdes intuitifs n'échouent pas tous de la même manière.

Le jugement présumé reçoit deux données d'entrée, et chacune a son propre paramètre : la présupposition, qui détermine dans quelle mesure il est construit à partir de l'auto-perception plutôt que de l'observation ; et la dépréciation, qui détermine dans quelle mesure la confirmation externe est dévalorisée.

Certains remèdes échouent parce qu'ils agissent sur le mauvais paramètre. D'autres échouent parce qu'ils aggravent précisément ce qu'ils voudraient corriger.

Les sections suivantes examinent les trois remèdes les plus courants à la lumière de cette distinction.

Pourquoi réfléchir davantage ne suffit pas

Beaucoup de gens essaient de sortir du problème en y réfléchissant encore plus.

Elles repassent mentalement une conversation. Elles analysent chaque phrase. Elles reconstituent chaque détail. Elles essaient de comprendre où elles se sont trompées.

Cette activité est appelée « réflexion », mais elle englobe des processus très différents les uns des autres.

Anthony Grant et ses collaborateurs ont distingué deux notions que le langage courant confond : la réflexion, c'est-à-dire la tendance à examiner ses propres pensées et comportements, et la prise de conscience, c'est-à-dire la clarté avec laquelle on les comprend. Ces deux dimensions sont bien moins corrélées entre elles qu'on pourrait s'y attendre, et seule la seconde est associée au bien-être, à des niveaux d'anxiété plus faibles et à une meilleure estime de soi. On peut réfléchir énormément sans parvenir à aucune prise de conscience.

Dans le cas du jugement présumé, cette distinction est particulièrement importante.

Continuer à réfléchir au problème n'augmente pas automatiquement la probabilité de le résoudre, et ce pour deux raisons distinctes.

La première est que la réflexion porte sur le matériel interne, et ce matériel interne est précisément ce que le chapitre cinq a montré être la source du problème. Repenser une conversation n'apporte aucune nouvelle information sur les autres : cela ne fait que retravailler ce que le système avait déjà produit. En termes des deux paramètres, cela ne réduit pas la présupposition. Au contraire, cela l'augmente.

La deuxième raison est moins évidente, et découle d'une observation que j'ai développée ailleurs. Le système ne communique pas ses résultats sous forme de pensées, mais sous forme d'émotions : une impulsion à agir lorsque la prévision est favorable, une résistance lorsqu'elle est défavorable. Ce qui parvient à la conscience, c'est le résultat, pas le calcul. Si tel est le cas, la réflexion interroge un niveau où l'information la réponse ne se trouve pas — et les explications qu'elle produit sont des reconstructions plausibles a posteriori, et non le récit de ce qui s'est réellement passé.

Cela ne rend pas la réflexion inutile. Cela rend inutile d'espérer qu'elle produise à elle seule le changement.

Pourquoi les paroles rassurantes ne fonctionnent pas

Une deuxième stratégie intuitive consiste à demander des confirmations.

« À ton avis, ai-je été convaincant ? » « Penses-tu que j'ai fait mauvaise impression ? » « Comment m'as-tu perçu ? »

L'idée est simple : si le système dispose d'informations erronées, il suffit de lui fournir les informations correctes.

Le problème, c'est que le système n'est pas dépourvu d'informations. Il les dévalorise.

Ici, le paramètre en jeu n'est pas la présupposition, mais la remise en perspective. La réassurance apporte exactement ce qui semble manquer — une information externe, explicite, provenant de la bonne personne — et ne produit pas l'effet escompté, car cette information est reçue et relativisée.

« Il dit ça par gentillesse. » « Il ne veut pas me blesser. » « Il n'a pas vu ce qui s'est réellement passé. »

L'information arrive. Elle ne modifie pas le système.

C'est le cas le plus instructif des trois, car il montre que le problème ne réside pas toujours dans la quantité d'informations disponibles. En ajouter davantage, lorsque le paramètre critique est la remise en question, ne sert à rien — et demander des confirmations à plusieurs reprises tend, pour des raisons que nous verrons, à renforcer justement l'habitude de les dévaloriser.

Pourquoi se contrôler nuit à la performance

Il existe enfin une stratégie encore plus intuitive. Mieux se contrôler.

Préparer chaque phrase. Surveiller en permanence son propre comportement. Éliminer toute erreur possible.

À court terme, cela peut même sembler efficace. À long terme, cela produit souvent l'effet inverse.

Chaque vérification augmente la quantité d'attention portée sur soi-même et réduit celle disponible pour l'environnement.

Le système perd justement les données dont il aurait besoin pour se mettre à jour. En termes des deux paramètres : l'autocontrôle augmente l'hypothèse, car il prive le système de la seule entrée qui pourrait le corriger.

C'est un cercle vicieux. Plus on cherche à vérifier comment on est évalué, moins on observe comment les autres réagissent réellement. Plus on se fie à ses sensations internes, plus le jugement présumé continue d'être construit de l'intérieur.

Il convient de noter que c'est également la raison pour laquelle les comportements protecteurs — préparer ses phrases, éviter les silences, tenir son verre pour cacher ses tremblements — figurent parmi les principales cibles de la thérapie de l'anxiété sociale. Non pas parce qu'ils sont inefficaces sur le moment, mais parce qu'ils empêchent de découvrir ce qui se serait passé sans eux.

Une conclusion provisoire

Les remèdes intuitifs ont une caractéristique commune.

Ils cherchent à améliorer le système en renforçant la surveillance. Or, au-delà d'un certain seuil, c'est précisément cette surveillance qui le maintient en place.

Considérés à travers ces deux paramètres, cependant, ils n'échouent pas tous de la même manière. Réfléchir et se contrôler renforcent la présupposition : ils privent le système d'informations externes et le contraignent à se nourrir encore davantage de lui-même. Demander des assurances ne touche pas à la présupposition — l'information externe la fournit bel et bien — mais se heurte à la dévalorisation.

Cette distinction n'est pas purement théorique. Elle indique qu'il existe au moins deux interventions utiles, et différentes l'une de l'autre : l'une qui restitue au système une information réelle, l'autre qui empêche sa dévalorisation.

Il vaut la peine de mettre en relation ce résultat avec celui du chapitre précédent, car les deux limites se cumulent.

Nous avons vu là-bas que le fait de constater rationnellement qu'une certaine personne ne peut rien décider à notre sujet corrige l'évaluation, mais n'empêche pas la réaction : lorsque cette constatation survient, la réaction a déjà eu lieu. Ici s'ajoute une deuxième limite : bon nombre des tentatives spontanées visant à corriger cette réaction finissent par alimenter le mécanisme qui la produit.

La raison pour laquelle le chapitre suivant est nécessaire réside dans cette somme. Il ne suffit pas de savoir comment fonctionne le système, et il ne suffit pas de s'efforcer de le corriger avec les moyens qui viennent naturellement.

Cela ne signifie pas que le système soit immuable. Cela signifie que la direction du changement n'est pas celle que l'on pourrait intuitivement imaginer.

C'est le thème du chapitre suivant.

11 Le principe du recalibrage

À la fin du chapitre précédent, nous sommes parvenus à une conclusion qui peut sembler décourageante.

Le jugement présumé continue de nous influencer même lorsque nous savons que le jugement des autres n'aura pas de conséquences réelles. Et les tentatives les plus spontanées pour y remédier finissent souvent par le renforcer.

On pourrait penser que le système est tout simplement immuable.

Ce n'est pas là la conclusion.

Le système ne change pas grâce au raisonnement par lequel nous essayons de le convaincre. Il change lorsque les prévisions qu'il utilise changent.

Un système qui apprend

Il est temps de donner un nom au système dont nous avons suivi le fonctionnement. Dans mes travaux précédents, je l'ai appelé « *superviseur social* » ; dans ces pages, j'ai préféré le décrire avant de le nommer, car une étiquette introduite trop tôt prend le pas sur la description. À partir de maintenant, ce nom est nécessaire, car la troisième partie traite de la manière de le modifier.

Appelons ainsi le système qui a été décrit pièce par pièce dans ces pages : celui qui construit le jugement présumé en puisant surtout dans la perception de soi, qui continue à fonctionner de manière atténuée lorsque personne ne nous observe, et qui fait parvenir à la conscience ses propres résultats — et non le calcul qui les produit — sous forme d'impulsions et de résistances émotionnelles. C'est un système de prévision.

Sa mission n'est pas de dire si quelque chose est juste ou faux. Elle consiste à anticiper les conséquences relationnelles de nos actions.

Comme tout système prédictif, il doit confronter ses attentes à ce qui se passe réellement. S'il ne le faisait pas, il ne pourrait pas s'adapter à un environnement en mutation.

Le superviseur social n'est donc pas rigide. Il est malléable.

Mais sa plasticité obéit à des règles différentes de celles de la réflexion consciente.

Les mots expliquent, l'expérience actualise

Si quelqu'un nous dit « tu ne dois pas te soucier du jugement des autres », nous comprenons parfaitement le sens de cette phrase. Nous pouvons même y adhérer.

Et pourtant, le système continue de réagir comme avant.

Ce n'est pas parce qu'il est irrationnel. C'est parce que cette phrase ne constitue pas une évidence. C'est une information.

Le superviseur social modifie ses prévisions surtout lorsque l'expérience s'écarte de manière significative de ce à quoi il s'attendait.

Il ne suffit pas de connaître une conclusion. Il faut que la prévision soit démentie.

L'erreur de prévision

De nombreux modèles contemporains d'apprentissage, bien que différents les uns des autres et ne pouvant être rattachés à une théorie unique, partagent l'idée que les prévisions sont mises à jour surtout lorsque l'expérience produit un résultat inattendu. Dans la suite, j'adopterai ce principe comme cadre d'interprétation, sans supposer qu'il décrive de manière exhaustive toutes les formes d'apprentissage.

Ce qui suit ne dépend d'ailleurs pas d'une formalisation particulière : cela exige simplement que le système actualise ses attentes plus rapidement face à une réfutation que face à une confirmation.

Un système formule une attente. L'expérience montre un résultat différent. L'écart entraîne une mise à jour.

Le principe est étonnamment simple. Aucun système ne modifie ses prévisions lorsque tout confirme ce qu'il croit déjà : il les modifie lorsque la réalité réserve une surprise.

Pour le superviseur social, cela signifie que le changement nécessite des situations dans lesquelles le jugement présumé est réellement mis à l'épreuve.

Pas imaginé. Pas discuté. Vécu.

Il s'ensuit que toute tentative de modifier le superviseur social n'a de sens que si elle crée des situations dans lesquelles l'une de ses prévisions peut être confrontée à ce qui se passe réellement. Si les deux coïncident, le système n'a aucune raison de changer ; s'ils divergent, l'information dont il avait besoin apparaît.

Pourquoi il ne suffit pas de s'exposer

Cette idée permet de comprendre un malentendu fréquent.

On entend souvent dire qu'il suffit de s'exposer aux situations redoutées.

L'exposition, cependant, ne garantit en soi aucun changement.

Prenons l'exemple d'une personne convaincue que, si elle prenait la parole lors d'une réunion, les autres la trouveraient incompétente. Elle parle, et pendant toute la durée de son intervention, elle observe sa voix, ses mains, ses hésitations. À la fin, elle se souvient surtout de son malaise, elle ne se souvient pas des visages de ceux qui l'écoutaient, et elle en conclut que cela s'est mal passé.

Elle a traversé la situation redoutée sans recueillir la moindre information nouvelle. Elle a simplement accumulé une nouvelle expérience cohérente avec son modèle antérieur.

Pour que l'exposition à une situation donne lieu à un apprentissage, le superviseur social doit pouvoir comparer une prévision suffisamment précise à un résultat observable.

Ce n'est pas le fait de rester dans la situation qui produit le changement. C'est la comparaison entre ce à quoi on s'attendait et ce qui s'est réellement passé.

