



UNIL | Université de Lausanne

Unicentre

CH-1015 Lausanne

<http://serval.unil.ch>

Year : 2025

Peut-on punir le coupable si son jumeau siamois est innocent ? Les transitions évolutives d'individualité en éthique

Tanguy Robyr

Tanguy Robyr, 2025, *Peut-on punir le coupable si son jumeau siamois est innocent ? Les transitions évolutives d'individualité en éthique*

Originally published at : Mémoire de maîtrise, Université de Lausanne

Posted at the University of Lausanne Open Archive.
<http://serval.unil.ch>

Droits d'auteur

L'Université de Lausanne attire expressément l'attention des utilisateurs sur le fait que tous les documents publiés dans l'Archive SERVAL sont protégés par le droit d'auteur, conformément à la loi fédérale sur le droit d'auteur et les droits voisins (LDA). A ce titre, il est indispensable d'obtenir le consentement préalable de l'auteur et/ou de l'éditeur avant toute utilisation d'une oeuvre ou d'une partie d'une oeuvre ne relevant pas d'une utilisation à des fins personnelles au sens de la LDA (art. 19, al. 1 lettre a). A défaut, tout contrevenant s'expose aux sanctions prévues par cette loi. Nous déclinons toute responsabilité en la matière.

Copyright

The University of Lausanne expressly draws the attention of users to the fact that all documents published in the SERVAL Archive are protected by copyright in accordance with federal law on copyright and similar rights (LDA). Accordingly it is indispensable to obtain prior consent from the author and/or publisher before any use of a work or part of a work for purposes other than personal use within the meaning of LDA (art. 19, para. 1 letter a). Failure to do so will expose offenders to the sanctions laid down by this law. We accept no liability in this respect.

UNIVERSITÉ DE LAUSANNE

FACULTÉ DES LETTRES

Mémoire de Maîtrise universitaire ès lettres en Philosophie

Peut-on punir le coupable si son jumeau siamois est innocent ?
Les transitions évolutives d'individualité en éthique

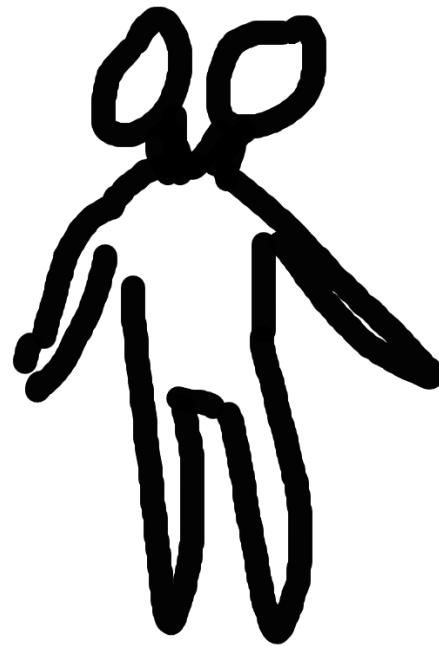
par

Tanguy Robyr

sous la direction du Professeur Christian Sachse

Session d'hiver 2025

Université de Lausanne
Faculté des lettres
Section de philosophie



Peut-on punir le coupable si son jumeau siamois est innocent ?

Les transitions évolutives
d'individualité en éthique

Mémoire de Master
Hiver 2025

Présenté par : Tanguy Robyr

Directeur : Christian Sachse

Experte : Simone Zurbuchen-Pittlik

Pour clarifier une phrase ou une idée, on a toujours besoin d'autrui. Je remercie mon directeur de mémoire, je remercie mes relectrices et relecteurs, et je remercie toutes et tous mes partenaires de débat. Grâce à votre aide, la réalité a résisté plus fort à mes premiers concepts préconçus.

Et si ce texte n'est plus un brouillon labyrinthique, inerte et perdu dans un coin de disque dur, c'est grâce à Zélie et son soutien infini.

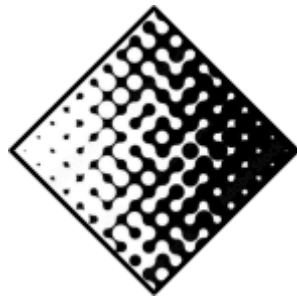


Table des matières

Table des figures	5
PROBLÉMATIQUE	6
1. LES BOUCLERS HUMAINS	8
1.1. Peut-on tuer un bouclier humain ?	8
1.2. Un problème insoluble	11
1.3. Un problème général du droit	13
2. LE PROBLÈME DES SIAMOIS	14
2.1. Peut-on emprisonner un innocent si son jumeau siamois est coupable ?	14
2.2. L'argument du standard de preuve	17
2.3. Théories philosophiques de la punition	19
3. LE CONSÉQUENTIALISME	21
3.1. Aperçu du conséquentialisme	21
3.2. Voile d'ignorance d'Harsanyi	23
3.3. Harsanyi appliqué aux tranches de vie	25
3.4. Harsanyi appliqué aux gènes	28
4. LES TRANSITIONS ÉVOLUTIONNAIRES D'INDIVIDUALITÉ	30
4.1. Définition fonctionnelle de Clarke	30
4.2. L'individualité est une poupée russe	31
4.3. L'utilitarisme comme transition évolutionnaire d'individualité	33
4.4. Peut-on sauver le conséquentialisme ?	34
4.5. Rawls comme une approximation limitée	37
5. LE RÉTRIBUTIVISME	39
5.1. Aperçu du rétributivisme classique	39
5.2. Théories hybrides conséquentialistes	41
5.3. Théories communicatives du mérite	45
6. LA LOI DE CONSERVATION DE LA PUNITION	49
6.1. Mérite de la communauté des prisonniers	49
6.2. Critique de cette loi de conservation	50
6.3. Au juste... qu'est-ce que le mérite ?	52

7. LA PHILOSOPHIE DE LA MONNAIE DE SIMMEL	54
7.1. Substantialisme : la monnaie comme commodité	54
7.2. Simmel : la monnaie comme symbole abstrait	56
7.3. La loi de conservation de la valeur monétaire	58
7.4. Transition fonctionnelle de la monnaie	59
7.5. Transition psychologique de la monnaie	61
8. UNE TRANSITION D'INDIVIDUALITÉ MONÉTAIRE	63
8.1. Une transition ontique de la monnaie	63
8.2. La transition évolutionnaire d'individualité de la monnaie	66
8.3. Histoire monétaire de la Suisse	68
8.4. Dernières remarques sur l'ontologie de la monnaie	72
9. UNE THÉORIE MONÉTAIRE DE LA PUNITION	74
9.1. Substantialisme : la valeur de la punition	74
9.2. Le shifgrethor : la punition comme symbole abstrait	75
9.3. Histoire de la rétribution en Suisse	77
9.4. La transition évolutionnaire d'individualité de la rétribution	79
CONCLUSION	83
Solution au problème des siamois	83
Limites de cette solution	85
Solution au problème des boucliers humains	88
BIBLIOGRAPHIE	91
Justice (théorie)	91
Justice (appliquée)	93
Biologie	95
Monnaie et économie	96
Divers	98

Table des figures

fig. I : Tableau des boucliers humains dans le droit humanitaire.....	10
fig. II : Argument du standard de preuve.....	18
fig. III : Schéma de l'individualité graduelle.....	32
fig. IV : Schéma de la punition graduelle.....	35
fig. V : Contre-argument du jeu à somme nulle.....	44
fig. VI : Schéma des fonctions de la monnaie.....	61
fig. VII : Trois alternatives à la "spiritualisation croissante" de la monnaie.....	64
fig. VIII : Schéma de la transition d'individualité monétaire.....	67
fig. IX : Évolution des masses monétaires en concurrence (XX ^e s.).....	70
fig. X : Schéma de la transition d'individualité de la punition.....	80
fig. XI : Schéma de la transition d'individualité de la justice punitive.....	87

PROBLÉMATIQUE

- *Well, what do I think?*
- *I think kill him.*
- *Oh, let's be nice to him!*
- *Oh, shut up.*

Three-Headed Giant, Monty Python's *The Holy Grail!*

Voici une expérience de pensée. Vous êtes juge de la Cour Suprême du Philosophistan et devant vous se présente Anne-Claire, une jumelle siamoise. De l'extérieur, elle ressemble à un être humain à deux têtes. Dans les faits, elle a certains organes internes à double, mais les autres en commun. Ses deux têtes ('Anne' et 'Claire') ont des caractères extrêmement différents. Il s'avère que Claire a commis un crime très grave, pour lequel vous donnez habituellement 10 ans de prison ferme. La police a prouvé *sans l'ombre d'un doute* que c'est Claire, seule et en toute discrétion, qui est responsable du crime. Pour Anne, la surprise est totale : elle ne savait rien de tout cela. La police a également prouvé *sans l'ombre d'un doute* qu'Anne est parfaitement innocente.

Or, Anne-Claire n'a qu'un seul corps : où une tête va, l'autre ira aussi. Il n'est pas possible de séparer son corps en deux par chirurgie, cela la tuerait. Pour des raisons juridico-techniques majeures, il n'est pas non plus possible d'aménager la peine. En tant que juge suprême, vous avez le devoir impératif d'établir et de rétablir la justice : que faites-vous ? Combien d'années de prison ferme donnez-vous à Anne-Claire ?

0 1 2 3 4 5 6 7 8 9 10

Appelons ce problème moral le *problème des siamois* ; l'objectif principal de ce travail sera d'y trouver une réponse objective. Pour y parvenir, je ferai une revue systématique des théories philosophiques de la punition, pour les mettre à l'épreuve de façon approfondie et pour les corriger de manière pertinente, le cas échéant. Ce travail procédera en huit étapes.

Tout d'abord (sections 1 et 2), je montre que le problème des siamois est analogue au problème des boucliers humains dans le droit de la guerre ; il s'agit d'un problème général du droit avec des occurrences réelles et régulières et sans solution confortable. Après une première tentative de solution (*argument du standard de preuve*), je conclus que remonter directement aux théories philosophiques de la punition est nécessaire.

Ensuite (section 3), je présente l'utilitarisme telle que décrit par John Harsanyi, qui confirme la pertinence de l'argument du standard de preuve. Néanmoins, nous verrons que son formalisme peut s'appliquer à la fois à la société, à la personne, au génome, etc. ce qui est devenu, depuis John Rawls, un axe de critique majeur de l'utilitarisme. Puis (section 4), je formalise cette critique éthique dans le vocabulaire de la philosophie de la biologie, et en particulier celui des *transitions évolutives d'individualité* telle qu'analysées par la philosophe Ellen Clarke. Je montre que l'utilitarisme traite la communauté comme la seule victime possible d'un crime. Je conclus par la nécessité d'également rendre justice à la personne humaine victime dudit crime.