Deux conditions nécessaires

Deux conditions se dégagent de ce qui précède, et elles correspondent aux deux paramètres présentés au chapitre cinq : la présupposition et la remise en question.

La première condition concerne la présupposition : le résultat doit être observé. Si l'attention reste tournée vers l'intérieur, le système continuera à se nourrir de ses propres sensations, et aucune expérience ne pourra le contredire car aucune expérience n'aura été réellement enregistrée.

La deuxième condition concerne l'anticipation : la prédiction doit être formulée à l'avance et sous une forme reconnaissable. Si elle reste vague, après l'expérience, il sera toujours possible de réinterpréter ce qui s'est passé — « ça s'est bien passé parce que c'était une situation facile », « ils ont été gentils » — et la réfutation sera absorbée sans laisser de trace.

C'est la raison pour laquelle, dans le traitement de l'anxiété sociale, les expériences comportementales demandent d'écrire la prédiction au préalable. Il ne s'agit pas d'un formalisme : une prédiction consignée par écrit ne peut plus être réécrite a posteriori.

Nous pouvons désormais énoncer ce principe sous sa forme la plus simple, avant d'y ajouter les précisions qu'il requiert : le superviseur social ne change pas parce que nous le convainquons, mais parce que l'expérience contredit de manière vérifiable ses prévisions.

Tout ce qui suit sert à préciser cette phrase dans trois directions : ce qu'il faut pour qu'une réfutation soit enregistrée, où interviennent les réfutations les plus efficaces, et dans quelle direction le système s'adapte le plus facilement.

Pour qu'une réfutation soit enregistrée

Il y a cependant un aspect que la notion d'erreur de prévision laisse de côté, et que j'ai abordé ailleurs en traitant de l'attention.

Pour qu'une prévision puisse être réfutée, il faut se rendre compte qu'elle est à l'œuvre. Et la reconnaissance, à elle seule, n'entraîne pas de changement.

Ce que la prise de conscience produit, c'est une brève fenêtre : un intervalle pendant lequel la réponse habituelle n'a pas encore été déclenchée et où une réponse différente est encore possible.

Pour que cette fenêtre soit exploitée, il faut avoir déjà décidé auparavant ce qu'on en fera. Reconnaître sans avoir d'alternative toute prête produit des spectateurs conscients de leurs automatismes, et non des personnes qui les modifient.

C'est aussi la raison pour laquelle la prise de conscience purement intellectuelle ne change souvent rien : on peut tout savoir de ses propres mécanismes et continuer à les reproduire, car savoir n'équivaut pas à avoir une réponse différente toute prête au moment où elle serait nécessaire.

Il en découle une distinction qui concerne précisément le moment de la prise de conscience, et qu'il vaut la peine de rendre explicite, car deux choses apparemment similaires tirent dans des directions opposées.

La première consiste à maintenir le sujet vivant : continuer à observer son propre fonctionnement sous différents angles, l'empêcher de sombrer dans l'insignifiance. Ici, la variété est précieuse, car l'attention s'éteint face à ce qui se répète à l'identique.

La seconde consiste à reconnaître le mécanisme au moment où il se met en marche. Ici, la variété est néfaste, et c'est le contraire qui est nécessaire : des signaux spécifiques et constants, ancrés dans des circonstances concrètes qui précèdent ou accompagnent la réaction. Une sensation physique, une situation récurrente, un type de pensée.

Les recherches sur la mémoire prospective — celle qui consiste à se souvenir de faire quelque chose au bon moment — indiquent que la reconnaissance en situation fonctionne mieux lorsque l'association entre le signal et la réponse est répétée de manière identique, et non lorsqu'elle est renouvelée.

C'est en confondant ces deux fonctions que de nombreuses résolutions échouent. Un rappel lu le matin arrive alors que le mécanisme n'est pas actif ; lorsqu'il s'activera, en plein cœur d'une situation, ce rappel ne sera plus présent.

Dans ce contexte, la conscience ne met pas le système à jour : elle planifie les conditions dans lesquelles la mise à jour peut avoir lieu, et décide à l'avance quoi faire de la fenêtre que la reconnaissance ouvre.

Où interviennent les réfutations les plus efficaces

Jusqu'ici, nous avons parlé d'interventions délibérées. Ce serait toutefois une erreur de conclure qu'elles sont les plus efficaces.

Ce n'est probablement pas le cas.

Les recalibrages les plus profonds semblent se produire par des voies que personne n'a prévues : un changement de position sociale, l'entrée dans un environnement aux critères différents, le vieillissement, une relation stable qui dure suffisamment longtemps pour produire des milliers de petits démentis.

Nous disposons d'ailleurs d'un indice empirique à ce sujet. Au chapitre 5, nous avons vu que la capacité à reconnaître quand notre estimation est juste est liée à la qualité de la relation avec la personne en question.

La raison en est plausible. Dans une relation étroite et durable, le retour d'information ne se présente pas en une seule occasion réévaluable, mais à de nombreuses reprises qu'il serait fastidieux de réinterpréter toutes de la même manière. La « décote » fonctionne bien sur un épisode isolé. Elle fonctionne mal sur une période de cinq ans.

Cela a une conséquence gênante pour quiconque recherche des méthodes rapides : le facteur qui bénéficie du plus grand soutien empirique n'est pas une technique, et il échappe en grande partie au contrôle de ceux qui souhaiteraient l'appliquer.

Dans quelle direction le système s'adapte-t-il le plus facilement ?

Il y a enfin un aspect qu'une analyse consacrée au changement a tendance à laisser de côté, et qu'il faut mentionner, car sinon le tableau serait incomplet.

Le superviseur social ne se recalibre pas uniquement dans le sens positif.

Une humiliation publique, une trahison, une période prolongée d'hostilité dans un environnement de travail entraînent des ajustements rapides et stables, dans la direction opposée.

Et cela se fait selon les mêmes mécanismes que ceux décrits dans ce chapitre : ce sont des expériences qui contredisent une prévision — la prévision d'être en sécurité, d'être apprécié, de pouvoir compter sur quelqu'un — avec une précision qu'aucun exercice ne parvient à égaler.

Il convient de noter l'asymétrie. La démenti dans le sens d'une détérioration survient généralement lors d'un épisode unique, intense et sans ambiguïté. La démenti dans le sens d'une amélioration survient presque toujours à l'occasion de nombreux petits événements facilement réinterprétables.

Il reste à se demander pourquoi ces deux démentis prennent des formes si différentes, et la réponse a déjà été donnée il y a deux chapitres. Si le fait d'exprimer un jugement expose celui qui l'émet, une réticence réciproque s'installe entre deux personnes qui se trouvent face à face : le jugement favorable n'est presque jamais déclaré ouvertement, et se présente sous la forme d'une courtoisie. Mais la courtoisie est précisément ce contre quoi le système se défend — « il le dit par gentillesse », « il ne veut pas me blesser » — de sorte qu'une information positive est d'emblée considérée comme prévisible, car la convention sociale rend sa source ambiguë.

L'humiliation publique, en revanche, rompt cette réticence : quelqu'un a payé le prix de s'exposer pour dire ce qu'il pense, et c'est pourquoi le signal est clair et se prête mal à une autre interprétation. En bref : personne ne nous dit clairement que tout va bien, alors qu'on nous le dit clairement quand ça va mal.

Il n'est donc pas vrai que le système soit réticent à évoluer. Il est réticent à évoluer vers le mieux.

Un critère, pas une méthode

Le principe du recalibrage peut se résumer en une phrase.

Le superviseur social évolue moins à travers les explications qu'il reçoit qu'à travers les erreurs de prévision qu'il subit.

Cette affirmation ne constitue pas encore une méthode. Elle définit toutefois le critère permettant d'évaluer toute proposition de changement.

Toute intervention corrective qui ne donne pas lieu à une confrontation entre une prévision du système et l'expérience aura du mal à en modifier le fonctionnement. Toute intervention corrective qui parvient à produire cette confrontation devient, du moins en principe, un candidat au recalibrage.

Traduit en termes opérationnels, ce critère s'articule autour de deux questions que l'on peut poser à toute proposition de changement, y compris la sienne.

Ce ne sont pas les deux questions que l'essai propose au lecteur, et dont traite le chapitre suivant : celles-ci s'adressent à soi-même face à un jugement, tandis que celles-ci concernent une méthode que l'on évalue.

- L'intervention augmente-t-elle la quantité d'informations externes qui parviennent au superviseur social ?
- L'intervention réduit-elle la tendance du superviseur à dévaloriser cette information ?

Il convient de les tester sur des cas dont le lecteur dispose déjà. Les trois remèdes du chapitre précédent constituent autant d'interventions, et le tableau reproduit ici les verdicts déjà rendus.

Réfléchir davantage n'ajoute pas d'informations externes — cela ne fait que retravailler des éléments internes — et ne réduit pas la décote : la réponse est « non » aux deux questions. Se contrôler aggrave la situation, car cela soustrait les informations qui auraient pu être disponibles. Demander des assurances, en revanche, apporte réellement des informations externes, et l'on s'arrête à la deuxième question : ces informations sont reçues et remises en perspective.

C'est le cas qui montre comment les deux réponses doivent être lues conjointement, et c'est le sujet du paragraphe suivant.

Si la réponse aux deux questions est négative, il est peu probable que la méthode produise un réajustement stable. Si les deux réponses sont positives, il existe une raison théorique de s'attendre à un effet. Si seule la première réponse est positive, le résultat dépend de l'importance de la décote : lorsqu'elle est élevée, l'information arrive mais n'est pas enregistrée, et c'est précisément ce qui se passe avec les assurances.

Ces deux critères ne garantissent pas qu'une intervention fonctionne : ils distinguent simplement celles qui sont compatibles avec le modèle de celles qui, bien que plausibles, ne pourraient guère en modifier le fonctionnement.

Où en est la méthode ?

Le fait que ce chapitre se limite à un seul critère ne signifie pas que la méthode n'existe pas.

Dans un ouvrage distinct, **Recalibrer le superviseur social**, j'ai proposé une démarche opérationnelle fondée sur ces mêmes prémisses : reconnaître la prédiction au moment où elle apparaît, la rendre explicite, reconstituer les éléments qui l'ont produite, envisager des scénarios alternatifs et comparer la prédiction notée à l'avance avec ce qui se passe réellement. Elle comprend un journal personnel, un catalogue des erreurs les plus fréquentes et les critères permettant d'interrompre le travail.

J'ai préféré ne pas l'inclure dans ces pages pour une raison de nature, et non d'espace. Un essai décrivant un mécanisme et un parcours d'autothérapie répondent à des questions différentes, s'adressent à des lecteurs dans des situations différentes et doivent être jugés selon des critères différents : le premier sur la fidélité aux données, le second sur les effets qu'il produit chez ceux qui l'appliquent.

Les confondre aurait affaibli les deux. Le chapitre qui suit propose donc deux outils minimaux, que tout le monde peut utiliser après avoir lu ces pages ; ceux qui souhaiteraient un travail plus approfondi le trouveront ailleurs.

Le chapitre suivant met ces critères à l'épreuve à travers les deux questions que cet essai pose au lecteur.

12 Deux questions fondamentales

Cet essai a promis, dès l'introduction, de converger vers deux questions fondamentales.

Le moment est venu de les formuler et de les soumettre au critère établi dans le chapitre précédent.

Le résultat, il faut le dire d'emblée, n'est pas uniforme. Les deux questions ne remplissent pas la même fonction, et l'une d'entre elles n'est pas un outil de recalibrage.

Les questions rassemblées dans l'annexe « *Questions pour approfondir* » ont un statut différent : ce sont des aides à consulter à tête reposée, et elles servent par la suite, lorsqu'il reste à décider quoi faire d'un jugement déjà formé.

Observations, déductions, prévisions

Avant de les formuler, il convient d'opérer une distinction que ces deux questions présupposent, et que nous perdons sans cesse de vue dans la vie quotidienne.

Les événements observables sont ceux qui peuvent être perçus ou documentés : une personne prononce une phrase, sourit, sort précipitamment, ne répond pas à un message. Les décrire ne nécessite pas de leur attribuer une signification.

En revanche, chaque fois que nous attribuons à quelqu'un une intention, une émotion ou un jugement, nous faisons une déduction. Et lorsque nous étendons cette déduction à l'avenir, nous formulons une prévision.

La différence apparaît plus clairement lorsqu'on les met en parallèle :

« Il m'a salué rapidement » est une observation.

« Il était agacé par moi » est une déduction.

« Il ne m'invitera plus » est une prédiction.

Dans le langage courant, nous passons de l'une à l'autre sans nous en rendre compte, et nous traitons la troisième affirmation avec la même certitude que la première. Les déductions sont indispensables — sans elles, la vie sociale serait impossible — et le problème n'est pas de déduire : c'est d'oublier que nous l'avons fait.

La même distinction, appliquée au sujet traité dans cet essai, distingue trois choses. Le jugement exprimé est une observation : quelqu'un dit quelque chose à notre sujet, et on l'a entendu. Le jugement réel — ce que cette personne pense vraiment — est une déduction : il n'est pas observable, et ne coïncide pas nécessairement avec celui qui est exprimé, lequel peut être réticent, de courtoisie, ironique ou faux. Le jugement présumé est une prédiction, et il se trouve dans un autre esprit : non pas dans celui de celui qui nous juge, mais dans le nôtre.

Confondre une prédiction avec une observation revient à attribuer à ce que notre esprit a produit la valeur d'un fait avéré. Les deux questions qui suivent servent à éviter cela.

La première question

Cette personne peut-elle changer quelque chose dans ma vie ?

Cette question découle du chapitre sur la valeur sociale. Si le poids d'un jugement dépend du pouvoir que celui qui le formule exerce sur nous, alors face à un jugement qui nous trouble, il existe un moyen de vérifier : déterminer si cette personne peut changer quelque chose dans notre vie.

Dans la plupart des cas, la réponse est non. Le commentaire d'un inconnu, l'avis de quelqu'un que nous ne reverrons pas, l'opinion de quelqu'un qui n'a aucune influence sur notre vie.

Ce n'est pas une question de consolation. Elle ne demande pas de minimiser, ni de se convaincre que le jugement n'a pas d'importance : elle demande une vérification, et la réponse peut très bien être oui. Quand c'est le cas, l'inquiétude est justifiée et le système fonctionne comme il se doit.

Ce que fait la première question, et ce qu'elle ne fait pas

Le critère du chapitre précédent sert à déterminer si une intervention peut recalibrer le système. Appliqué à cette question, la réponse est non — et il convient de le préciser avant d'expliquer à quoi cela sert.

Le chapitre sur le dysfonctionnement du mécanisme en a déjà expliqué la raison. La réaction précède la vérification. Lorsque la vérification intervient, le système a déjà produit sa prévision et l'organisme a déjà commencé à réagir.

La vérification corrige l'évaluation, pas la réaction.

Et au regard des deux critères opérationnels, elle ne s'en sort pas très bien. Elle n'augmente pas l'information externe : elle n'apporte aucune donnée sur la façon dont cette personne nous perçoit réellement. Elle ne réduit l'« escompte » que partiellement, et de manière indirecte : elle rend moins automatique le fait d'attribuer au jugement un poids qu'il n'a pas.

Il serait donc erroné de la présenter comme une technique capable de faire cesser le malaise. Elle ne le fait pas, et le modèle prévoit qu'elle ne peut pas le faire, du moins dans l'immédiat.

Sa fonction est autre, et il convient de la définir avec précision plutôt que de la juger à l'aune d'un critère qui ne relève pas de son domaine. Elle ne sert pas à modifier le superviseur social : elle sert à empêcher que la réaction initiale ne soit amplifiée par le traitement ultérieur — les conséquences imaginées, les scénarios, le ruminement des jours suivants, qui sont presque toujours plus longs et plus coûteux que la réaction elle-même. C'est un outil de maîtrise, pas de correction. Il n'empêche pas le premier coup. Il empêche la suite.

La deuxième question

Que penserais-je si mon jugement n'était observable par personne ?

Cette question concerne l'autre aspect : non pas comment nous sommes jugés, mais comment nous jugeons.

Elle découle du chapitre sur la récursivité et de celui sur les jugements en tant que signaux. Si la manière dont nous évaluons quelque chose est elle-même observable, et si ce que nous disons tend à devenir ce que nous pensons, alors certains de nos jugements sont tels que nous ne les aurions pas formulés dans d'autres circonstances.

La question invite à distinguer ce que nous évaluons de ce qu'il nous convient d'évaluer.

Pourquoi on n'y répond pas en se tournant vers son for intérieur

Il existe cependant une difficulté que l'essai lui-même a créée, et qu'il serait incorrect de contourner.

Le chapitre sur l'adaptation à l'auditeur a montré que les jugements de convenance ne restent pas reconnaissables comme tels. Ils deviennent sincères. Si tel est le cas, la question ne peut être résolue par l'introspection : en nous regardant en nous-mêmes, nous trouverons un jugement qui nous semble authentique, et c'est exactement ce à quoi on devrait s'attendre dans les deux cas.

Posée comme un examen de conscience, la question ne mène à rien.

Elle peut produire quelque chose si elle est posée comme une comparaison entre des situations, c'est-à-dire comme une prédiction à vérifier plutôt que comme une introspection. Que penses-tu de ce livre, de cette personne, de cette décision, lorsque tu en parles avec quelqu'un qui pense différemment de toi ? Et lorsque tu en parles avec quelqu'un qui pense comme toi ? Les deux réponses coïncident-elles ?

Formulée ainsi, la question répond au premier critère : elle produit une information qui ne provient pas de la perception de soi. Il s'agit d'une comparaison entre deux comportements observables, et non entre deux intuitions.

Quant au deuxième critère, l'effet est plus incertain. La question ne réduit l'écart que si la divergence entre les deux situations est suffisamment nette pour ne pas pouvoir être résorbée — et il est facile de la résorber : « je me suis exprimé différemment », « j'ai été plus prudent par politesse », « c'était un contexte différent ». Ce sont des explications presque toujours disponibles, et presque toujours en partie vraies. C'est la raison pour laquelle il vaut mieux effectuer la comparaison sur des cas où la différence est flagrante, et non sur des nuances.

Ce que fait la deuxième question, et ce qu'elle ne fait pas

Pour cette question également, il convient de préciser ce qu'elle ne promet pas.

Elle ne permet pas de déterminer quel est le jugement vrai. C'est la conséquence la plus directe de ce que nous avons vu concernant l'adaptation à l'auditeur : si le jugement modulé tend à devenir sincère, il n'existe nulle part un jugement original intact à retrouver. Constaté que son opinion varie en fonction de son interlocuteur ne signifie pas qu'une des deux versions soit authentique et l'autre feinte. Cela signifie que le jugement est sensible au contexte — ce qui est une information différente, et plus utile.

Cela n'élimine pas le conformisme, et ne le réduit pas automatiquement. Savoir qu'une opinion s'est formée en fonction de l'environnement ne la rend pas fausse, ni n'oblige à la changer : de nombreuses convictions héritées de son environnement sont d'excellentes convictions.

Cela ne garantit pas l'authenticité, et ce serait suspect si c'était le cas. Un essai qui soutient l'inexistence d'un point d'observation extérieur au système ne peut pas ensuite offrir un outil pour y parvenir.

Ce qu'il fait, c'est rendre observable quelque chose qui, normalement, ne laisse aucune trace : l'influence du contexte social sur nos jugements. Il ne précise pas l'ampleur de cette influence en général, mais la met en évidence dans un cas concret, à travers deux comportements que l'on peut comparer.

C'est moins que ce que le lecteur pourrait souhaiter. C'est aussi la seule chose que le modèle autorise à promettre.

Pourquoi faut-il les ancrer ?

Il reste le problème le plus pratique, et c'est celui dont le chapitre précédent a déjà indiqué la solution.

Considérées à tête reposée, ces deux questions relèvent de ce que nous avons appelé l'entraînement : elles maintiennent le sujet à l'esprit, sans rien modifier au moment où elles seraient utiles. Le moment où elles seraient utiles est un autre — lorsque le trouble commence, lorsque l'on se rend compte que l'on est en train de formuler un jugement devant quelqu'un.

Si elles doivent fonctionner comme des déclencheurs, ce que nous avons vu dans le chapitre précédent indique une condition : qu'elles soient associées à un signal spécifique et constant, et que cette association soit toujours répétée de la même manière plutôt que renouvelée.

Le signal ne doit pas être cherché dans la pensée, mais là où le système se manifeste en premier lieu : dans l'essai sur l'autogouvernance sociale, j'ai soutenu que ce mécanisme ne communique pas par le biais de contenus mentaux, mais par le biais de pulsions et de résistances. Ce sont elles qui sont reconnaissables — le nœud à l'estomac, l'impulsion de répondre, l'hésitation avant de parler, le besoin soudain de s'expliquer.

Et il faut la dernière condition, la plus facile à oublier : avoir décidé au préalable quoi faire de la fenêtre que la reconnaissance ouvre. Reconnaître sans alternative toute prête produit des spectateurs conscients de leurs propres automatismes.

Dans le cas de la première question, l'alternative peut être minimale : répondre à la constatation et s'arrêter là, sans poursuivre la construction de scénarios. Dans le cas de la seconde, noter le jugement avant de l'exprimer — car un jugement mis par écrit ne peut plus être réécrit par la suite, et c'est la seule façon de se rendre compte à quel point on a changé d'avis.

Si une prévision revient sans cesse

Il peut arriver qu'une prévision se répète toujours de la même manière, dans des situations différentes, sans que la vérification ne l'ébranle. Dans ce cas, ces deux questions ne suffisent pas, et il est utile de savoir ce que l'on pourrait faire de plus.

Une prévision ne naît pas de rien : elle se construit à partir de quelque chose. On peut essayer de la reconstituer : quels faits observables l'ont fait naître, quelles inférences y ont été ajoutées, quelle part provient de la situation présente et quelle part d'expériences antérieures, quelles informations

disponibles ont été négligées. Cette opération ne permet pas de déterminer si la prédiction est juste ou fausse : elle rend visible le processus qui l'a produite, et une prédiction dont on entrevoit la construction ressemble déjà moins à un fait.

C'est un travail plus long qu'une simple question, et qui nécessite de prendre des notes. Ceux qui souhaiteraient s'y atteler de manière systématique trouveront dans l'essai **Recalibrer le superviseur social** un parcours conçu à cet effet.

Ce qu'elles promettent réellement

Il vaut la peine de tirer le bilan sans l'embellir.

Aucune de ces deux questions n'empêche le système de réagir. Aucune des deux ne le recalibre à elle seule. Par rapport aux voies décrites dans le chapitre précédent — les relations prolongées, les changements d' t d'environnement, le temps —, ce sont de petits outils, qui sont de surcroît sujets à la dégradation dont nous parlerons plus loin.

Ce qu'elles font, c'est rendre observable un mécanisme qui opère à l'abri de sa propre inspection. Ce n'est pas rien : c'est la condition pour que quoi que ce soit d'autre puisse se produire. Mais ce n'est pas la même chose que de le résoudre, et les présenter autrement trahirait le critère que cet essai s'est fixé.

Et ils le font en déplaçant l'attention vers l'extérieur, et non vers l'intérieur : ils servent à interrompre l'introspection, et non à l'affiner.

Ceux qui les utilisent avec une certaine régularité devraient s'attendre à deux effets modestes et vérifiables : moins de temps passé à reconstituer des scénarios après un jugement reçu, et quelques cas où l'on se rend compte d'avoir changé d'avis sans s'en rendre compte.