Ainsi (section 5), je présente les théories rétributives de la punition. Malheureusement, elles échouent toutes à résoudre le problème des siamois de manière satisfaisante, car elles sont contradictoires, extrêmes ou stochastiques. Malgré tout, les plus prometteuses sont les théories communicatives du mérite. En particulier (section 6), si on ajoute une nouvelle règle à ces dernières (*loi de conservation de la punition*), on conclut qu'il faut répartir la peine des siamois en deux. Néanmoins, cette règle part du principe que toute autre peine lèse la communauté des prisonniers, ce qui est douteux en apparence.

Pour défendre la réalité de la loi de conservation de la punition, je procède en deux temps. D'abord (section 7), je m'appuie sur la philosophie de la monnaie de Georg Simmel. Son approche défend qu'une partie de la valeur de l'argent est relative à une masse monétaire. Par ailleurs (section 8), je montre que la théorie de Simmel se laisse formaliser, elle aussi, comme une transition évolutive d'individualité. Ensuite (section 9), pour appliquer les idées de Simmel à la punition, j'introduis le concept de *shifgrethor*, dont la valeur est relative à une masse punitive. Je montre que les théories communicatives du mérite impliquent le shifgrethor, et que le shifgrethor implique la loi de conservation de la punition. Par ailleurs, je montre que c'est la dénégation du shifgrethor qui empêche les autres théories rétributives de résoudre le problème des siamois.

En conclusion, je montre que, dans notre société, une peine juste au problème des siamois tourne à peu près autour de deux à quatre ans de prison ferme, sur les dix ans mérités par le coupable. Par ailleurs, je propose des pistes de solutions pour le problème des boucliers humains.

1. LES BOUCLIERES HUMAINS

1.1. Peut-on tuer un bouclier humain ?

Sous l'Irak de Saddam Hussein, les soldats avaient pour consigne d'utiliser tous les moyens nécessaires à la victoire. Durant la guerre du Golfe de 1990, le gouvernement irakien s'empare de citoyens étrangers, pour les utiliser comme boucliers humains¹. Durant l'invasion américaine de 2003, le commandement irakien déplace même des bus remplis de femmes et d'enfants dans les zones de conflits, pour couvrir les déplacements et les attaques de ses troupes². En réaction, l'Organisation des Nations unies (ONU) condamne ces violations graves du droit humanitaire.

Autre contexte : le 3 novembre 2006, dans le cadre du conflit israélo-palestinien, des soldats israéliens piègent une trentaine de militants du Hamas dans une mosquée. Le Hamas lance alors un appel radiophonique pour que des femmes convergent en nombre vers la mosquée et cachent les combattants parmi elles. Alors que les militants s'échappent déguisés parmi la foule d'une cinquantaine de femmes, la troupe israélienne ouvre le feu : une palestinienne est tuée et dix autres sont blessées (la plupart à hauteur des pieds)³.

Dans ces différentes situations, les civils jouent le rôle de *boucliers humains*. Or, « la présence de boucliers humains nécessite la pratique d'une guerre 'vertueuse' »⁴ qui

¹ SCHMITT Michael N., « Human shields in international humanitarian law », *The Columbia journal of transnational law*, vol. 47, n° 2, 2009, p. 19.

² *idem*.

³ McCARTHY Rory, « Israeli troops open fire on women outside mosque », *The Guardian*, 03.11.2006, cité dans SCHMITT Michael, « Human shields in international humanitarian law », *art. cit.*, p. 39.

⁴ BARGU Banu, « Human shields », *Contemporary Political Theory*, vol. 12, n° 4, 2013, p. 282.

cherche à épargner autant que possible les civils⁵. Concrètement, le droit international humanitaire octroie des statuts spéciaux aux civils et aux prisonniers de guerre, pour les distinguer des combattants. Il arrive, fortuitement, que les deux populations se mélangent ; dans ce cas, il fait consensus que la priorité est de ne pas causer de victimes collatérales en trop grand nombre (cela serait injustifiable pour le sens moral des militaires, pour les tribunaux et pour les médias).

Il devient alors criminel d'exploiter la droiture de son adversaire à son avantage en cherchant délibérément à se protéger derrière des civils. En effet, comme l'explique Michael N. Schmitt, professeur de droit international au *U.S. Naval War College*, « d'un point de vue militaire pratique, un civil qui prend les armes pourrait bien être *moins* efficace pour dissuader ou défendre contre une attaque que le civil qui fait bouclier »⁶. On peut dire qu'un civil se transforme un bouclier humain lorsque "être une victime collatérale"—ou plutôt, menacer de le devenir—devient sa fonction principale, le cœur de sa participation à l'effort de guerre d'un des camps.

Comment réagir face à cette situation ? Michael N. Schmitt distingue trois approches, qui sont trois rapports à la perfidie de celui qui utilise des boucliers humains. La première approche, c'est de traiter les boucliers humains comme des civils ordinaires. Cela ne veut pas dire que leur vie et leur intégrité physique est garantie : pour parvenir à un objectif militaire, on autorise de toute façon N victimes collatérales potentielles. C'est l'approche défendue par les États-Unis et, en général, par la grande majorité des experts du droit humanitaire⁷. La deuxième approche consiste à « tenir compte » des décisions illégales du commandement ennemi. Dans les faits, pour le même objectif militaire, on autorise cette fois $N*k$ victimes collatérales ; les boucliers humains sont des "civils au rabais", « faute d'un meilleur terme »⁸. C'est notamment ce que tolèrent le Royaume-Uni

⁵ Plus exactement, le droit de la guerre autorise quelques 'victimes collatérales', tant que des objectifs militaires légitimes le justifient, d'autant plus si le brouillard de guerre les rend inévitables : cela s'appelle le principe de proportionnalité. Plus que de tuer des innocents, c'est violer le principe de proportionnalité qui est un crime de guerre. Mais cela ne change rien au raisonnement général présenté ici.

⁶ « Indeed, from a practical military point of view, a civilian who takes up arms may well be *less* effective in deterring or defending against attack than one who shields. »

SCHMITT Michael, « Human shields in international humanitarian law », *art. cit.*, p. 41.

⁷ *ibid.*, p. 50.

⁸ « [This approach] agrees that involuntary human shields retain immunity from attack, but suggests that they should, for the lack of a better term, "be discounted" when calculating incidental injury for proportionality and precautions in attack purposes. »

Ibid., p. 51.

et le Comité International de la Croix-Rouge. Ce genre d'approche qui construit une hiérarchie des civils n'est pas si rare. Par exemple, bombarder une usine d'armement est toléré malgré le nombre de victimes collatérales, notamment car les ouvriers de ces usines-là ont un statut "au rabais", comparé aux ouvriers qui travaillent dans d'autres industries⁹. Dans la troisième approche, la plus extrême, on ignore tout simplement la présence des boucliers humains. Ils ne perdent pas exactement leur statut de civil ; on déplace simplement l'entière responsabilité de leur mort vers l'ennemi. Cette interprétation a été largement exploitée par Israël après ses opérations militaires dans la bande de Gaza en 2014, à la fois devant l'ONU et sur les réseaux sociaux¹⁰. Cette approche-là est encadrée par le droit des représailles, qui autorise à violer une loi de la guerre pour forcer son ennemi à la respecter à nouveau. Mais cette exception est également très controversée¹¹.

STATUT DES BOUCLIERES HUMAINS	civils	civils "au rabais"	participants au conflit	<i>l'ennemi a toute la responsabilité</i>
FORTUIT	consensuel et inscrit dans le droit			
SOUS LA CONTRAINTE	« <i>dominate among international humanitarian law experts</i> »	« <i>enjoys significant support</i> »		« <i>extreme view</i> »
VOLON-TAIRES	« <i>enjoys strong support</i> »	« <i>enjoys significant support</i> »	approche de Schmitt	

fig. 1 : Tableau des boucliers humains dans le droit humanitaire¹².

Mais Michael Schmitt veut ajouter une distinction supplémentaire, entre boucliers humains contraints et boucliers humains volontaires. Ce n'est pas une distinction présente dans le droit de la guerre ; ses rédacteurs n'envisageaient pas vraiment une situation comme celle des boucliers humains palestiniens que j'ai présentée plus tôt. Les

⁹ *Ibid.*, p. 53.

¹⁰ GORDON Neve et PERUGINI Nicola, « The politics of human shielding: On the resignification of space and the constitution of civilians as shields in liberal wars », *Environment and Planning D: Society and Space*, vol. 34, n° 1, 2016, p. 170.

¹¹ SCHMITT Michael, « Human shields in international humanitarian law », *art. cit.*, p. 50.

¹² Informations tirées de SCHMITT Michael, « Human shields in international humanitarian law », *art. cit.*

soldats israéliens ont-ils mal agi en ouvrant le feu ? La plupart des interprètes du droit les condamneraient. Cependant, Michael Schmitt arrive à l'interprétation inverse : les boucliers humains volontaires, qui agissent de leur plein gré et en connaissance de cause, sont des *participants* au conflit et agissent sur le champ de bataille pour influencer son issue : en tant que tels, ils ne sont plus des civils. Ils ne peuvent pas être des victimes collatérales. Le tableau ci-dessus (fig. 1) résume les différentes positions et leur popularité respective dans les années 2000.

1.2. Un problème insoluble

Le but de ces couches de distinctions successives, c'est de maintenir une binarité la plus nette possible entre zone civile et théâtre des opérations. Pour ce faire, la droiture morale consiste à chercher à *collapser* toute situation intermédiaire dans une de ces deux catégories antithétiques, soit par une action sur le terrain, soit dans l'analyse conceptuelle de la situation. À l'inverse, des antagonistes cherchent constamment à déstabiliser cette binarité pour leur bénéfice. Le droit humanitaire est un système complexe et instable qui, pour survivre, doit être maintenu par un afflux constant de législation et de concepts ; ou, dit plus simplement, le droit de la guerre est dialectique.