Si rien de tel ne se produit, ces outils ne produisent pas les effets prévus par le modèle. Il se peut qu'ils n'aient pas été utilisés comme décrit ici ; il se peut que le modèle soit erroné. La deuxième possibilité doit être envisagée au même titre que la première. S'il n'en était pas ainsi, cet essai demanderait de la confiance plutôt qu'une vérification.

13 Le déclin du recalibrage

Si le modèle proposé dans cet essai est correct, il faut en accepter une conséquence désagréable.

Aucun recalibrage n'est définitif.

Il serait rassurant de penser qu'une fois modifiée, une prévision du superviseur social reste stable. Mais un système dont la tâche est de s'adapter à l'environnement ne peut pas fonctionner ainsi : il continuerait à utiliser des prévisions conçues pour un monde qui pourrait ne plus exister.

La dégradation ne se présente toutefois pas sous une seule forme. Il en existe deux, qui agissent de manière différente et se cumulent.

La première forme de dégradation : le système continue d'apprendre

Le superviseur social ne cesse jamais de confronter ses attentes à l'expérience.

Chaque nouvelle relation, chaque changement d'environnement, chaque modification de sa position sociale génère de nouvelles informations. Certaines confirment les prévisions antérieures. D'autres les modifient.

Le réajustement n'est donc pas un événement ponctuel, mais un processus continu.

Le même mécanisme qui permet de corriger une prévision erronée peut, par la suite, en construire une autre tout aussi erronée.

Pourquoi on retombe facilement

Beaucoup de personnes décrivent la même expérience.

Pendant un certain temps, elles parviennent à s'exprimer plus librement, à s'ouvrir sans gêne particulière, à accorder moins d'importance au jugement des autres.

Puis quelque chose change. Un incident désagréable. Un nouvel environnement. Une personne particulièrement critique.

Et la gêne semble revenir.

Ce phénomène ne prouve pas que le changement précédent était illusoire. Il peut simplement indiquer que le superviseur social a continué à faire ce pour quoi il existe : actualiser ses prévisions.

Le problème n'est pas qu'il ait recommencé à apprendre. Le problème, c'est ce qu'il a appris.

Il convient de rappeler, à cet égard, l'asymétrie décrite dans le chapitre consacré au principe de recalibrage. La réfutation dans le sens d'une détérioration se produit généralement lors d'un épisode unique, intense et sans ambiguïté ; celle dans le sens d'une amélioration se produit à de nombreuses occasions mineures et facilement réinterprétables. Il s'ensuit qu'un seul épisode peut rapidement défaire ce que de nombreuses expériences avaient lentement construit — non pas parce que le système est fragile, mais parce que les deux directions ne l'atteignent pas avec la même efficacité.

L'importance de l'environnement

Cette observation redonne de l'importance à un thème déjà abordé.

Dans le chapitre consacré au principe de recalibrage, nous avons vu que les changements les plus profonds semblent se produire au sein de relations durables et d'environnements relativement stables.

La raison en devient désormais plus claire.

Un environnement stable ne se contente pas de modifier les prévisions : il ne cesse également de les confirmer. Chaque confirmation successive rend moins probable qu'un seul épisode parvienne à les renverser.

Un environnement imprévisible, au contraire, oblige le superviseur à s'adapter en permanence, et aucune prévision n'a le temps de se consolider.

La deuxième forme de dégradation : on s'habitue aux outils

Il existe cependant une deuxième manière dont le recalibrage s'use, et elle concerne directement ce que cet essai a proposé au lecteur dans le chapitre précédent.

L'attention volontaire est une ressource limitée, et elle s'éteint lorsque le stimulus ne change pas. C'est un principe général du fonctionnement mental : un bruit constant cesse d'être perçu, le contact des vêtements avec la peau disparaît de la conscience, un panneau vu tous les jours devient invisible.

Il en va de même pour les mécanismes psychologiques que nous nous proposons d'observer, ainsi que pour les outils avec lesquels nous nous proposons de les observer.

Il en résulte un paradoxe que j'ai abordé ailleurs dans le cadre de mes réflexions sur l'attention. Pour modifier un automatisme, il faut l'observer ; mais c'est précisément parce qu'il est répétitif qu'il tend à échapper à l'attention. Et il en va de même pour la question qui devrait nous aider à le saisir : au bout de quelques semaines, elle cesse d'être remarquée.

Deux précisions rendent ce phénomène plus insidieux qu'il n'y paraît.

La première est que l'accoutumance ne concerne pas seulement le stimulus individuel, mais sa catégorie. On ne s'habitue pas seulement à une formule répétée : on s'habitue à « la question que je me pose quand quelqu'un me juge », même si elle est formulée différemment à chaque fois.

La seconde est que l'accoutumance n'est pas passive. Face à une répétition excessive, l'esprit ne se contente pas de désactiver la perception : il réagit par l'ennui, qui pousse activement à chercher autre chose. C'est pourquoi les pratiques d'auto-observation ne sont pas seulement oubliées : elles sont abandonnées. D'abord, l'outil perd de son intensité, puis il commence à repousser.

Il convient donc de dire sans détours à quoi le lecteur doit s'attendre. Les deux questions du chapitre précédent cesseront de fonctionner. Non pas parce qu'elles sont erronées, mais parce que ce sont des outils d'attention, et qu'aucun outil d'attention ne résiste à la répétition.

Le dire n'est pas une mise en garde : c'est une condition pour que le déclin, lorsqu'il surviendra, soit reconnu pour ce qu'il est, au lieu d'être interprété comme une réfutation du modèle ou comme un échec personnel.

Pas de point d'arrivée

De ce point de vue, parler de guérison définitive est trompeur.

Un système prédictif n'atteint pas un état final. Il atteint, à chaque fois, un équilibre entre ce qu'il a appris et l'environnement dans lequel il se trouve.

Lorsque l'environnement change profondément, cet équilibre change lui aussi.

Cela ne signifie pas que tout progrès soit précaire. Cela signifie qu'aucun progrès ne supprime la nécessité de continuer à apprendre.

Une attitude différente

Cette conclusion modifie également la façon dont nous nous évaluons nous-mêmes.

Si le malaise réapparaît, on a naturellement tendance à l'interpréter comme un échec personnel.

Le modèle suggère une interprétation différente. Le superviseur social ne fait pas nécessairement marche arrière : il cherche à s'adapter à des informations qu'il considère comme nouvelles.

La question utile n'est donc pas « pourquoi ai-je rechuté ? », mais « quelles prévisions le système cherche-t-il à actualiser ? ».

Cette deuxième question n'élimine pas le problème. Elle permet toutefois de l'observer de la même manière que nous avons observé tout le reste du parcours : comme le fonctionnement d'un système qui formule des prévisions et les modifie à la lumière de l'expérience.

Ce modèle nous apprend également quelque chose sur la deuxième dégradation. Si l'on s'habitue à la catégorie et non à la formulation particulière, alors remplacer un outil usé par un autre mieux formulé ne sert à rien : ce n'est pas ce que l'outil demande — qui reste valable — qui s'est usé, mais la capacité de l'attention à s'en rendre compte — et celle-ci ne revient pas en changeant les mots, car ce ne sont pas les mots qui l'ont usée.

Reconnaître ce deuxième déclin signifie avant tout éviter d'attribuer aux outils un échec qui relève en réalité du fonctionnement normal de l'attention.

14 Quand mettre fin à l'observation

Cet essai a soutenu, en s'appuyant sur la littérature clinique, que l'attention portée à soi-même est l'un des mécanismes qui entretiennent le malaise social.

Il est donc contraint de dire quelque chose de dérangeant sur soi-même.

Un texte qui décrit longuement la manière dont nous imaginons être jugés, et qui propose deux questions à se poser lorsque l'on se sent observé, entraîne inévitablement une augmentation de l'attention portée à soi-même.

Pour la plupart des lecteurs, cela ne pose pas de problème. Pour certains, cela peut en être un.

Le mécanisme dont il faut se méfier

Le chapitre sur les remèdes intuitifs a déjà décrit ce mécanisme : celui qui reporte son attention sur lui-même cesse de recueillir des informations sur la façon dont les autres réagissent, et finit par utiliser ses propres sensations comme preuve de ce que les autres pensent. C'est le mécanisme qui, dans le modèle de Clark et Wells, entretient l'anxiété sociale.

Le chapitre précédent a ensuite montré que même les outils attentionnels perdent de leur efficacité, car on s'y habitue. Le risque dont il est question ici est différent : il ne concerne pas leur efficacité, qui s'épuise, mais leur utilisation, qui, dans certaines conditions, peut s'avérer néfaste.

C'est le mécanisme que le chapitre cinq a décrit comme une « présupposition élevée ».

Un outil qui invite à observer son propre fonctionnement social peut, en principe, l'alimenter. Il n'y a là rien de paradoxal : c'est la conséquence directe de ce qu'affirme le modèle.

D'où la recommandation la plus importante de tout cet essai, et la plus simple. L'attention ne doit pas être tournée vers l'intérieur, mais vers l'extérieur. Ces deux questions servent à reconnaître que le jugement présumé est en grande partie une construction, et non à mieux surveiller la manière dont on est évalué.

Le but de la méthode n'est pas d'augmenter le temps passé à observer le superviseur social. C'est de le réduire.

Que faut-il observer ?

Avant même de déterminer quand s'arrêter, il convient de dissiper un malentendu qui rend l'observation néfaste dès le départ.

Cet essai invite à observer des automatismes, et non à s'évaluer soi-même.

Il s'agit de deux activités mentales totalement différentes.

La première se présente ainsi : dans cette situation, le système a produit cette prévision.

La seconde se présente ainsi : « Je suis nul. »

La seconde n'est pas une observation : c'est un jugement, et qui plus est un jugement sur soi-même formulé à l'aide du dispositif que l'essai vient de décrire comme peu fiable. Si l'auto-observation prend cette forme, elle n'applique pas le modèle : elle en fournit un exemple.

Quand s'arrêter

L'observation devrait produire au moins l'un des deux effets suivants : une compréhension plus précise du fonctionnement de son propre système, ou une diminution progressive des réflexions qui font suite à un jugement reçu.

Si, au bout de quelques semaines, elle ne produit ni l'un ni l'autre, mais engendre au contraire davantage d'inquiétude, de rumination ou d'autocontrôle, il vaut mieux y mettre fin.

Le critère peut être formulé de manière encore plus simple : si le malaise augmente et que les prises de conscience ne suivent pas — c'est-à-dire si le malaise s'accroît sans que le nombre de fois où l'on saisit le mécanisme au moment où il se met en marche n'augmente pas —, il faut arrêter.

Ce critère n'admet pas de demi-mesure : réduire l'observation ne suffit pas, il faut y mettre fin.

C'est un critère grossier, mais il a l'avantage d'être observable : ces deux éléments peuvent être constatés en quelques semaines, sans qu'il soit nécessaire de se juger soi-même.

Et interrompre ne signifie pas avoir échoué. Cela signifie avoir appliqué ce que cet essai soutient : qu'une intervention s'évalue à ses effets, et non à sa plausibilité.

Quand ne pas l'utiliser du tout

Certaines conditions rendent l'auto-observation intensifiée plus risquée qu'utile. Il ne s'agit pas de catégories diagnostiques, mais de situations dans lesquelles l'attention risque de devenir une partie du problème.

Une forte tendance à la rumination. Une anxiété intense. Une dépression cliniquement significative. Un fonctionnement de type obsessionnel, dans lequel le contrôle est déjà l'habitude dont on souhaiterait se libérer.

Dans tous ces cas, il existe déjà une activité mentale qui s'autoalimente, et un sujet supplémentaire sur lequel l'exercer n'est pas une ressource.

Le cadre approprié de ce qui a été proposé concerne les automatismes ordinaires de la vie psychique. Pas les troubles. Pour ceux-ci, il existe des traitements éprouvés depuis des décennies, et un essai théorique ne remplace pas un professionnel — en particulier pour l'anxiété sociale, qui apparaît dans ces pages comme un cas extrême d'un mécanisme ordinaire, et non comme quelque chose à traiter par soi-même.

Quand demander de l'aide

Il existe enfin des signes qui ne concernent plus l'utilisation d'un outil, mais l'état de celui qui l'utilise.

Si l'utilisation de cet outil accroît durablement le mal-être au lieu de le réduire.

Si cela perturbe le travail.

Si elle perturbe les relations.

Si elle occupe une part croissante de la journée.

Dans ces cas-là, ce qui est utile, ce n'est pas de changer de méthode, ni de lire un autre essai : c'est d'en parler à quelqu'un qui s'en occupe de manière professionnelle.

Il convient d'ajouter que les critères exposés dans ce chapitre s'appliquent également, et à plus forte raison, au parcours opérationnel proposé dans **Recalibrer le superviseur social**. Une méthode qui demande de noter quotidiennement ses prévisions implique une exposition à l'auto-observation bien supérieure à celle de deux questions, et les mêmes contre-indications s'appliquent avec plus de force.

La limite la plus générale

Il subsiste une limite qu'aucune précaution ne peut éliminer, et qui découle du principe même de cet essai.

Il n'existe pas, au sein du système, de point d'observation extérieur. Celui qui observe son propre jugement présumé le fait à l'aide du même appareil qui le produit. Même l'évaluation « je m'observe trop » est, à son tour, une évaluation.

C'est pourquoi les critères proposés ici s'appuient sur des faits vérifiables — le malaise s'accroît, les reconnaissances ne le font pas, la journée se remplit — et non sur une impression subjective. Les impressions subjectives, en la matière, sont précisément ce dont l'essai recommande de se méfier.

Et c'est aussi la raison pour laquelle une personne extérieure, lorsqu'elle est disponible, reste le meilleur outil : non pas parce qu'elle voit juste, mais parce qu'elle voit depuis un point que le système ne peut atteindre.

Sur ce point, je peux m'exprimer de manière plus directe, et à titre personnel.

Le modèle cognitif de la panique le plus reconnu est celui de David Clark — le même auteur qui, avec Adrian Wells, a formulé le modèle de l'anxiété sociale maintes fois cité dans ces pages. Il décrit la crise comme une boucle de rétroaction : certaines sensations corporelles ordinaires sont interprétées comme le signe d'une catastrophe imminente, l'interprétation augmente l'activation, l'activation confirme l'interprétation. Il s'ensuit que la boucle se rompt lorsque l'interprétation change, et c'est un résultat avéré.

J'ai moi-même souffert de crises de panique, et celles-ci ont cessé lorsque j'ai compris ce qui les alimentait : la peur, jamais formulée, que les autres se rendent compte des jugements que je portais sur eux et me le fassent payer. C'est ce que décrit le chapitre consacré au jugement en tant qu'exposition, vu du point de vue de celui qui craint d'être démasqué.

Deux précisions s'imposent toutefois, sans lesquelles ce récit ferait plus de mal que de bien.

La première est que je n'y suis pas parvenu tout seul, ni par la lecture. Il m'a fallu une dizaine de séances de psychothérapie. Cela est cohérent avec ce que j'ai affirmé quelques lignes plus haut : le

cercle vicieux se nourrit surtout de l'intérieur, et ce qui le brise vient de l'extérieur. Je ne connais pas de moyen d'y parvenir par l'introspection.

Si un lecteur souffrant de crises de panique ne les voyait pas cesser en comprenant cela, qu'il n'en conclue pas qu'il a mal compris, ni que ce que j'ai raconté est faux : un cas ne prouve rien, et pour la même, il n'est pas contredit par un autre cas. Il se peut que sa panique trouve son origine ailleurs ; et il se peut qu'il ait très bien compris, et que comprendre, à lui seul, ne suffise pas. Pour moi, cela n'a pas suffi : cette compréhension est venue au cours d'une psychothérapie, et non à la place d'une psychothérapie.

Conclusion

Tout modèle de l'esprit court le même risque : transformer un nom en explication. On commence par essayer de rendre compte d'un phénomène, et on finit par une étiquette qui semble l'expliquer du simple fait de le nommer.

Le superviseur social peut subir ce sort, et dans le chapitre consacré à l'observateur intériorisé, j'avais tenté de l'éviter en précisant qu'il ne s'agissait pas de la voix d'une norme, mais de l'évaluation d'une réaction. C'était une promesse, et il convient de la vérifier à présent.

Si, à la fin de cet essai, le lecteur se contentait de penser « c'est mon superviseur social », le modèle aurait échoué. Non pas parce que le nom est erroné, mais parce qu'aucune étiquette n'apporte de connaissance si elle ne modifie pas la manière dont on observe les faits.

Le but de ces pages n'était pas d'introduire une nouvelle faculté mentale. Il s'agissait de proposer une manière différente d'interpréter des phénomènes que nous connaissons bien : la gêne, l'autocensure, le besoin d'approbation, le soulagement qui suit une confirmation, l'importance accordée à certaines personnes et pas à d'autres, la tendance à modifier ses opinions en fonction de son interlocuteur.

Ce modèle peut se résumer en quelques affirmations.

Le comportement social est guidé avant tout par le jugement que nous anticipons, et non par celui que nous recevons.

Cette prévision est le produit d'un système qui fonctionne de manière récursive sur une donnée non observée mais déduite, et construite principalement à partir de la façon dont nous nous percevons nous-mêmes.

Le poids attribué à un jugement ne dépend pas du jugement en soi, mais du pouvoir que celui qui le formule exerce sur nous — même si le système, au moment où il réagit, ne vérifie pas ce pouvoir.

Ce même mécanisme agit dans les deux sens : il façonne ce que nous montrons et les jugements que nous exprimons. Et les jugements que nous avons adaptés à ceux qui nous écoutaient finissent par nous sembler avoir toujours été les nôtres.

Le système continue de s'actualiser, mais pas de manière symétrique : certaines expériences modifient ses prévisions beaucoup plus facilement que d'autres, et aucune modification n'est définitive.

Si ces affirmations sont correctes, le malaise social ne dépend pas seulement de ce que les autres pensent de nous. Il dépend surtout de ce que notre système prévoit qu'ils penseront — et cette prévision, dans la plupart des cas, n'est pas une interprétation des autres.

Cette distinction n'élimine pas le problème. Elle en modifie la signification.

Lorsqu'une prévision est reconnue comme telle, elle cesse, au moins en partie, de se présenter comme une description inévitable de la réalité.

Ce n'est pas un savant qui peut décider si cette interprétation est correcte. Seule la confrontation avec les faits peut en décider.

Le critère le plus important n'est donc pas le degré de conviction que suscite le modèle, mais sa capacité à produire de nouvelles observations, des prévisions vérifiables et des questions que nous ne savions pas formuler auparavant.

S'il y parvient, il aura eu une valeur. S'il n'y parvient pas, il devra être corrigé ou remplacé.

Tel est le sort habituel de tout modèle qui prétend décrire le fonctionnement de l'esprit. Et c'est, au fond, la même règle que cet essai attribue au superviseur social : modifier ses prévisions lorsque l'expérience les contredit.

Annexe I — Références

Cette liste rassemble les travaux sur lesquels reposent les affirmations empiriques de cet essai, afin que le lecteur puisse les vérifier — et, s'il le juge nécessaire, les contester.

L'ordre est thématique et suit le fil de l'argumentation.

Les fondements sociaux du « moi »

A. Smith, *Théorie des sentiments moraux*, 1759 — le spectateur impartial.

C. H. Cooley, *Human Nature and the Social Order*, 1902 — le « looking-glass self ».

G. H. Mead, *Mind, Self, and Society*, 1934 — l'autre généralisé.

E. Goffman, *La mise en scène de la vie quotidienne*, 1956 — la gestion de l'impression.

P. Rochat, **Others in Mind : Social Origins of Self-Consciousness**, 2009.

Valeur relationnelle, appartenance, choix du partenaire

M. R. Leary et R. F. Baumeister, « The nature and function of self-esteem: Sociometer theory », dans *Advances in Experimental Social Psychology*, 32, 2000, pp. 1-62.

M. A. Nowak et K. Sigmund, « Evolution of indirect reciprocity », *Nature*, 437, 2005, pp. 1291-1298.

R. Bénabou et J. Tirole, « Incentives and prosocial behavior », *American Economic Review*, 96, 5, 2006, p. 1652-1678 — l'utilité de réputation.

Récurtivité et croyances d'ordre supérieur

J. M. Keynes, *Théorie générale de l'emploi, de l'intérêt et de la monnaie*, 1936, chap. 12 — le concours de beauté.

R. D. Laing, H. Phillipson et A. R. Lee, *Interpersonal Perception: A Theory and a Method of Research*, 1966.

N. Luhmann, *Soziale Systeme : Grundriß einer allgemeinen Theorie*, 1984 — la double contingence.

R. Nagel, « Unraveling in guessing games : An experimental study », *American Economic Review*, 85, 5, 1995, pp. 1313-1326 — le modèle à plusieurs niveaux.

T. Mauersberger et R. Nagel, « Levels of Reasoning in Keynesian Beauty Contests : A Generative Framework », 2018 — revue de la littérature.

Perception des personnes

J. S. Uleman et ses collaborateurs, sur les inférences spontanées de traits de personnalité : l'attribution automatique et non intentionnelle de caractéristiques stables à partir de l'observation d'un comportement.

A. Bott, L. Brockmann, I. Denneberg, E. Henken, N. Kuper, F. Kruse et J. Degner, « Spontaneous Trait Inferences From Behavior: A Systematic Meta-Analysis », *Personality and Social Psychology Bulletin*, 2022 — quatre-vingt-six publications, 13 630 participants, effet moyen $d_z = 0,59$.

J. Degner, « Revisiting Spontaneous Trait Inferences: Questions, Challenges, and Intradisciplinary Connections », *Social and Personality Psychology Compass*, 2025 — revue critique.

Métaperceptions et précision

D. A. Kenny et B. M. DePaulo, « Do people know how others view them? An empirical and theoretical account », *Psychological Bulletin*, 114, 1, 1993, pp. 145-161.

E. N. Carlson, R. M. Furr et S. Vazire, sur la méta-intuition : la capacité à reconnaître quand notre estimation du jugement d'autrui est plus ou moins exacte.

E. J. Boothby, G. Cooney, G. M. Sandstrom et M. S. Clark, « The Liking Gap in Conversations : Do People Like Us More Than We Think ? », *Psychological Science*, 29, 11, 2018, p. 1742-1756.

Harper et ses collaborateurs, « How Well Do We Know How Others See Us? A Systematic Review of Meta-Accuracy Across Relational Contexts, Research Foci, Attribute Domains, and Measurement Approaches », *Journal of Personality*, doi 10.1111/jopy.70075.

Narcissisme et métaperception

E. N. Carlson, S. Vazire et T. F. Oltmanns, « You probably think this paper's about you: Narcissists' perceptions of their personality and reputation », *Journal of Personality and Social Psychology*, 101, 1, 2011, p. 185-201.

Il convient de noter que la littérature scientifique n'est pas unanime : il existe un courant opposé, selon lequel les traits narcissiques s'accompagneraient également d'une surévaluation du jugement d'autrui. Le résultat utilisé dans cet essai est celui qui bénéficie du plus grand soutien, mais ce n'est pas le seul disponible.

Le jugement comme signal

P. Bourdieu, *La distinction : critique sociale du jugement*, 1979.

T. Kuran, *Private Truths, Public Lies : The Social Consequences of Preference Falsification*, 1995.

J. Tosi et B. Warmke, « Moral Grandstanding », *Philosophy & Public Affairs*, 44, 3, 2016, p. 197-217.

J. J. Jordan, R. Sommers, P. Bloom et D. G. Rand, « Why Do We Hate Hypocrites? Evidence for a Theory of False Signaling », *Psychological Science*, 28, 3, 2017, p. 356-368.