Sans surprise, la distinction entre boucliers humains "volontaires" et "involontaires", peut donc à son tour être exploitée et brouillée par des agents sans scrupules. Les philosophes Judith Butler et Vasuki Nesiiah ont notamment relevé depuis une décennie que cela « fait partie de la stratégie militaire » d'Israël¹³. En effet, dans le discours militaire et étatique israélien, d'une part, les habitants, les écoles et les hôpitaux de la ville de Gaza sont présentés comme des boucliers humains pour les opérations militaires du Hamas¹⁴ ; d'autre part, toujours dans ce discours, ces mêmes habitants ont

¹³ BUTLER Judith, « Human shields », *London Review of International Law*, vol. 3, n° 2, 2015, p. 228.

¹⁴ « Invoking the prohibition of human shields, Israel argued that they were targeting the enemy that happened to be operating in a world that included residential apartment buildings, schools and hospitals. In other words, the logic of Israeli military discourse on human shields is that if any actor in Gaza were fighting a national liberation struggle, the legal principle of human shields requires that that fight take place outside of Gaza, outside a city of residential apartment buildings, schools and hospitals »

NESIAH Vasuki, « Human shields, human heresies », *International Politics Reviews*, vol. 10, n° 1, 2022, p. 39.

choisi collectivement de devenir des boucliers humains, en soutenant le Hamas¹⁵. De proche en proche, tous les civils deviennent des participants actifs au conflit. Par exemple, dans un cas d'enfants palestiniens tués par des bombardements israéliens, le premier ministre d'Israël Benyamin Netanyahu s'est vigoureusement plaint de ces 'morts télégraphiques'. Judith Butler interprète ainsi ces protestations : « l'implication claire c'est que le Hamas ou, en fait, les Palestiniens en général, envoient leurs enfants jouer sur la plage ou sur une place de jeu afin qu'ils soient tués—que c'est 'le genre de personne qu'ils sont' et que la mort mise en scène de leurs enfants est ensuite utilisée contre Israël, sans justification »¹⁶.

Sans doute que des concepts supplémentaires sont nécessaires, pour à nouveau distinguer clairement civils et militaires dans le droit... mais pour un temps seulement. Dans sa critique d'un livre récent sur le sujet des boucliers humains, Vasuki Nesiah conclut en se demandant « vaut-il la peine de sauver le droit humanitaire international ? »¹⁷. C'est une question légitime. Cependant, comme j'ai tenté de le souligner, tout système de droit qui veut opérer une distinction entre 'zone civile' et 'théâtre d'opération militaire' sera confronté au même problème, à la même lutte interminable au sein du droit. Il n'y a pas de raison de penser qu'une des parties puisse un jour définitivement s'imposer dans cette 'course aux armements évolutionnaire' des concepts—pour reprendre le mot de Daniel Dennett¹⁸. Ou, à nouveau dit plus simplement, il n'y a pas de raison que la dialectique se termine. Faut-il donc renoncer tout court à distinguer civils et militaires ? Et peut-on remplacer cette distinction par une autre ?

Si une réponse évidente existait, elle ferait déjà consensus depuis longtemps. Mais si une réponse existe tout de même, alors c'est sûrement par un chemin détourné et inattendu

¹⁵ « At one point, the Israeli government defended itself against the allegation of war crimes by pointing out that the general population had voted for Hamas, suggesting, it seemed, that those who voted for Hamas not only endorsed its role in the conflict with Israel, but by the very act of voting or voting in that way, became active participants in the conflict themselves. Voting is conceived, then, as signing up for the armed forces. »

BUTLER Judith, « Human shields », *art. cit.*, p. 240.

¹⁶ « Yet the clear implication is that Hamas or, indeed, Palestinians more generally, send their children out to play on the beach or in the playground in order that they will be killed—that this is 'the kind of people they are' and that the staged death of their children is then used unjustifiably against Israel. »

Idem.

¹⁷ NESIAH Vasuki, « Human shields, human heresies », *art. cit.*, p. 41.

¹⁸ DENNETT Daniel C., *Freedom Evolves*, London : Penguin Books, 2004, p. 21.

qu'il faut la dénicher. Ce qu'il faudrait, c'est pouvoir décomposer le problème des boucliers humains en plusieurs problèmes plus simples. C'est plus facile à dire qu'à faire. Dans tous les cas, il faut commencer par formaliser le problème de la manière la plus générale possible.

1.3. Un problème général du droit

Comme nous l'avons vu, dans le droit humanitaire, le *problème des boucliers humains* tient à trois choses. La première, c'est que le droit distingue radicalement le statut de plusieurs individus différents (civil/combattant). La deuxième, c'est qu'en pratique, le bras armé des États ne peut pas les distinguer : il n'existe pas de frappe assez 'chirurgicale' pour cibler chacun selon son statut. La troisième, c'est le *mens rea* des agents, leur disposition d'esprit : le soldat et/ou le civil peuvent agir sciemment pour que le droit de la guerre se retrouve dans l'embarras et pour chercher à exploiter la clémence de l'ennemi.

On peut généraliser la première étape du problème : distinguer des statuts, ce n'est pas spécifique au droit humanitaire, mais c'est le fondement de tous les types de théories morales et juridiques. Pensons, par exemple, à une distinction aussi universelle que innocent/coupable. Qui plus est, tous ces systèmes moraux et juridiques visent à avoir un effet sur le monde : par conséquent, tous prévoient l'existence d'un bras armé, pour imposer cet ordre juridique, même aux récalcitrants.

On peut généraliser la deuxième étape du problème : le bras armé d'une théorie morale (la punition ou la récompense) n'est jamais parfait non plus, et cette imperfection est parfaitement prévisible. Une amende ou une récompense pécuniaire, par exemple, ne visent que des porte-monnaies ; que se passe-t-il si plusieurs personnes partagent le même compte en banque ou les mêmes dettes ? Cette idée n'est pas si exotique, puisqu'elle est au fondement du contrat de mariage. Quant au système carcéral, il ne peut enfermer que des corps. Que se passe-t-il si plusieurs personnes partagent le même corps ? C'est une situation plus rare, mais pas impossible. On peut penser aux jumeaux siamois et aux personnes atteintes d'un trouble dissociatif de l'identité, qui y seront confrontés toute leur vie ou presque.

On peut généraliser la troisième étape du problème : par définition ou presque, toutes les théories morales ne portent que sur des agents libres. De même, par définition ou presque, les théories morales ne sont pas “libres”, mais des règles prévisibles et déterminées. Il y a asymétrie : l’un peut s’adapter, l’autre pas. Dès que les conditions matérielles sont réunies (comme être un jumeau siamois ou souffrir d’un trouble dissociatif de l’identité), l’agent est libre de mettre le système moral ou juridique dans l’embarras—ce n’est qu’une question de volonté.

Tout cela signifie une chose : le problème des boucliers humains est un problème général. Pensez-vous, comme Nesiah, qu’il sonne le glas du droit humanitaire international ? Prenez garde, car il peut emporter dans sa tombe les concepts mêmes de loi, de droit et de chose publique. Mais l’inverse est vrai aussi ; car si nous parvenons à résoudre ce problème de manière générale, à renforcer la loi, le droit et la chose publique, alors le droit de la guerre en ressortira plus fort lui aussi.

2. LE PROBLÈME DES SIAMOIS

2.1. Peut-on emprisonner un innocent si son jumeau siamois est coupable ?

Comment pourrions-nous résoudre le problème des boucliers humains ? Existe-t-il vraiment une solution univoque, satisfaisante et prévisible, mais qu’aucun agent libre ne peut exploiter à son avantage ? Pour le savoir, il faut s’attaquer à une version plus simple du problème, sans le contexte, la complexité morale et empirique et les intérêts croisés des guerres et des luttes armées. Commençons donc par un problème de justice pénale, le cas des jumeaux siamois.

Au XIX^e siècle, deux frères du Siam (actuelle Thaïlande) font carrière aux États-Unis dans le cirque. Eng et Chang Bunker (1811–1874) sont en effet des jumeaux dont le corps n’est pas complètement séparé : au niveau de la hanche, ils sont reliés par un pont de tissus fibreux. Après 63 ans de vie commune, avoir amassé une grande fortune, de nombreux esclaves et la nationalité américaine, les frères Bunker meurent à quelques

heures d'intervalle, auprès de leurs femmes respectives et de leurs nombreux enfants. On doit à leur célébrité le terme de 'jumeaux siamois', mais aussi une anecdote judiciaire troublante. Un jour qu'un médecin indiscret voulait ausculter les jumeaux en public, Chang Bunker le frappa. Le médecin porta plainte : il voulait faire arrêter Chang (et donc Eng Bunker, à son corps défendant). Le juge n'en fit rien, et expliqua même au plaignant qu'il risquait d'être arrêté pour fausse accusation contre le jumeau innocent ! Les frères Bunker repartirent libres¹⁹. Autre exemple : au xx^e siècle, et par deux fois, Lucio Godina se retrouva devant un juge pour avoir gravement violé le code de la route, au volant de sa voiture conçue sur mesure. Les deux fois, parce que son jumeau siamois Simplicio Godina était un simple passager innocent, les juges les laissèrent libres²⁰. Par leur naissance et la forme de leurs corps, Chang Bunker et Lucio Godina ont échappé à la loi. Était-ce justifié ? Pour des crimes en série plus graves, serait-ce toujours justifié ?

De la même manière que le droit humanitaire se fonde sur la séparation morale et géographique entre militaires et civils, notre système légal part du principe que l'innocent et le coupable ne partagent pas le même corps. Et le problème que les jumeaux siamois posent au droit pénal est inéluctable, plus inéluctable encore que celui des boucliers humains en droit international humanitaire. En effet, en dernier recours, le droit humanitaire menace de punir ceux qui violent ses commandements : les criminels de guerre ont beau se camoufler au sein des civils, la communauté internationale est capable de les en extraire, pour les emprisonner ou les pendre. En temps normal, aucun bouclier humain n'est éternel et infaillible.

Par contraste, comment punir le jumeau coupable sans, du même coup, punir le jumeau innocent ? Quand le magazine *Explainer* posa cette question en 2009, 6000 personnes tentèrent puis échouèrent, indépendamment, à trouver une solution pratique ou technique à ce problème²¹. Même une lourde amende, par exemple, impacte le niveau de vie commun des deux jumeaux, en plus d'être une peine inadaptée au crime commis. Il n'est pas possible de rendre le système carcéral plus chirurgical et de ne punir que le siamois coupable. Plus fort qu'un bouclier, les jumeaux siamois sont un cas d'armure humaine.

¹⁹ DAVIS Colleen, « Complicity, Crime and Conjoined Twins », *Alternative Law Journal*, vol. 42, n° 1, 2017, p. 19.