J. J. Jordan et D. G. Rand, « Signaling when no one is watching : A reputation heuristics account of outrage and punishment in one-shot anonymous interactions », *Journal of Personality and Social Psychology*, 117, 1, 2019, p. 57-88.

E. Goffman, « On Face-Work : An Analysis of Ritual Elements in Social Interaction », *Psychiatry*, 18, 1955, p. 213-231 — le consensus de travail.

P. Brown et S. C. Levinson, *Politeness: Some Universals in Language Usage*, Cambridge University Press, 1987 — la protection réciproque de la face.

J. A. Minson et B. Monin, « Do-Gooder Derogation : Disparaging Morally Motivated Minorities to Defuse Anticipated Reproach », *Social Psychological and Personality Science*, 3, 2, 2012 — la dévalorisation de ceux qui adoptent un critère plus exigeant.

Comment la communication modifie les convictions

E. T. Higgins et W. S. Rholes, « "Saying is believing" : Effects of message modification on memory and liking for the person described », *Journal of Experimental Social Psychology*, 14, 1978, pp. 363-378.

G. Echterhoff, E. T. Higgins et J. M. Levine, « Shared reality : Experiencing commonality with others' inner states about the world », *Perspectives on Psychological Science*, 4, 2009, pp. 496-521.

L. Festinger, *A Theory of Cognitive Dissonance*, 1957 — cité à titre de distinction, et non comme fondement.

D. J. Bem, « Self-perception theory », dans *Advances in Experimental Social Psychology*, 6, 1972 — idem.

F. Heider, « Attitudes and cognitive organization », *Journal of Psychology*, 1946, et *The Psychology of Interpersonal Relations*, Wiley, New York, 1958 — la théorie de l'équilibre.

G. Simmel, « Fashion », *International Quarterly*, 10, 1904 — imitation et distinction.

Anonymat, observation, comportement

E. Hoffman, K. McCabe et V. L. Smith, « Social distance and other-regarding behavior in dictator games », *American Economic Review*, 86, 3, 1996, p. 653-660.

M. Bateson, D. Nettle et G. Roberts, « Cues of being watched enhance cooperation in a real-world setting », *Biology Letters*, 2, 3, 2006, pp. 412-414 — ce qu'on appelle l'effet « watching-eyes », selon lequel de simples images d'yeux augmenteraient les comportements coopératifs. Ce serait l'expérience naturelle pour un modèle tel que celui présenté dans ces pages, et le modèle ne s'appuie pas sur elle : les méta-analyses ci-dessous ne le confirment pas.

Concernant les méta-analyses qui ne confirment pas cet effet : « Artificial surveillance cues do not increase generosity: two meta-analyses » ; et K. Dear, K. Dutton et E. Fox, « Do watching eyes influence antisocial behavior? A systematic review and meta-analysis », *Evolution and Human Behavior*, 40, 3, 2019, pp. 269-280.

Conscience de soi, attention, introspection

S. Duval et R. A. Wicklund, *A Theory of Objective Self Awareness*, Academic Press, New York, 1972.

A. Fenigstein, M. F. Scheier et A. H. Buss, « Public and private self-consciousness: Assessment and theory », *Journal of Consulting and Clinical Psychology*, 43, 1975, pp. 522-527.

T. Gilovich et K. Savitsky, « The spotlight effect and the illusion of transparency : Egocentric assessments of how we are seen by others », *Current Directions in Psychological Science*, 8, 1999, pp. 165-168.

A. M. Grant, J. Franklin et P. Langford, « L'échelle d'introspection et de prise de conscience : une nouvelle mesure de la conscience de soi privée », *Social Behavior and Personality*, 30, 8, 2002, p. 821-836.

R. Harrington et D. A. Loffredo, « Insight, Rumination, and Self-Reflection as Predictors of Well-Being », *The Journal of Psychology*, 145, 1, 2011, p. 39-57 — sur le paradoxe de l'égocentrisme.

Clinique de l'anxiété sociale

D. M. Clark et A. Wells, « A cognitive model of social phobia », dans R. G. Heimberg, M. R. Liebowitz, D. A. Hope et F. R. Schneier (dir.), *Social Phobia: Diagnosis, Assessment, and Treatment*, Guilford Press, New York, 1995, p. 69-93.

J. P. Tangney et R. L. Dearing, *Shame and Guilt*, Guilford Press, New York, 2002 — la distinction entre la honte et la culpabilité.

D. M. Clark, « A Cognitive Approach to Panic », *Behaviour Research and Therapy*, 24, 4, 1986, pp. 461-470.

Structure dimensionnelle de l'anxiété sociale

A. M. Ruscio, « The latent structure of social anxiety disorder: consequences of shifting to a dimensional diagnosis », *Journal of Abnormal Psychology*, 119, 4, 2010, p. 662-671 — analyses taxométriques des données de la National Comorbidity Survey Replication.

S. Crome, A. Baillie, T. Slade et A. M. Ruscio, « Phobie sociale : nouvelles preuves d'une structure dimensionnelle », *Australian and New Zealand Journal of Psychiatry*, 44, 11, 2010, p. 1012-1020.

Traitement prédictif

A. Clark, « Whatever next? Predictive brains, situated agents, and the future of cognitive science », *Behavioral and Brain Sciences*, 36, 3, 2013, p. 181-204.

Il existe une littérature récente et en pleine évolution sur l'interprétation de l'anxiété sociale et de la solitude en termes de précision attribuée aux signaux sociaux, sur laquelle cet essai s'appuie comme cadre d'interprétation et non comme fondement.

Dépression et précision

M. T. Moore et D. M. Fresco, « Depressive realism: A meta-analytic review », *Clinical Psychology Review*, 32, 6, 2012, pp. 496-509 — soixante-quinze études, plus de sept mille participants, effet d'ampleur négligeable.

« Sadder ≠ Wiser : Depressive Realism Is Not Robust to Replication », *Collabra : Psychology*, 8, 1, 2022 — reproduction infructueuse de l'étude originale de L. B. Alloy et L. Y. Abramson (1979).

Sur la distorsion négative des métaperceptions associée aux symptômes dépressifs et à une faible estime de soi : les études sur la concordance entre soi et les autres et sur la précision des métaperceptions dans les traits du modèle à cinq facteurs.

Variations culturelles

R. A. Kleinknecht, D. L. Dinnel, E. E. Kleinknecht, N. Hiruma et N. Harada, « Cultural factors in social anxiety: a comparison of social phobia symptoms and Taijin kyofusho », *Journal of Anxiety Disorders*, 11, 2, 1997, pp. 157-177.

B. Zhou, F. Lacroix, J. Sasaki, Y. Peng, X. Wang et A. G. Ryder, « Analyse des variations culturelles dans l'anxiété sociale et le type « offensif » du Taijin Kyofusho à travers les effets indirects de l'intolérance à l'incertitude et des conceptions de soi », *Journal of Cross-Cultural Psychology*, 45, 10, 2014, pp. 1561-1578.

Évaluation sociale dans la petite enfance

J. K. Hamlin, K. Wynn et P. Bloom, « Évaluation sociale chez les nourrissons préverbaux », *Nature*, 450, 2007, pp. 557-559.

K. Lucca et ses collaborateurs (ManyBabies Consortium), « Évaluation sociale par les nourrissons des personnes qui les aident et de celles qui les gênent : une étude de réplication coordonnée à grande échelle et multi-laboratoires », *Developmental Science*, 2025 — réplication infructueuse du résultat précédent.

N. Botto et P. Rochat, « Sensitivity to the Evaluation of Others Emerges by 24 Months », *Developmental Psychology*, 54, 9, 2018, pp. 1723-1734.

Travaux antérieurs de l'auteur

Les travaux suivants ne constituent pas des sources indépendantes à l'appui du modèle. Ils représentent les étapes de son élaboration et sont cités dans le texte afin de ne pas répéter des arguments déjà développés ailleurs.

La distinction est importante : un modèle ne se confirme pas en citant ses propres versions antérieures.

Psychologie des besoins — pour les notions de besoin et de désir, qui ne sont pas redéfinies ici.

Le jugement présumé, 2026 — première formulation du concept, du superviseur du jugement social et du coût relationnel attendu. Il désignait dans l'effet « *watching-eyes* » la prédiction la plus

caractéristique du modèle ; les méta-analyses ultérieures ne l'ont pas confirmée, et ces pages ne la reprennent pas.

Autogouvernance sociale, 2026 — le système qui oriente à travers les émotions, et l'extension à ce que l'on est disposé à prendre en considération.

L'autothérapie comme gestion de l'attention, 2026 — la fenêtre de choix, la distinction entre entraîner et déclencher, l'accoutumance aux outils attentionnels.

Recalibrer le superviseur social, 2026 — le parcours opérationnel dérivé de ce modèle : reconnaissance de la prévision, reconstruction des éléments qui l'ont produite, scénarios alternatifs, expériences comportementales, registre personnel, erreurs les plus fréquentes et critères d' t d'interruption. C'est le travail auquel le présent essai renvoie pour la partie pratique, et d'où provient la distinction entre observations, inférences et prévisions, avec son application au jugement exprimé, réel et présumé.

Un précédent non retenu

I. Kant, *Critique de la faculté de jugement (Kritik der Urteilkraft)*, 1790.

Lorsque l'idée de cet essai m'est venue, je me suis demandé si je n'étais pas en train de réinventer une roue déjà inventée par Kant. Je suis allé vérifier. La réponse est non, et je la rapporte ici car la question est légitime et d'autres pourraient se la poser.

La *Critique de la faculté de jugement* porte sur la faculté de jugement et traite du beau, du sublime et de la finalité dans la nature. Elle ne dit rien des conséquences relationnelles du jugement. On y trouve une observation qui tient toujours : celui qui dit « ceci est beau » n'exprime pas une préférence personnelle, mais revendique l'assentiment d'autrui sans pouvoir présenter quoi que ce soit qui l'impose. C'est le précédent le plus illustre du critère de démarcation que cet essai a examiné et écarté.

Je dois toutefois ajouter mon propre jugement, qui est sévère. La distinction entre « j'aime » et « c'est beau » était nouvelle en 1790 et est aujourd'hui une banalité — ce qui est le sort habituel des idées réussies, et non un défaut. Le reste m'a semblé être un exercice analytique arbitraire et inutilement compliqué, conçu davantage pour maintenir la cohérence d'un système que pour comprendre comment les gens jugent. Je n'ai rien retiré de ces pages. Il s'agit d'une impression de lecture, et non d'une thèse d'histoire de la philosophie, et je le précise clairement.

Remarque sur l'exhaustivité de cette liste

Les vérifications sur lesquelles s'appuie cet essai ont été menées à partir de sources universitaires accessibles, en anglais et en italien. Elles ne constituent pas une revue systématique, et les principales bases de données spécialisées n'ont pas été consultées.

Il existe donc, selon toute vraisemblance, des travaux pertinents qui ne figurent pas ici — y compris des travaux qui pourraient anticiper certaines parties du modèle ou les contredire.

Annexe II — Ce qu'affirme le modèle

Au cours de cet essai, plusieurs prévisions ont été formulées. Dispersées dans les chapitres, elles restent toutefois invisibles en tant qu'ensemble : on les remarque une à une, puis on les oublie.

Cette annexe les rassemble en un seul endroit.

Les prévisions sont classées par ordre décroissant de risque : de la plus exposée, dont l'échec compromettrait l'ensemble du système, à celles dont l'échec ne nécessiterait que des corrections locales.

La généralisation du pathologique à l'ordinaire

Les différences entre une anxiété sociale élevée et faible sont principalement quantitatives — poids des a priori, précision attribuée à l'information externe, profondeur de la récursivité — et non qualitatives.

C'est la prédiction la plus risquée de cet essai. L'architecture décrite existe déjà dans la littérature en tant que modèle d'un trouble ; ce que l'on soutient ici, c'est qu'elle décrit le fonctionnement ordinaire, dont l'anxiété sociale serait le cas extrême.