²⁰ *Idem.*

²¹ *Ibid.*, p. 18.

Il n'est pas non plus possible de rendre le droit moins chirurgical et d'établir systématiquement une responsabilité partagée. Dans un article de 2017, la juriste australienne Colleen Davis détaille le problème, avec quelques réserves. Certes, il est parfois possible de condamner l'autre siamois comme complice ; cependant, la simple présence sur les lieux du crime ne suffit pas. Au minimum, dans la loi australienne en tout cas, il faut que le complice présumé encourage délibérément et volontairement le coupable par un signe, des mots, une expression du visage²². Qui plus est, il existe de nombreux crimes où les simples témoins n'ont pas l'obligation légale d'intervenir²³. Pour certains types de crime, le jumeau innocent peut ne même pas être témoin—si son jumeau empoisonne quelqu'un, par exemple, ou s'il commande un assassinat.

Le problème s'applique d'ailleurs à la punition dans d'autres contextes. Ainsi, quand une des sœurs Ruthie et Verena Cady (1984–1991) était punie par ses parents (par exemple, en devant rester debout au coin sans bouger), l'innocente n'y voyait pas nécessairement de problème. Il arrivait que Verena dise à sa soeur « tu ne peux pas encore partir, Ruthie, tu es toujours punie »²⁴.

Si je précise tout cela, c'est pour montrer qu'il s'agit d'un problème moral et philosophique profond, pas d'un problème technique. Le cas des siamois met à l'épreuve le droit parce qu'il ébranle l'une des fondations de ce dernier : notre *théorie naïve de la punition*. Cela ne veut pas dire qu'une solution technique est impossible et que le droit ne peut pas s'adapter. Ce que cela veut dire, c'est que le droit pourrait s'adapter de plein de manières différentes, et que ces différentes solutions découlent de différentes théories de la punition. Pour l'illustrer, prenons une certaine interprétation du problème des siamois—l'argument du standard de preuve—et montrons qu'il se rattache à une théorie particulière de la justice.

²² *Ibid.*, pp. 21–22.

²³ *Ibid.*, p. 22.

²⁴ « What about disciplining attached children ? “When one misbehaves”, Marlene admitted, “of course the other has to suffer some of the consequences too. Ruthie has had to go stand in the corner a few times for being naughty, and Verena just goes along with her. She’ll say ‘You can’t leave the corner yet, Ruthie, you’re still naughty’, even though she is there too. She just understands that Ruthie has got to be punished and that’s the way it is.” »

CADY Marlene, « The pure joy of being alive », *People Weekly*, vol. 32, 03.07.1989, p. 70, cité dans DREGER Alice Domurat, *One of Us: Conjoined Twins and the Future of Normal*, Cambridge Mass : Harvard University Press, 2004, p. 34.

2.2. L'argument du standard de preuve

Au fond, qu'est-ce que le problème des siamois ? C'est une situation qui nous force à commettre une *erreur judiciaire*, que ce soit en condamnant un innocent avéré ou en libérant un coupable avéré. Plus exactement, il nous demande quel type d'erreur judiciaire nous préférons. Pour chaque année de prison subie par l'innocent, à combien d'années de prison le coupable doit-il échapper ? De manière analogue l'incertitude épistémique force—statistiquement—les institutions pénales à commettre des erreurs judiciaires. C'est hélas une nécessité du métier.

Néanmoins, il est possible de contrôler ce ratio d'erreur judiciaire, en modifiant le *standard de preuve* en vigueur. Par exemple, si je me contente d'indices anecdotiques pour rendre mon jugement, peu de criminels m'échapperont ; mais je ferai beaucoup de victimes collatérales. À l'inverse, si je demande la Lune à chaque procès, les innocents ne risquent rien ; mais l'écrasante majorité des crimes resteront impunis. Or, la préférence des juges est célèbre et très claire : « mieux vaut que dix coupables s'échappent, qu'un seul innocent souffre »²⁵ avait écrit le juriste anglais William Blackstone (1723–1780). Pour chaque année de prison subie par un innocent, les institutions ont choisi à combien d'années de prison les coupables vont échapper. Il suffit d'appliquer ce même ratio au problème des siamois : pour chaque jour de prison subi par l'innocent, les institutions ont choisi à combien de jours de prison le coupable doit échapper.

Concrètement, maintenant, à quelle peine peuvent s'attendre les siamois si on applique ce ratio ? Si une personne est prête à libérer une cinquantaine de coupables pour sauver un innocent—un ratio de (1 : 50)—elle condamnerait tout de même le duo à une peine de deux mois et demi (pour un crime qui vaut 10 ans de réclusion). Même Blackstone n'est pas aussi magnanime avec son ratio de (1 : 10). Dans le même ordre d'idée, il est extrêmement difficile de condamner le duo à la peine maximale. À titre de comparaison, le haut commandement nazi défendait un ratio de (50 : 1)²⁶ ce qui revient à condamner le duo à une peine de 9 ans et neuf mois et demi. Le seul interdit absolu, c'est de gracier

²⁵ « ... better that ten guilty persons escape, than that one innocent suffer »

BLACKSTONE William, *Commentaries on the laws of England*, vol. 4, Oxford : Clarendon Press, 1766.

²⁶ En octobre 1941, par exemple, un petit groupe de résistants communistes assassine Karl Hotz (1877–1941), le gouverneur militaire de la ville de Nantes. Puisque les tueurs se sont échappés, le haut commandement nazi veut fusiller de 50 à 150 otages français ; au final, 48 périrent.

GILDEA Robert, « Resistance, Reprisals and Community in Occupied France », *Transactions of the Royal Historical Society*, vol. 13, [Cambridge University Press, Royal Historical Society], 2003, p. 166.

les deux siamois ou de les condamner à la peine du coupable, qui représentent un ratio de $(1 : \infty)$ ou de $(\infty : 1)$, respectivement. Cela reviendrait à renoncer purement et simplement au concept de standard de preuve et, par généralisation, à celui de preuve et donc de justice. Voilà pour l'aperçu (fig. 11).

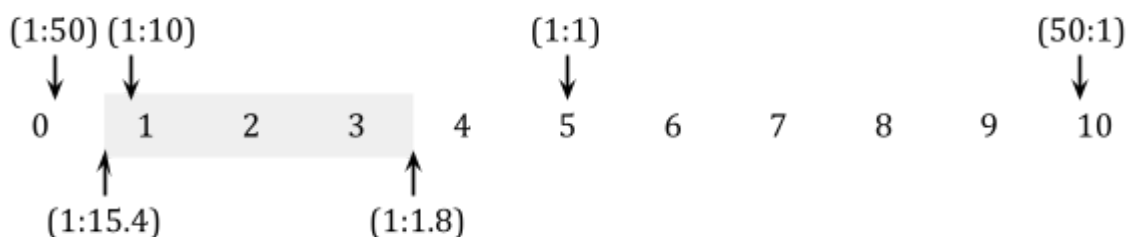


fig. 11 : Argument du standard de preuve.

Passons maintenant à la pratique. Comment mesurer le standard de preuve en vigueur dans nos institutions ? Pour cela, il faudrait connaître le nombre d'innocents en prison, et le nombre de coupables en liberté. Comment mesurer ces données ? C'est un problème difficile et un champ de recherche assez pauvre. Un autre problème, c'est que ce standard de preuve change selon la gravité des accusations : en général, plus la punition attendue est sévère, plus le standard est élevé. Ici, nous choisissons de nous concentrer sur le standard pour des crimes aussi graves qu'un meurtre. Malgré tout, on peut proposer un ordre de grandeur.

Commençons par compter le nombre de criminels qui échappent à toute punition. D'abord, en général, dans les pays occidentaux d'Europe, entre 2 à 23% des crimes graves restent parfaitement irrésolus (la police a échoué à trouver quelqu'un à accuser)²⁷. Notons que certains crimes sont pris pour des accidents et des disparitions ; et que des accidents ou des disparitions sont pris pour des crimes. Faute de pouvoir quantifier l'erreur policière, nous partons du principe que ses effets se compensent dans les calculs. Ensuite, on sait que parmi les accusés, un certain nombre sont acquittés, à peu près 2%²⁸ dont la plupart sont innocents²⁹. Dans tous les cas, partons du

²⁷ LIEM Marieke *et alii*, « Homicide clearance in Western Europe », *European Journal of Criminology*, vol. 16, n° 1, 2019, p. 82.

²⁸ « there are more than fifty judgments for guilt for every not guilty verdict »

GIVELBER Daniel, « Lost Innocence: Speculation and Data about the Acquitted », *American Criminal Law Review*, vol. 42, n° 4, Fall 2005, p. 1167.

²⁹ « The most likely reason for an acquittal in such circumstances, I suggest, is that the defendant is innocent. »

Ibid., pp. 1169-1170

principe qu'en cas d'acquittement le crime reste irrésolu³⁰. Sur 1000 meurtres, cela représente donc entre 245 et 40 crimes impunis.

Connaître le nombre d'innocents en prison est plus difficile. La seule approche assez fiable pour être mentionnée, c'est l'analyse ADN a posteriori. Par ce moyen, on sait que parmi les américains condamnés à mort entre 1973 et 1989, 2.3% à 5% étaient innocents³¹. Faute de mieux, extrapolons un peu, et imaginons que pour tous les meurtres, entre 2.3% et 5% des années d'emprisonnement sont purgées par des innocents et des innocentes. Qui plus est, partons du principe que quand un innocent est condamné, en moyenne, le vrai coupable court toujours.

Fort de ces nombres très approximatifs, nous pouvons estimer une échelle de grandeur de ce fameux *standard de preuve*. En temps normal, pour chaque année de prison subie par un innocent pour un crime grave, les coupables ont échappé de 1,8³² à 15,4³³ années de prison. Nous sommes donc probablement un peu plus sévères que le principe de Blackstone. Dans tous les cas, ces nombres nous permettent de répondre au problème des siamois. D'après l'estimation haute, il faut condamner le duo à moins de 35% de la peine du coupable, car $(35 : 65) \cong (1 : 1,8)$. D'après l'estimation basse, il faut condamner le duo à un peu plus de 6% de la peine du coupable, car $(6 : 94) \cong (1 : 15,4)$. En clair, pour un crime qui mériterait normalement dix ans de réclusion, une institution judiciaire condamnerait probablement les siamois à une peine qui tourne autour de 7 mois et 3 ans et demi de prison (fig. 11).

2.3. Théories philosophiques de la punition

Magnifique ! En apparence, le problème des siamois est donc clos. Cependant, nous avons passé sous silence deux objections majeures. Premièrement, ce n'est pas une

³⁰ Cette hypothèse est fautive, mais nécessaire. Dans certains cas, le juge acquitte car aucun crime n'a été commis par personne. Dans d'autres, c'est parce que le vrai coupable est condamné que l'acquittement est possible.

³¹ GROSS Samuel R. et O'BRIEN Barbara, « Frequency and Predictors of False Conviction: Why We Know So Little, and New Data on Capital Cases », *Journal of Empirical Legal Studies*, vol. 5, n° 4, 2008, p. 947.

³² On cherche le ratio (INNOCENTS_EN_PRISON *max* : COUPABLES_LIBRES *min*). Sur 1000 crimes, il y a 40 crimes impunis minimum. Sur les 960 crimes punis, il y a 48 innocents maximum (5%). On obtient donc le ratio $(48 : 40+48) \cong (1 : 1,83)$.

³³ On cherche le ratio (INNOCENTS_EN_PRISON *minimum* : COUPABLES_LIBRES *maximum*). Sur 1000 crimes, il y a 245 crimes impunis maximum. Sur les 755 crimes punis, il y a 17 innocents minimum (2,3%). On obtient donc le ratio $(17 : 245+17) \cong (1 : 15,4)$

institution judiciaire qui condamne les siamois, c'est une juge. De même, ce ne sont pas les juges qui décident du standard de preuve adéquat : ce dernier émerge à l'échelle institutionnelle. *Pourquoi est-ce que les juges devraient se soumettre aux préférences institutionnelles ?* Deuxièmement, les préférences institutionnelles de nos États occidentaux sont peut-être très immorales. *Pourquoi est-ce que les préférences institutionnelles devraient se soumettre aux pratiques actuelles ?*

Répondre à ces questions nécessite une réflexion philosophique. Nous pouvons préciser un peu mieux nos questions. En effet, dans un article de 2010, Rizzolli et Saraceno pensent avoir montré que, lorsqu'il s'agit de choisir un standard de preuve, le système judiciaire se comporte de manière *utilitariste*—même si les individus, eux, n'utilisent pas de raisonnements utilitaristes pour se justifier³⁴. Sans épiloguer, le raisonnement met dans la balance les coûts sociaux et financiers du système policier, judiciaire et carcéral, par contraste au risque de victimes de récidives, ce qui représente également des coûts sociaux et financiers pour la communauté. Or, l'utilitarisme est une théorie morale. La question centrale devient donc *pourquoi faudrait-il utiliser un raisonnement utilitariste pour punir les jumeaux siamois, par contraste avec d'autres raisonnements moraux ?*

Jusqu'ici, la question que nous nous sommes posées, implicitement, c'est « qu'est-ce qui distingue une punition juste d'une punition injuste ? ». Néanmoins, quand les philosophes se penchent sur cette question, ils réalisent rapidement qu'il faut d'abord répondre à une question plus générale : « est-ce qu'une punition *peut* être juste ? ». En effet, punir quelqu'un, c'est (a) lui infliger volontairement un tort, afin de (b) condamner un comportement passé de cette personne³⁵. Ces deux propriétés combinées semblent difficiles à concilier. D'après le théoricien du droit H.L.A Hart (1907–1992), il faudrait même décomposer cette nouvelle question en trois : « pourquoi créer et maintenir des institutions punitives ? », « qui peut être puni ? » et « combien de punition faut-il donner ? »³⁶. Nous sommes partis du principe que, pour un crime donné, un coupable mérite d'être puni de 10 ans de prison et qu'un innocent en mérite 0 ; cependant, cette affirmation n'a rien de trivial. Une mauvaise théorie de la punition

³⁴ RIZZOLLI Matteo et SARACENO Margherita, « Better that ten guilty persons escape: punishment costs explain the standard of evidence », *Public Choice*, vol. 155, n° 3–4, 2013, pp. 395–411.

³⁵ HOSKINS Zachary et DUFF Antony, « Legal Punishment », in ZALTA Edward N. et NODELMAN Uri (dir.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, 2024.

³⁶ *Idem*.

échouerait à nous expliquer pourquoi cette situation de départ est possible, tout simplement.

Pour tenter de répondre à ces quelques questions, il y a deux grandes familles de doctrines : le conséquentialisme (qui contient notamment l'utilitarisme) et le rétributivisme. Le conséquentialisme demande que la juge justifie, en fin de compte, sa décision d'après les conséquences de cette dernière. Plus exactement, la meilleure décision, c'est celle qui mène au meilleur futur. D'après Jeremy Bentham (1748–1732), un des conséquentialistes les plus connus, « en soi, punir est toujours mal. [...] [Mais] s'il fallait admettre de punir quoi que ce soit, ce serait à condition que cela promette d'éviter un mal plus grand encore »³⁷. Qu'est-ce qu'un malheur plus grand encore ? Comment faut-il le mesurer ? Ces questions génèrent de nombreux désaccords entre les théoriciens du conséquentialisme. Nous verrons dans quelle mesure et à quel prix elles peuvent résoudre le problème des siamois.

Le rétributivisme est la seconde approche classique pour résoudre ce problème. Il présente le problème autrement : à la fin d'une décision juste, chacun a eu *ce qu'il mérite* par rapport à ses actions passées. Magnifique ! Mais qu'est-ce que le mérite ? Peut-il se mesurer ? Ces questions forment le fond d'un grand nombre de désaccords entre tous les théoriciens de la rétribution ; pour cette raison, il existe une grande diversité de doctrines ici aussi, que nous tenterons d'explorer et de mettre à l'épreuve par la suite. Mais pour l'instant, explorons le conséquentialisme et, en particulier, l'utilitarisme.

³⁷ « [...] all punishment in itself is evil. [...] if it ought at all to be admitted, it ought only to be admitted in as far as it promises to exclude some greater evil ».

BENTHAM Jeremy, *An Introduction to the Principles of Morals and Legislation*, Londres : W. Pickering, 1780–1789, ch. XIII.2.

3. LE CONSÉQUENTIALISME

3.1. Aperçu du conséquentialisme

Il existe plusieurs versions du conséquentialisme. Toutes partagent cette idée centrale : la meilleure décision est celle qui mène au meilleur futur. Ainsi, il faut pondérer les conséquences *et seulement les conséquences* d'un choix. Cela n'implique pas d'être altruiste : un être machiavélique est conséquentialiste, par exemple. Plus largement, on peut distinguer deux types de conséquentialisme : *agent-relative* et *agent-neutral*. Comme exemples de conséquentialisme 'relatif', on peut citer l'égoïsme moral ou les conséquentialistes relativistes, comme Sen, Broome ou Portmore³⁸. Cependant, dans notre cas, nous nous intéressons à la décision d'une juge : par définition, cette dernière doit agir comme un agent neutre et désintéressé (même si elle fait cela uniquement pour recevoir son salaire, égoïstement). Par la suite, nous ne parlerons donc que du conséquentialisme 'neutre'.

L'exemple paradigmatique du conséquentialisme *agent-neutral* est l'utilitarisme dit 'classique', qui remonte à Jeremy Bentham (1748–1732). Les utilitaristes cherchent à maximiser, en moyenne, le bonheur (ou l'absence de souffrance ou les aspirations psychologiques, etc..) de tous les êtres sentients. C'est une théorie qui mène naturellement Bentham et ses suiveurs au féminisme, à l'antiracisme, au véganisme, etc. Cependant, contre cet utilitarisme classique une multitude de contre-exemples problématiques ont été imaginés au fil des années (il arrive que la chose morale à faire, par exemple, ce soit de découper des patients sains pour greffer leurs organes à d'autres). En réaction, une myriade de modifications à l'utilitarisme de Bentham ont été proposées, qui forment autant de versions du conséquentialisme. Pour la plupart, ces changements n'ont aucune incidence sur le problème des siamois. Par la suite, partez du principe que nous analysons toutes ces versions d'un même mouvement.

Il vaut tout de même la peine de mentionner l'*utilitarisme de la règle*, l'alternative la plus populaire actuellement et dont Peter Singer (1946–...) est un des défenseurs actuels les plus fameux³⁹. L'idée est de mettre en concurrence différents systèmes de règles morales

³⁸ SINNOTT-ARMSTRONG Walter, « Consequentialism », in ZALTA Edward N. et NODELMAN Uri (dir.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, 2023,

³⁹ En l'occurrence, Peter Singer défend l'utilitarisme "à deux niveaux" de Richard Hare (1919–2002), qui permet d'opter soit pour l'utilitarisme classique, soit pour l'utilitarisme de la règle, selon le contexte.

(qu'ils soient conséquentialistes, déontologiques, basées sur l'éthique des vertus, etc.), de mesurer la combinaison qui mène aux meilleures conséquences pour les êtres sentients, puis de se tenir mordicus à ces règles-là. En somme, l'utilitariste de la règle attend d'avoir fini de lire ce mémoire pour se décider sur la ou les règles à appliquer envers les siamois. Nous n'en parlerons donc pas spécifiquement ici.

Une dernière catégorie de conséquentialisme *agent-neutral* qu'il faut mentionner est le *contractarianisme* (ainsi que le contractualisme, un proche cousin). Cette doctrine s'appuie sur la métaphore du contrat social pour réfléchir à nos intuitions morales et à la structure de la société. Le contractarianisme avance que l'État est le produit d'un contrat implicite de chacun envers chacun. Dans ce contrat, chacun renonce à certaines libertés en échange de protection—un échange bénéfique pour des raisons égoïstes et rationnelles. Cette idée a été utilisée par des ténors de la philosophie politique moderne comme Thomas Hobbes (1588–1679), John Locke (1632–1704) ou Jean-Jacques Rousseau (1712–1778)⁴⁰. Néanmoins, leur formalisme est limité à délimiter les droits politiques ; ce qui ne nous aide pas vraiment pour le problème des siamois.

En revanche, la version contemporaine du contractarianisme a un formalisme plus riche. Elle consiste à imaginer tous les humains rassemblés, possédant toutes les connaissances nécessaires pour réfléchir à l'organisation d'une société. La seule chose que tous ignorent, c'est la place qu'ils auront dans cette société : c'est la *voile d'ignorance*. La formulation de John Rawls (1921–2002) est la plus connue⁴¹, mais sa conclusion est controversée, puisque selon lui un agent rationnel cherche à maximiser la situation du membre de la société le moins bien loti. Mais elle a été précédée par une version mathématique plus formelle, celle de John Harsanyi⁴² (1920–2000)—améliorée par John A. Weymark (1950–...) en 1991⁴³—qui arrive quant à elle à une conclusion utilitariste.