Justification disponible : les études taxométriques convergent vers une structure dimensionnelle de l'anxiété sociale, continue avec le fonctionnement normal.

Limite de ces éléments à l'appui : la dimensionnalité concerne les symptômes, et non les mécanismes. Un phénotype continu ne démontre pas que les sujets présentant un faible niveau d'anxiété utilisent les mêmes processus, mais à un degré moindre.

Ces deux paramètres sont bel et bien distincts

supposition et la dévalorisation — dans quelle mesure le jugement présumé est construit à partir de l'auto-perception, et dans quelle mesure le retour externe est dévalorisé — varient de manière relativement indépendante au sein de la population.

Il s'ensuit que les quatre combinaisons possibles devraient toutes être observées : niveau élevé d'anxiété sociale et niveau élevé de narcissisme, niveau faible d'anxiété sociale et niveau élevé de narcissisme, ainsi que les deux combinaisons mixtes à vérifier.

Les deux combinaisons non encore observées

Une présupposition élevée et une minimisation faible devraient produire une image de soi instable : la perception de la façon dont on est perçu oscille à chaque retour d'information externe.

Une présupposition faible et une dévalorisation faible devraient correspondre au profil le mieux calibré, dans lequel le jugement présumé est construit à partir à la fois de la perception de soi et des retours disponibles. Ce n'est pas un trouble : c'est le cas où le système fait ce pour quoi il semble avoir été conçu.

Ces deux combinaisons ont été formulées avant de consulter la littérature, précisément pour ne pas les adapter aux résultats.

La profondeur de la récursivité

Le nombre de niveaux itérés — « que vont-ils penser de moi », « que vont-ils penser de mon apparence » — augmente à mesure que l'anxiété sociale s'accroît.

Cette prévision naît de la rencontre entre deux domaines de la littérature qui ne se citent pas mutuellement : le modèle à niveaux dans les jeux de devinettes, où les personnes s'arrêtent après un ou deux passages, et la recherche clinique sur l'anxiété sociale.

Elle est vérifiable à l'aide d'outils déjà existants.

Le poids du jugement dépend du pouvoir de celui qui le formule

Le poids attribué à un jugement dépend du pouvoir que celui qui le formule exerce sur nous, et non de la probabilité d'interactions futures.

Il s'agit d'une correction d'une formulation antérieure de l'auteur, et la différence est vérifiable : en manipulant expérimentalement le pouvoir qu'un observateur exerce sur le sujet tout en maintenant constante la fréquence des interactions, l'autocensure, le conformisme et l'activation émotionnelle devraient varier de manière prévisible.

La distorsion du jugement exprimé

La composante d'autoprésentation d'un jugement augmente proportionnellement à son observabilité et à l'importance de l'auditeur, et ne tombe pas à zéro, même dans des conditions d'anonymat.

Le coût de la violation

Celui qui présente un critère comme valable pour tout le monde devrait être jugé plus sévèrement, lorsqu'il l'enfreint, que celui qui le présente comme une préférence personnelle ou comme une règle valable uniquement dans certaines circonstances.

Il convient de distinguer cela de ce qui a déjà été observé : à savoir que la sanction de l'hypocrisie dépend de la falsification du signal et non d'une simple incohérence. C'est là un résultat acquis. La proportionnalité entre la portée déclarée du critère et la sévérité de la sanction ne l'est pas.

Le comportement privé et le comportement public

L'observateur intériorisé est systématiquement plus faible que l'observateur extérieur. Il s'ensuit que le comportement dans des conditions d'anonymat réel devrait être moins marqué que le comportement public, sans pour autant disparaître complètement.

C'est cette prévision qui empêche d'utiliser l'observateur intériorisé comme explication universelle.

Les effets attendus des deux questions

Quiconque utilise avec une certaine régularité les deux questions proposées au douzième chapitre devrait observer deux effets modestes : moins de temps passé à reconstituer des scénarios après avoir reçu un jugement, et quelques cas où l'on se rend compte d'avoir changé d'avis sans s'en rendre compte.

C'est la seule prévision que le lecteur peut vérifier par lui-même, et c'est aussi la plus fragile, car elle ne fait pas la distinction entre un modèle erroné et un outil mal utilisé.

La dépression parmi les quatre combinaisons

Le modèle soutient que la représentation négative du jugement d'autrui est en grande partie déduite plutôt qu'observée. Une tradition de recherche aurait pu le réfuter — le réalisme dépressif, selon lequel, dans certaines conditions dépressives, certaines évaluations seraient plus précises plutôt que plus faussées — mais l'effet n'a pas tenu : une méta-analyse portant sur soixante-quinze études lui attribue une ampleur négligeable, et une réplique récente ne le reproduit pas.

Les données pertinentes vont au contraire dans le sens inverse : les symptômes dépressifs et la faible estime de soi faussent les métaperceptions dans un sens négatif, c'est-à-dire que la personne qui en souffre croit être perçue plus négativement qu'elle ne l'est réellement. C'est ce que le modèle appelle une « présupposition élevée », et la dépression occupe donc la même combinaison que l'anxiété sociale : il n'est pas nécessaire d'ajouter un paramètre. Ce qui ne prouve pas que le mécanisme soit identique — cela montre simplement que les deux tableaux partagent les valeurs des deux paramètres.

Le mécanisme hors du monde occidental

Le modèle a été élaboré à partir d'une littérature en grande partie occidentale, et le poids normatif du jugement d'autrui varie sensiblement entre les sociétés plus individualistes et celles plus collectivistes. Il fallait vérifier s'il était transposable.

Le cas le plus étudié est le *taijin kyofusho*, la forme japonaise de l'anxiété sociale. La différence par rapport à la description occidentale ne concerne pas l'existence du mécanisme, mais son contenu : la peur est centrée sur le fait d'offenser les autres plutôt que sur la gêne que l'on ressent soi-même. L'anticipation du jugement est la même ; ce qui change, c'est ce que l'on craint que le jugement porte sur.

La variable qui médie l'association avec la culture n'est pas la présence ou l'absence du mécanisme, mais le type de construction du soi — indépendant ou interdépendant — et le phénomène a également été observé en dehors du Japon, dans d'autres contextes collectivistes.

Cela confirme de manière indépendante une conséquence du modèle qui n'avait pas encore été vérifiée : le mécanisme est général, les critères qu'il applique sont importés de l'environnement. Il en découle une prédiction déjà formulable et désormais corroborée : la variabilité interculturelle devrait être élevée quant au contenu du jugement redouté et faible quant à la structure du processus.

Convergence et divergence des jugements

Chez ceux qui coopèrent, les jugements devraient converger au-delà de ce que justifient les informations partagées ; chez ceux qui sont en conflit, ils devraient diverger au-delà de ce que justifient les informations.

La prédiction ne concerne pas la convergence en soi, qui s'expliquerait par le partage des données, mais son excès par rapport aux preuves disponibles.

La sélection préalable des idées

Le coût social attendu lié au fait d'être associé à une idée intervient avant son évaluation et réduit la probabilité qu'elle soit prise en considération, indépendamment de son contenu.

Il s'agit de l'extension proposée dans le chapitre sur l'adaptation à l'auditeur, et c'est la moins étayée parmi celles avancées dans ces pages : elle découle du modèle par analogie avec ce qui se passe pour les jugements exprimés, et non d'un résultat.

Une anomalie que le modèle n'explique pas

L'étude des cultures a toutefois produit un résultat que le modèle ne prévoit pas.

Les scores moyens d'anxiété sociale s'avèrent plus élevés dans les pays collectivistes, tandis que la prévalence clinique est plus élevée dans les pays individualistes. Ces deux mesures vont dans des directions opposées.

Le modèle décrit dans ces pages ne l'explique pas. Il suggère une interprétation possible — ce qui varierait, ce ne serait pas seulement l'intensité du mécanisme, mais le seuil au-delà duquel un environnement le considère comme pathologique — et cette interprétation est cohérente avec l'idée, développée dans le chapitre sur la désadaptation, selon laquelle le système doit être évalué par rapport à l'environnement dans lequel il opère. Mais il s'agit d'une interprétation plausible, et non d'une conséquence du modèle : rien de ce qui précède ne la prévoyait.

Elle est mentionnée ici car une dissociation de ce type est précisément le genre de donnée qu'une théorie devrait expliquer plutôt que d'absorber.

Annexe III — Questions pour approfondir

Les deux questions du chapitre douze sont des déclencheurs : elles fonctionnent si elles sont toujours associées au même signal corporel — le nœud à l'estomac, l'hésitation avant de parler — au moment où le système s'active.

Les questions rassemblées ici ont un statut différent. Ce sont des aides à la réflexion : on les utilise à tête reposée, face à une décision à prendre — exprimer un jugement ou le taire, comment l'exprimer, essayer d'en vérifier la validité, le modifier, en réduire le poids. Elles ne doivent être ancrées à rien et ne prétendent pas intervenir sur la réaction : elles interviennent après, lorsque la réaction a déjà eu lieu et qu'il reste à décider quoi en faire.

Il convient toutefois de préciser en quoi elles approfondissent la réflexion, car le chapitre dix a établi que réfléchir davantage, en soi, ne produit aucun changement : la réflexion retravaille la matière interne et, au regard des deux paramètres, renforce la présupposition au lieu de l'affaiblir.

L'approfondissement proposé ici est d'un autre type, dans le douzième chapitre, à la section « *Si une prévision revient sans cesse* ». Il ne consiste pas à revenir sur un jugement jusqu'à ce qu'il soit convaincant, mais à en rendre visible la construction : quels faits observables l'ont engendré, quelles inférences y ont été ajoutées, qu'est-ce qui pourrait le réfuter. Un jugement dont on entrevoit la construction ressemble déjà moins à un fait. Si, en revanche, l'utilisation de ces questions se réduit à repenser plus longuement la même chose, elles produisent l'effet que l'essai déconseille, et il vaut mieux y renoncer.

Il ne faut pas non plus toutes les adopter. Dix questions utilisées quotidiennement cessent rapidement d'avoir un sens : la dégradation décrite au chapitre treize s'applique également à elles, et à plus forte raison. Cette annexe est un répertoire à consulter en cas de besoin, et non un outil à utiliser en permanence.

Chaque question découle d'un point précis de l'essai, indiqué à chaque fois. Cette indication n'est pas une simple fioriture : séparée du chapitre qui la justifie, chacune de ces questions devient une maxime de bon sens, et le critère permettant de l'utiliser — avec ses limites — se trouve là.

Ce que vous avez entre les mains

Le jugement dont il est question ici a-t-il été prononcé, ou est-ce moi qui l'ai anticipé ?

D'après le douzième chapitre, qui distingue le jugement exprimé, le jugement réel et le jugement présumé. Une partie des décisions prises « en réponse à un jugement » répond en réalité à une prévision : personne n'a rien dit. Être le premier à l'établir modifie la valeur de toutes les questions suivantes, car on peut raisonner sur un jugement exprimé comme sur un fait, tandis que sur un jugement présumé, on raisonne sur un produit de son propre système.

Qu'ai-je observé, et qu'ai-je ajouté ?

Même origine : la distinction entre observations, inférences et prévisions, appliquée cette fois à son propre verdict. « Il m'a salué rapidement » et « il ne m'invitera plus » semblent pour l'instant être la même chose ; une décision fondée sur la première et une fondée sur la seconde n'ont pas la même base. La question ne demande pas de renoncer aux inférences — sans elles, la vie sociale serait impossible — mais de savoir combien on en a faites.

Faut-il l'exprimer, et comment ?

Si je l'exprime, que révèle-t-elle de moi à celui qui m'écoute ?

Extrait du chapitre sept. Tout critère rendu public est une information sur celui qui le propose, et à partir de ce moment-là, il s'applique à lui. La question ne suggère pas de se taire : elle demande de prendre en compte le coût du signal, qui existe toujours, même lorsque le jugement est fondé.