DARDENNE Emilie, « From Jeremy Bentham to Peter Singer », *Revue d'études benthamiennes*, n° 7, 2010.

⁴⁰ CUDD Ann et SEENA Eftekhari, « Contractarianism », in ZALTA Edward N. et NODELMAN Uri (dir.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, 2021.

⁴¹ RAWLS John, *A Theory of Justice*, Oxford : Clarendon Press, 1972, xv+607 p.

⁴² HARSANYI John C., « Cardinal Utility in Welfare Economics and in the Theory of Risk-taking », *Journal of Political Economy*, vol. 61, n° 5, University of Chicago Press, 1953, p. 434–435. et

HARSANYI John C., « Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility », *Journal of Political Economy*, vol. 63, n° 4, University of Chicago Press, 1955, p. 309–321.

⁴³ WEYMARK John A., « A reconsideration of the Harsanyi–Sen debate on utilitarianism », in ROEMER John E. et ELSTER Jon (dir.), *Interpersonal Comparisons of Well-Being*, Cambridge : Cambridge University Press, 1991, p. 255–320.

3.2. Voile d'ignorance d'Harsanyi

Comme nous l'avons expliqué plus tôt, le standard de preuve des institutions permet de résoudre le problème des siamois (section 2.2), et il découlerait d'un raisonnement utilitariste (section 2.3). Mais en quoi est-ce bien ? Et surtout, pourquoi est-ce que les juges devraient se soumettre aux préférences morales institutionnelles ? Il s'avère que la théorie utilitariste et contractarianiste de John Harsanyi permet de répondre à ces deux questions d'une traite. Nous allons d'abord la présenter, dans tout son formalisme mathématique ; ensuite, nous allons l'interpréter de manière critique.

L'approche de Harsanyi est la suivante. Disons qu'il y a trois êtres humains qui hésitent entre deux modèles de société, x et y . Chaque personne cherche à maximiser son bien-être personnel. Chacune donne une note à chaque modèle de société, qu'on représente sous forme de cotes⁴⁴. Voici par exemple les Préférences de monsieur A : $\langle P_a(x) = 2 : P_a(y) = 1 \rangle$. Ensuite, ajoutons une Observatrice impartiale, qui choisit le modèle de société. Si elle était certaine de "s'incarner" en monsieur A, elle proposerait elle aussi $\langle P(x) = 2 : P(y) = 1 \rangle$ et le modèle de société imposé serait x , car il a la meilleure note. Sauf que l'Observatrice est cachée derrière un *voile d'ignorance* : elle ne sait pas en qui elle va s'incarner. Elle considère qu'elle a autant de chance d'être l'un ou l'autre des trois humains. Autrement dit, c'est une loterie. Quel est le modèle de société que préfère l'Observatrice ?

D'après la *théorie de l'utilité espérée* (expected utility theory) de John von Neumann (1903–1957) et Oskar Morgenstern (1902–1977), un agent rationnel ne fait pas de différence entre un gain w et une loterie avec une espérance de gain w . On appelle "espérance de gain" le gain moyen qu'on peut espérer en tirant au sort. L'approche de von Neumann et Morgenstern est encore aujourd'hui la théorie orthodoxe de la rationalité, et elle est autant critiquée et amendée que défendue⁴⁵. Ainsi, à partir de ce postulat et de quelques subtilités mathématiques que j'élude, on peut montrer "simplement" que la note de bien-être d'une société est la *moyenne* (arithmétique) des

⁴⁴ J'insiste : ce sont des *cotes*, pas des probabilités. Il est facile de convertir l'une en l'autre, mais les maths seront plus limpides avec des cotes.

⁴⁵ VON NEUMANN John et MORGENSTERN Oskar, *Theory of Games and Economic Behavior*, Princeton : Princeton University Press, 2004 (1944), 776 p.

PETTIGREW Richard, « Risk, rationality and expected utility theory », *Canadian Journal of Philosophy*, vol. 45, n° 5–6, Routledge, 2015, p. 798.

notes des individus. Et que le modèle de société que l'Observatrice préfère est celui à la note la plus haute, celui que la collectivité préfère en moyenne : c'est-à-dire que l'utilitarisme l'emporte.

D'après les utilitaristes, nous devrions tous nous comporter comme l'Observatrice impartiale : c'est la seule à ne pas avoir de préférence envers l'une ou l'autre membre de la communauté. Cela est encore plus vrai et nécessaire pour une juge par métier. Elle ne fait qu'adopter les vœux moyens d'une foule composée d'égoïstes rationnels et bien informés.

Appliquée aux collectivités humaines, la théorie de Harsanyi justifie tellement bien l'utilitarisme qu'elle nous ferait oublier qu'elle est un formalisme abstrait. Cela veut dire qu'on peut remplacer 'le bien-être' par n'importe quelle unité de mesure objective et empirique que tout individu désirerait maximiser (ou minimiser), même si cette mesure est difficile en pratique. On pourrait, par exemple, vouloir maximiser le nombre de trombones⁴⁶ que chacun possède en moyenne—peu importe. On peut aussi remplacer 'individu' par n'importe quel agent capable de préférences. On pourrait, par exemple, traiter comme des agents les planeurs du *jeu de la Vie* de John Horton Conway (1937–2020) qui préfèrent, peut-on supposer, persévérer dans leur être plutôt que de se désintégrer⁴⁷—peu importe à nouveau. Autrement dit, pour toutes ces situations, la théorie de Harsanyi proclame assez naïvement les mêmes devoirs moraux : il faut que ces agents se comportent comme une Observatrice impartiale. Pour l'illustrer, prenons deux exemples, la personne métaphysique et la communauté des gènes.

3.3. Harsanyi appliqué aux tranches de vie

Pour analyser la notion de personne, comme point de départ métaphysique, le plus simple est de dire qu'une personne est une succession temporelle (censément continue) de tranches de vie. Une tranche de vie n'a pas de contrôle sur les choix de ses prédécesseurs et elle ne ressentira jamais l'impact futur de ses choix. Pour cette raison,

⁴⁶ Cet exemple est inspiré de l'expérience de pensée des trombones à papier universels de Nick Bostrom.

BOSTROM Nick, « Ethical Issues in Advanced Artificial Intelligence », in SMIT Iva *et alii* (dir.), *Cognitive, Emotive, and Ethical Aspects of Decision Making in Humans and in Ai*, Windsor (Ontario) : International Institute for Advanced Studies in Systems Research and Cybernetics, 2005, p. 16.

⁴⁷ BEER Randall D., « Autopoiesis and Cognition in the Game of Life », *Artificial Life*, vol. 10, n° 3, 2004, p. 309–326.

on peut traiter ces tranches de vie comme des agents avec des préférences égoïstes, qui devraient rationnellement viser une gratification instantanée (avant de disparaître pour toujours dans le passé). Comment, alors, expliquer la gratification différée ? Prenons l'exemple le plus classique, le *test du marshmallow* : si l'enfant résiste à l'envie d'en manger un pendant n minutes, l'expérimentateur lui en offrira deux. Le système de Harsanyi permet alors de formaliser le problème, en cachant toutes les tranches de vie derrière un voile d'ignorance. C'est d'ailleurs souvent pour étudier ce problème-là que le principe de von Neumann et Morgenstern est le plus souvent utilisé. Dans tous les cas, d'après les utilitaristes, toutes ces tranches de vie devraient se comporter comme l'Observatrice impartiale, et ne pas avoir de préférence envers l'une ou l'autre partie temporelle de l'être humain. Rationnellement, une personne doit maximiser son bien-être moyen sur toute une vie.

Cette analyse a des ramifications inattendues. En effet, dans les théories du contrat social, une des conditions préalables pour entrer dans un tel contrat, c'est justement d'être un individu rationnel. Les agents humains qu'imagine Harsanyi, ce sont des piles de tranches de vie dans un trench-coat. Par conséquent, on pourrait demander à l'Observatrice impartiale de se projeter dans les tranches de vie directement—peu importe, en soi, car le calcul final de la moyenne sera strictement le même. En fait, c'est justement cette propriété-là qui va être critiquée par d'autres théoriciens du contrat social : sans le vouloir, Harsanyi fait disparaître les personnes de toute réflexion morale, car elles ne sont qu'une étape intermédiaire et inutile des calculs, entre tranches de vie et société.

Comme nous l'avons mentionné plus tôt, il existe deux utilisations fameuses du *voile d'ignorance*, celle de John Harsanyi—dont nous venons de parler—et celle de John Rawls. Or, si le premier John a conclu en faveur de l'utilitarisme, le second John a radicalement rejeté cette conclusion. Un argument au cœur de ce rejet tient justement à la définition de la personne et de la rationalité. Bart Schultz synthétise de manière tellement claire le reproche de Rawls, que je ne peux pas faire mieux qu'une longue citation :

Rawls affirme que l'approche utilitariste étend à toute la société le principe de prudence rationnelle de une personne. [...] Naturellement, le problème c'est qu'une société, ce n'est pas comme une personne. Adopter cette approche, c'est traiter les identités différentes des personnes comme si elles n'avaient pas plus d'importance, pour l'ordre de la société, que n'ont les différentes étapes de la vie d'un individu

*n'ont dans sa prudente organisation. Mais une société est composée de personnes distinctes, chacune avec une vie à mener, un point de vue, etc. En tant que 'système-à-buts' distincts, les personnes ne peuvent pas simplement être liquidées comme autant de récipients à expériences riches. En effet, les frontières entre individus ont plus qu'une importance indirecte et définissent en fait l'objet même de la préoccupation morale.*⁴⁸

Pour sauver de l'élimination le concept de personne, Rawls propose une nouvelle règle : pour passer des tranches de vie à la personne, il faut utiliser un principe de rationalité A ; pour passer des personnes à la société, il faut utiliser un principe de rationalité B. C'est le seul moyen de forcer un calcul par étape, et donc de préserver l'importance conceptuelle des personnes. D'après Rawls, le principe A, c'est de maximiser le bien-être moyen (celui de von Neuman et Morgenstern)⁴⁹ ; tandis que le principe B, c'est de maximiser le bien-être du plus mal loti. L'analyse du problème et cette solution proposée par Rawls est géniale, mais elle est tout autant controversée que le problème à résoudre. Pourquoi, en effet, faudrait-il multiplier les principes de rationalité ? Et pourquoi ces principes-là plutôt que d'autres ?