Celui qui m'écoute a-t-il du pouvoir sur moi ?

C'est la première question de l'essai, abordée sous son angle actif. Le coût d'un jugement exprimé ne dépend pas de l'ampleur de l'auditoire, mais du pouvoir de ceux qui en font partie : un même jugement peut être gratuit dans un contexte et coûteux dans un autre. Là encore, la réponse peut très bien être oui — et alors la réticence n'est pas de la lâcheté : c'est l'équilibre décrit dans **Le pacte que personne ne conclut**, vu de son propre point de vue.

Mon silence sera-t-il interprété comme un jugement ?

Extrait du chapitre VII, section « *Le critère que l'on n'adopte pas* ». L'abstention n'est pas perçue comme une absence de jugement : celui qui ne suit pas ce que les autres suivent, ou qui se tait là où tout le monde s'exprime, est interprété comme quelqu'un qui a porté un jugement. Le véritable choix ne se situe donc pas entre s'exprimer et ne pas communiquer, mais entre deux communications différentes — et il convient d'évaluer comment la seconde sera interprétée avant de la préférer par prudence.

Ce que je m'apprête à exprimer, est-ce le jugement que j'ai, ou celui qu'il convient d'avoir ici ?

Extrait du chapitre huit. La question ne demande pas si l'adaptation aura lieu : le chapitre soutient qu'elle tend à se produire de toute façon, et sans laisser de trace. Elle demande une vérification de l'écart entre le jugement que l'on a et celui que l'on s'apprête à exprimer, car c'est le seul moment où cet écart est encore visible. S'il y a un écart, la décision change de nature : il ne s'agit plus seulement de s'exprimer, mais d'accepter que la version convenable, répétée, devienne la sienne. S'il n'y en a pas, la question se résout en un instant, comme il se doit.

Faut-il le vérifier, le modifier, en réduire le poids ?

Dirais-je la même chose devant ceux qui pensent différemment ?

C'est la deuxième question posée par l'essai sous la forme d'une comparaison entre des situations : non pas une inspection, mais la comparaison entre deux comportements observables. Le douzième chapitre en indique également la limite : il vaut mieux effectuer cette comparaison sur des cas où la différence est flagrante, car sur les nuances, le sceptique a toujours une explication toute prête.

Ma prévision peut-elle être contredite par quelque chose ?

Par le onzième chapitre. Si l'on veut vérifier la validité d'un jugement, la condition n'est pas d'y réfléchir davantage : c'est de le formuler au préalable, par écrit, sous une forme que l'expérience puisse contredire. Une prévision vague est toujours réabsorbée ; une prévision consignée par écrit ne peut plus être réécrite a posteriori. Si l'on n'est pas disposé à cela, la vérification n'a pas encore commencé.

Dans quelle mesure mon attachement à ce jugement est-il un attachement à ce que j'ai déclaré ?

Extrait des deux chapitres centraux de la deuxième partie, pris ensemble. Un jugement énoncé est un signal déjà émis : le modifier expose à l'incohérence, et le système le sait avant nous. Dans la résistance à changer un jugement, il convient donc de distinguer la partie épistémique — les raisons pour lesquelles il semble encore valable — de la partie liée à la réputation, c'est-à-dire le coût de se contredire. Seule la première concerne la validité.

En minimiser l'importance est-ce une constatation ou une consolation ?

Extrait du chapitre dix. Les assurances que le système dévalorise peuvent aussi être auto-administrées : « ça ne compte pas », dit pour se calmer, est une information qui arrive et ne change rien. L'alternative est la constatation — le pouvoir sur nous existe ou n'existe pas — avec la fonction que le douzième chapitre lui reconnaît : contenir l'élaboration ultérieure, non pas effacer le premier choc. Réduire l'importance d'un jugement est une conclusion à laquelle on parvient ; lorsqu'il s'agit d'un point de départ, c'est une consolation.

Enfin, l'avertissement du chapitre quatorze s'applique également à cette annexe. Approfondir, ici, ne signifie pas revenir en arrière : ces questions servent à rendre un jugement observable et à décider quoi en faire, et une fois la décision prise, elles ont épuisé leur fonction. Si l'on se rend compte que l'on y revient par la suite, ce n'est plus approfondir — c'est ruminer avec un matériau de meilleure qualité. Les critères pour s'arrêter sont ceux déjà donnés, et ils n'admettent pas de demi-mesures.

Annexe IV — Glossaire

Certains termes utilisés dans cet essai ont un sens plus restreint que celui qui leur est couramment attribué. La précision d'un modèle dépendant de la stabilité de son lexique, il convient de les définir.

Les entrées sont de trois types, et la distinction est précisée au cas par cas : les termes définis dans ces pages, les termes tirés de la littérature scientifique, et les termes provenant d'ouvrages antérieurs de l'auteur.

Entraînement et déclenchement

Termes définis. Deux fonctions opposées de l'attention portée à l'un de ses propres automatismes. L'entraînement maintient le sujet vivant en l'observant sous des angles toujours différents, et exige de la variété. Le déclenchement sert à reconnaître l'automatisme au moment où il se met en marche, et exige le contraire : un signal spécifique et constant, répété toujours de la même manière.

Les confondre est la raison pour laquelle de nombreuses tentatives d'auto-observation échouent. Ces termes proviennent d'un travail antérieur de l'auteur sur l'attention.

Être choisi

C'est un fait établi. La principale voie par laquelle le jugement d'autrui produit des conséquences : la probabilité d'être préféré en tant qu'allié, ami, partenaire, collaborateur, employé ou modèle. Ce qui est en jeu, c'est la valeur sociale, c'est-à-dire à quel point une personne est désirable en tant que partenaire d'interaction.

Ce n'est pas la seule voie — celui qui juge peut aussi entraver, punir ou dénoncer à d'autres — mais c'est celle autour de laquelle le système s'est construit. La compétition ne fait pas exception : il s'agit d'une compétition pour être choisi dans des conditions de rareté.

Jugement

Au sens strict, et de manière plus précise que dans l'usage courant. Une évaluation qui devient un verdict et qui influence la relation avec la personne jugée, ou avec d'autres juges si l'objet du jugement est une chose.

Le verdict se reconnaît à l'emploi du mot. Le jugement consiste en des qualifications et non en des mesures. Et c'est un point d'arrivée, pas un parcours : « je suis encore en train d'évaluer » a un sens, « je suis encore en train de juger » n'en a pas.

La conséquence relationnelle est ce qui le distingue de l'évaluation. À l'égard d'une personne jugée, on adopte une attitude plus ou moins favorable, qui tend vers l'inclusion ou l'exclusion, la confiance ou la méfiance, et, dans le cas moral, le désir de punir ou de récompenser. Lorsque l'objet n'est pas une personne, la conséquence retombe sur celui qui a émis le jugement.

La distinction ne porte pas sur l'objet — on peut juger une œuvre et évaluer une personne — mais sur la nature de l'acte. Et elle n'implique pas nécessairement une prise de conscience : vis-à-vis des

personnes, le verdict se forme automatiquement, comme le montrent les recherches sur les inférences spontanées de traits de personnalité évoquées dans le premier chapitre.

Jugement présumé

D'après des travaux antérieurs de l'auteur. La représentation anticipée, généralement implicite et inconsciente, du jugement que d'autres pourraient porter sur notre comportement.

Le terme « jugement » doit ici être compris au sens large : il ne s'agit pas de l'opinion exprimée verbalement, mais de l'ensemble de la réponse sociale attendue — approbation ou désapprobation, acceptation ou rejet, disposition à coopérer ou perte de confiance.

Il s'agit d'une troisième grandeur, distincte tant du jugement réel que de l'auto-évaluation, et c'est elle qui oriente le comportement.

Observateur intériorisé

Terme à l'histoire ancienne, utilisé ici dans son sens propre. Mécanisme qui continue d'évaluer le jugement d'autrui même en l'absence d'observateurs réels.

Il convient de le distinguer de trois concepts antérieurs auxquels il est apparenté mais avec lesquels il ne coïncide pas : le spectateur impartial d'Adam Smith et l'autre généralisé de Mead ont une cible unique et stable — un idéal moral, la communauté — tandis que l'observateur décrit ici change de cible en fonction de la personne dont on prévoit qu'elle jugera. Ce n'est pas la voix d'une norme : c'est l'estimation d'une réaction.

Dans ce modèle, il est systématiquement plus faible que l'observateur extérieur.

Hypothèse

Stipulée. L'un des deux paramètres qui régissent le jugement présumé — l'autre étant l'actualisation. Elle mesure dans quelle mesure l'estimation de la façon dont les autres nous évaluent découle de la façon dont nous nous évaluons nous-mêmes, plutôt que de ce qui est observable dans la situation.

Lorsqu'elle est élevée, on part du principe que les autres nous voient tels que nous nous voyons — pour le meilleur comme pour le pire — et non par distraction : on regarde, et on trouve ce que l'on s'attendait à trouver.

Cela concerne la manière dont la représentation est construite ; la décote concerne la manière dont elle est maintenue.

Recalibrage

Convenu. La modification des prévisions du superviseur social à la suite d'expériences qui les contredisent de manière vérifiable.

Elle ne consiste pas à apporter de nouvelles informations, mais à empêcher qu'elles ne soient dévalorisées. Ce n'est pas un événement achevé : le système continue de se mettre à jour, et peut également se recalibrer dans un sens péjoratif.

Récurtivité

Stipulé, avec des antécédents en économie et en sociologie. Propriété d'une évaluation qui peut à son tour devenir l'objet d'une évaluation.

Elle n'indique pas une réflexion consciente du type « je pense que tu penses que je pense ». La récurtivité spontanée est faible : dans la plupart des situations, elle ne dépasse pas un ou deux niveaux, et c'est un processus rapide et en grande partie inconscient.

C'est le critère qui délimite le domaine de l'essai : les évaluations non récurtives n'en font pas partie.

Remise

Stipulé. L'un des deux paramètres qui régissent le jugement présumé — l'autre étant la présupposition. Il mesure dans quelle mesure le retour d'information provenant des autres est relativisé, réinterprété ou ignoré lorsqu'il le contredit.

Cela concerne la manière dont la représentation est maintenue. C'est le paramètre sur lequel se brisent les assurances : l'information arrive, mais ne modifie pas le système.

Superviseur social

Tiré de travaux antérieurs de l'auteur. Le système qui évalue les conséquences relationnelles des comportements possibles et oriente leur sélection.

Il ne contrôle pas le comportement de manière directe : il l'oriente par le biais des émotions, sous forme d'une incitation à agir lorsque la prévision est favorable et d'une résistance lorsqu'elle est défavorable. Ce qui parvient à la conscience, c'est le résultat, et non le calcul.

Le nom est une abréviation, et comporte un risque explicitement mentionné dans la conclusion : que l'étiquette se substitue à l'explication.

Valeur sociale

Convenue. Dans quelle mesure une personne est jugée désirable en tant que partenaire d'interaction : allié, ami, collaborateur, employé, employeur, modèle à imiter.

Ce n'est ni une propriété de la personne ni une valeur unique : cela dépend des besoins de celui qui évalue, et une même personne peut avoir beaucoup de valeur pour l'un et peu pour l'autre sans que l'un ou l'autre ait tort.

Cela concerne les deux aspects. Ce qui compte, c'est ce que l'on représente pour les autres et la valeur de ses propres jugements, car la manière dont on juge est elle-même évaluée.

Évaluation

Convenu. L'attribution d'une valeur positive ou négative, d'une certaine intensité, en fonction de ses propres besoins et désirs.

Ce critère relève de l'évaluateur lui-même et ne prétend pas s'appliquer à autrui. Il s'agit d'un principe fondamental, qui s'applique également en dehors du domaine social.

Les notions de besoin et de désir ne sont pas redéfinies dans cet essai : elles renvoient à un ouvrage précédent de l'auteur.