Ces questions sont autrement plus techniques⁵⁰. Pour ce qui nous concerne, bornons-nous à souligner que la théorie de la justice de Rawls ne nous aide pas à résoudre le problème des siamois. Le principal problème, c'est son objectif de "maximiser le bien-être du plus mal loti". En effet, qui est le moins bien loti dans cette affaire ? Le siamois innocent ou la victime du crime ? Dans son fondement, la théorie de Rawls est pluraliste et instrumentaliste, c'est-à-dire qu'*a priori*, elle accepte d'appliquer à peu près n'importe quelle théorie de la punition, conséquentialiste ou rétributive—tant que cela sert son but ultime, maximiser le bien-être du plus mal loti. En somme, Rawls attend lui aussi d'avoir fini de lire ce mémoire pour se décider sur la règle à appliquer envers les siamois. Nous ne la développerons donc pas plus pour l'instant.

⁴⁸ SCHULTZ Bart, « Persons, Selves, and Utilitarianism », *Ethics*, vol. 96, n° 4, 1986, p. 721.

⁴⁹ VON NEUMANN John et MORGENSTERN Oskar, *Theory of Games and Economic Behavior*, Princeton : Princeton University Press, 2004 (1944), 776 p.

⁵⁰ En théorie, on pourrait encore étendre ce principe : pour passer des individus à la société, il faut d'abord passer par les familles (et par un principe de rationalité C) ; et pour passer des sociétés à l'Organisation des Nations-Unies, il faut utiliser un principe de rationalité D, etc.

Par exemple, on pourrait proposer qu'une personne rationnelle doit chercher à maximiser le bonheur de la tranche de vie la mieux lotie ; qu'une famille rationnelle doit maximiser le bonheur moyen des personnes (et donc, de ces tranches de vie les mieux loties) ; et qu'une ONU rationnelle doit maximiser le bonheur de la famille la moins bien lotie. Avec quatre principes de rationalité différents et jusqu'à trois individus intermédiaires, potentiellement irréductibles, la combinatoire monte déjà à $4^4 = 256$ possibilités.

3.4. Harsanyi appliqué aux gènes

L'exact même débat existe en biologie ; mais plutôt que de découper les humains en tranches temporelles, on les découpe en tranches fonctionnelles. C'est en particulier le cas en biologie évolutionnaire, où la génétique joue un rôle central. L'idée de traiter les gènes comme des agents égoïstes remonte aux années 1960, par une 'heuristique de personification' popularisée par Dawkins et W.D. Hamilton⁵¹. Dans les années 2010, le philosophe Samir Okasha a fait le pas de plus, en appliquant le formalisme de Harsanyi à un exemple simple : la loi de ségrégation de Mendel⁵². Cette dernière dit que lors de la reproduction sexuée, chaque gène a 50% de chance de se retrouver dans une descendance. En somme, il s'agit d'une loterie. Cette situation se laisse donc facilement formaliser dans le système de Harsanyi. Certains gènes, cependant, trichent en tout égoïsme et montent les cotes jusqu'à 90:10 en leur faveur⁵³. Or, le bénéfice principal de la reproduction sexuée reste le brassage génétique : cette triche lèse donc la communauté des gènes. Dans cette situation, l'Observatrice impartiale, qui ne sait pas dans quel gène elle s'incarnera, a toutes les raisons de promulguer la loi de ségrégation de Mendel—ce qui est exactement ce qu'on retrouve en réalité. La loi de Mendel est donc un exemple d'utilitarisme.

Cette réduction de l'agentivité des organismes à l'agentivité des gènes a elle aussi fait couler beaucoup d'encre. Une critique analogue à celle de Rawls est adressée à Dawkins—il éliminerait la notion d'organisme—mais la contre-analyse proposée par les biologistes n'a rien de rawlsienne. En effet, dans *Unto Others*, le philosophe Elliott Sober et le biologiste David Sloan Wilson écrivent que les gènes forment une « communauté » qui a « évolué un système de règles qui leur permet de fonctionner comme une unité adaptative » ; pour cette raison, ajoutent-ils, il faut parler de *transition évolutionnaire de l'individualité* (ETI), un phénomène qui fait que « les groupes seront comme des

⁵¹ DAWKINS Richard, *The Selfish Gene*, Oxford University Press, 2016, 497 p.

HAMILTON W. D., « The genetical evolution of social behaviour. I », *Journal of Theoretical Biology*, vol. 7, n° 1, 1964, pp. 1–16.

⁵² OKASHA Samir, « Social Justice, Genomic Justice and the Veil of Ignorance: Harsanyi Meets Mendel », vol. 28, n° 1, Cambridge University Press, Cambridge, United Kingdom, 2012, pp. 43–71.

⁵³ KOZIELSKA M. *et alii*, « Segregation distortion and the evolution of sex-determining mechanisms », *Heredity*, vol. 104, n° 1, 2010, p. 100.

individus »⁵⁴. Parmi ces transitions d'individualité, citons quelques unes des plus importantes : le génome (l'union des brins d'ADN en chromosomes), les eucaryotes (par endosymbiose), la multicellularité, les sociétés eusociales, et le mutualisme interespèces⁵⁵.

Si nous reformulons la critique rawlsienne dans le vocabulaire des biologistes, nous en tirons trois thèses clés :

1. Le formalisme de Harsanyi décrit une *transition d'individualité*.
2. Au sein du formalisme de Harsanyi, un mécanisme spécifique génère cette transition : l'Observatrice impartiale.
3. Le formalisme de Harsanyi ne nous dit pas sous quelles conditions est-ce qu'une *transition d'individualité* est une bonne ou une mauvaise chose (alors que c'est précisément cette question-là qui obnubile la philosophie morale).

Il semble que définir l'individu à partir des *transitions évolutives de l'individualité* soit une voie prometteuse pour formaliser la critique de Rawls contre Harsanyi. Cela nous permet aussi de vérifier qu'il dit vrai : est-ce que cette Observatrice impartiale est suffisante pour faire une transition d'individualité ? Et si oui, est-ce que ce nouveau contexte nous permet de savoir si l'utilitarisme a un rôle à jouer—et lequel—pour résoudre le problème des siamois ?

⁵⁴ « If communities of genes and cells can evolve a system of rules that allow them to function as adaptive units, then why can't communities of individuals do the same? If they do, then *groups will be like individuals*, which is the proposition that we are seeking to establish. »

SOBER Elliott et WILSON David Sloan, *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Cambridge, Mass. : Harvard Univ. Press, 1999, 394 p., cité dans DENNETT Daniel C., *Freedom Evolves*, New York : Viking, 2003, p. 193.

⁵⁵ WEST Stuart A. *et alii*, « Major evolutionary transitions in individuality », *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, n° 33, National Academy of Sciences, 2015, p. 10113.

4. LES TRANSITIONS ÉVOLUTIONNAIRES D'INDIVIDUALITÉ

4.1. Définition fonctionnelle de Clarke

En biologie, il existe au moins une quinzaine de définitions différentes et concurrentes de l'organisme individuel, centrées sur autant de concepts : 'reproduction', 'cycles de vie', 'génétique', 'sexe', 'goulots développementaux', 'séparation germe-soma', 'mécanismes de service d'ordre', 'cloture ou contiguïté spatiale', 'réponse immunitaire', 'maximisation d'adaptation', 'coopération' et/ou 'conflit', 'codispersion', 'adaptations', 'autonomie métabolique', 'intégration fonctionnelle'⁵⁶. Cependant, aucune de ces définitions n'est consensuelle et la plupart parviennent de temps en temps à des conclusions étranges. Surtout, il y a de nombreux cas où elles ne sont pas d'accord entre elles ; or, l'individu biologique est aussi l'individu évolutionnaire, celui sur lequel le processus de sélection naturelle agit⁵⁷. En définissant et délimitant différemment les populations vivantes, les prédictions changent—ce qui est un problème.

Par chance, la philosophe Ellen Clarke prétend pouvoir unifier toutes ces approches concurrentes dans un seul paradigme. Sa stratégie, c'est de retourner au concept même de *transition évolutionnaire d'individualité*⁵⁸ : ces transitions ont des effets variés sur les systèmes biologiques, et utilisent des mécanismes très hétéroclites. Mais tous ces effets ont, au fond, les mêmes fonctions. Pourrait-on définir la transition d'individualité (et donc l'individu) à partir de son fonctionnement ?

D'après Clarke, un organisme individuel se définit par la possession de deux types de mécanismes fonctionnels⁵⁹. (1) Les mécanismes de *démarcation* aident l'objet à participer au processus de sélection naturelle—voire fondent sa capacité à le faire—en dissociant d'autant plus son destin de celui des autres objets. À l'inverse, (2) les *services d'ordres* assistent l'objet pour empêcher—voire, modifier à son profit⁶⁰—les processus

⁵⁶ CLARKE Ellen, « The Multiple Realizability of Biological Individuals », *The Journal of Philosophy*, vol. 110, n° 8, 2013, p. 415.

⁵⁷ CLARKE Ellen, « A levels-of-selection approach to evolutionary individuality », *Biology & Philosophy*, vol. 31, n° 6, 2016, p. 894.

⁵⁸ CLARKE Ellen, « The Multiple Realizability of Biological Individuals », *art.cit.*, p. 413.

⁵⁹ *Ibid.*, p. 427.

⁶⁰ BAILLON Ludovic et BASLER Konrad, « Reflections on cell competition », *Seminars in Cell & Developmental Biology*, vol. 32, 2014, pp. 137–144.

de sélection naturelle en son sein, en liant d'autant plus le destin de toutes ses parties entre elles. En bref, un individu concurrence dehors et coopère dedans. Plus l'organisme est capable de ces deux actions, plus son caractère d'individu est affirmé. À l'inverse, il perd son individualité propre, si ses services d'ordre perdent contre les mécanismes de démarcation de ses sous-parties (l'individu se disloque), ou si ses propres mécanismes de démarcation perdent contre le service d'ordre d'un superorganisme (l'individu fusionne avec d'autres).

Il vaut la peine de signaler deux cas-limites à cette définition. La première exception, c'est le premier individu, le proto-organisme atomique qu'on ne peut pas décomposer plus avant (par définition)⁶¹. Chez ce dernier il n'y a aucun service d'ordre, mais des mécanismes de démarcation seulement—il s'agit pourtant clairement d'un organisme individuel. C'est ce que Clarke appelle le *Life problem*, par opposition au *Levels problem* dont elle s'est occupée. La seconde exception, c'est un hypothétique dernier individu, un superorganisme qui contient tout le vivant et a donc atteint la dernière transition d'individualité possible (par définition). C'est ce qu'on pourrait appeler un *Gaia problem*, que nous pouvons ajouter à la typologie de Clarke. Chez ce dernier, il n'y a que des services d'ordre, mais pas de mécanismes de démarcation. Pour autant, ce superorganisme final serait toujours bel et bien un individu, de la même manière qu'un proto-organisme était un individu, lui aussi.

4.2. L'individualité est une poupée russe

Si l'on se penche un peu sur la manière dont Ellen Clarke décrit le mécanisme de sélection naturelle, on peut identifier « une circularité vicieuse » qu'elle relève elle-même⁶². Prenons l'exemple de la transition des eukaryotes à la multicellularité. Les cellules eukaryotes sont des individus, elles possèdent donc chacune un mécanisme de

NELSON Paul et MASEL Joanna, « Intercellular competition and the inevitability of multicellular aging », *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, n° 49, 2017, pp. 12982–12987.

⁶¹ Il est difficile de savoir qui est cet individu atomique. Peut-être que les niveaux cessent lorsque commence la sélection naturelle et la proto-cellule. Alternativement, nous pouvons continuer jusqu'à la sélection chimique, ce qui nous arrête au niveau du gène, du brin d'ARN, de la protéine, qui ne sont pas décomposables. Mais il est possible que les principes de Clarke puissent être adaptés pour se passer de sélection chimique et puissent s'appliquer à des entités chimiques et physiques plus basiques encore.

⁶² CLARKE Ellen, « The Multiple Realizability of Biological Individuals », *art. cit.*, p. 430.

démarcation qui les distingue les unes des autres. En bref, ce trait résiste aux transitions évolutives d'individualité, c'est presque sa fonction. Pour que la multicellularité apparaisse, il faudrait qu'un mécanisme de *services d'ordre* venu d'en haut s'oppose aux nombreuses *démarcations* d'en bas. Or, ces *services d'ordres* sont des traits fonctionnels, qui sont sélectionnés par l'évolution, au niveau d'un superindividu. Mais quel superindividu ? La transition évolutive d'individualité n'a même pas commencé !

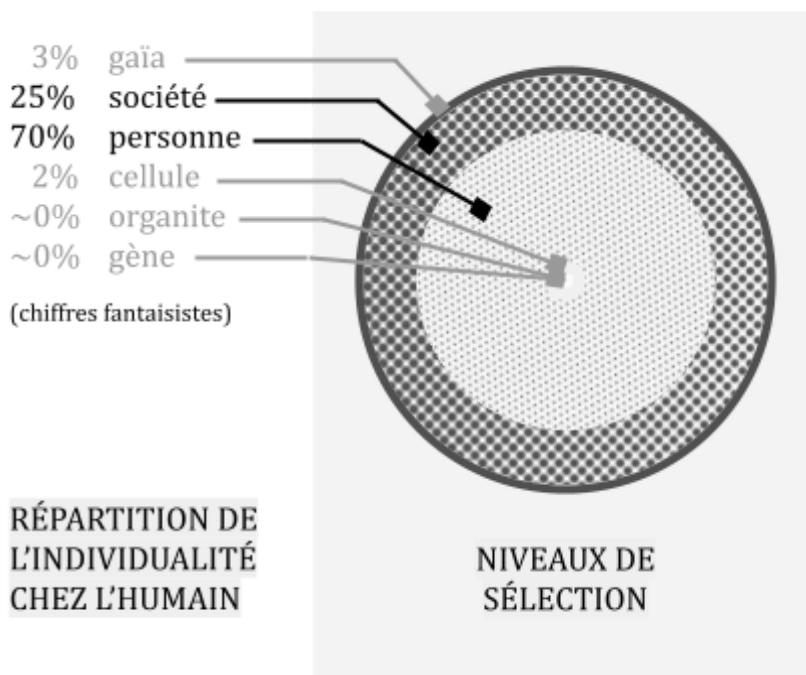


fig. III : Schéma de l'individualité graduelle.

Pour sortir de ce problème, Clarke utilise une définition gradualiste de l'individualité (fig. III). D'abord, on suppose que par le fruit du hasard (comme une mutation de la membrane qui acolle des cellules-filles entre elles), la *transition évolutive* fait 1% du chemin. Dès lors, la sélection naturelle agit à 99% au niveau des eukaryotes individuels, et 1% au niveau multicellulaire. Cela suffit pour lancer l'effet boule-de-neige.

Cette concession de Clarke est très crédible ! Car en pratique, en biologie, on n'est jamais un pur individu. Des *services d'ordre* parfaits n'existent pas : tous les multicellulaires sont à risque de cancer, tous les génomes risquent que certains gènes se dupliquent de manière abusive durant la méiose. Nous ne sommes pas nous-mêmes de purs individus ! Des *mécanismes de démarcation* parfaits n'existent pas : tous les humains finissent par se spécialiser au sein de la société et en dépendent largement pour survivre. Autrement

dit, une définition biologique de l'individualité doit pouvoir, et doit chercher à distinguer différents niveaux de sélection et pondérer leur importance respective. En retour, cela permet d'affirmer que par nature tout organisme vivant a toujours plusieurs degrés d'individualité concentriques, comme des poupées russes dont certaines couches sont plus épaisses et pertinentes que d'autres (fig. 111).

4.3. L'utilitarisme comme transition évolutionnaire d'individualité

Nous avons maintenant un moyen de confirmer, ou réfuter, l'argument de Rawls contre Harsanyi. Est-ce que l'utilitarisme mène à une *transition d'individualité* ? Est-ce que l'Observatrice impartiale sert à la fois de mécanisme de *service d'ordre* et de mécanisme de *démarcation* ? Cherche-t-elle à la fois à lier le destin évolutionnaire de certains objets, et à distinguer leur destin commun du destin du reste du monde ?

La première action de l'Observatrice impartiale, c'est de choisir qui elle écoute et qui elle ignore. Les utilitaristes, par exemple, ne tiennent compte que des agents capables de sensations, de souffrance, d'aspirations psychologiques. Les fourmis, elles, ne tiennent compte que de leurs sœurs ou leurs proches cousines. Exploiter le plus possible l'exogroupe au bénéfice de l'endogroupe n'est pas seulement autorisé, *il est moralement requis*, car c'est ce que l'Observatrice préfère par principe. Ainsi, chez Harsanyi, l'Observatrice réalise la fonction de *démarcation* dans sa forme la plus extrême. Cette puissante fonction est bel et bien présente, même quand elle nous sert à dire qu'il est bon d'exploiter radicalement un animal non-conscient, une plante, une bactérie, un écosystème, un paysage, un artefact, un robot, une IA, etc. pour le bénéfice de l'endogroupe.

Comme nous l'avons vu, le trait caractéristique d'une transition évolutionnaire d'identité, c'est que le mécanisme de *démarcation* des sous-individus est massivement affaibli par l'action des nouveaux *services d'ordre* du superindividu. Or, c'est exactement ce qu'on observe. Le but même de l'expérience de pensée du voile d'ignorance, c'est de nous demander de renoncer à nos vœux égoïstes rationnels, en faveur du vœu de l'Observatrice impartiale. C'est-à-dire que nous renonçons à la capacité de distinguer activement notre destin, nos aspirations psychologiques et notre sens moral de celui des

autres : c'est un *service d'ordre*. La définition de l'organisme individuel de Clarke démontre donc qu'aux yeux de la science, le formalisme de Harsanyi crée bel et bien un nouveau superorganisme.

C'est donc raisonnablement établi : les utilitaristes classiques postulent l'existence d'un superorganisme qui comprend tous les êtres sentients. Or, un tel superorganisme n'existe certainement pas et il n'apparaîtra jamais. Se comporter comme si cet organisme existait réellement est assez étrange, sur le plan ontologique et sur le plan moral. Au mieux, ce superorganisme-là est une fiction philosophique un peu douteuse—mais peut-être moralement justifiée. Au pire, et comme le défend Rawls, c'est une erreur morale de la plus grande envergure. Mais a-t-on vraiment besoin d'avalier une couleuvre aussi énorme pour résoudre le problème des siamois ? Peut-on rendre le conséquentialisme plus parcimonieux sur le plan ontologique ?

4.4. Peut-on sauver le conséquentialisme ?

Une première observation c'est que dans le problème des siamois, l'avis égoïste d'une mésange ou d'un poulpe nous importe peu ! Ils sont indifférents au sort des jumeaux ou de leurs futures victimes. Concrètement, nous n'avons besoin de récolter que l'avis d'humains qui peuvent être victimes d'une erreur judiciaire, dans un système pénal donné. Réduisons donc l'envergure minimale de ce superindividu : plutôt que de faire entrer tous les êtres sentients dans l'assemblée, restreignons le nombre de gens derrière le voile d'ignorance et réfléchissons en termes de communautés humaines. La question devient alors « les citoyens soumis au droit suisse (par exemple) ont-ils renoncé à toute leur individualité pour former un unique superorganisme, le *corps social* ? ». Même après cette reformulation, le problème reste presque entier : clairement, vous et moi sommes toujours des individus, au moins en partie. Traiter le corps social comme le *seul* individu que l'on peut léser reste une fiction douteuse, et une erreur morale considérable.

D'un autre côté, comme nous venons de le voir, toute transition d'individualité est graduelle. Est-il raisonnable de croire que le corps social n'est *pas* un individu, même en petite partie ? Que les êtres humains ne peuvent pas, par nature, être engagés dans une nouvelle transition évolutionnaire d'individualité ? Plusieurs philosophes et théoriciens

de la sociologie, comme Auguste Comte (1798–1857), Albert Schäffle (1831–1903) ou Émile Durkheim (1858–1916), ont utilisé une vision très ‘organismique’ des groupes humains⁶³. Cette analogie avec l’organisme a été vertement critiquée notamment par Max Weber—qui lui reconnaissait tout de même une valeur heuristique⁶⁴. Clairement, le corps social est très loin d’être aussi nettement un individu que le corps d’un animal. Mais la fécondité de cette analogie suggère qu’une transition d’individualité a au moins commencé.

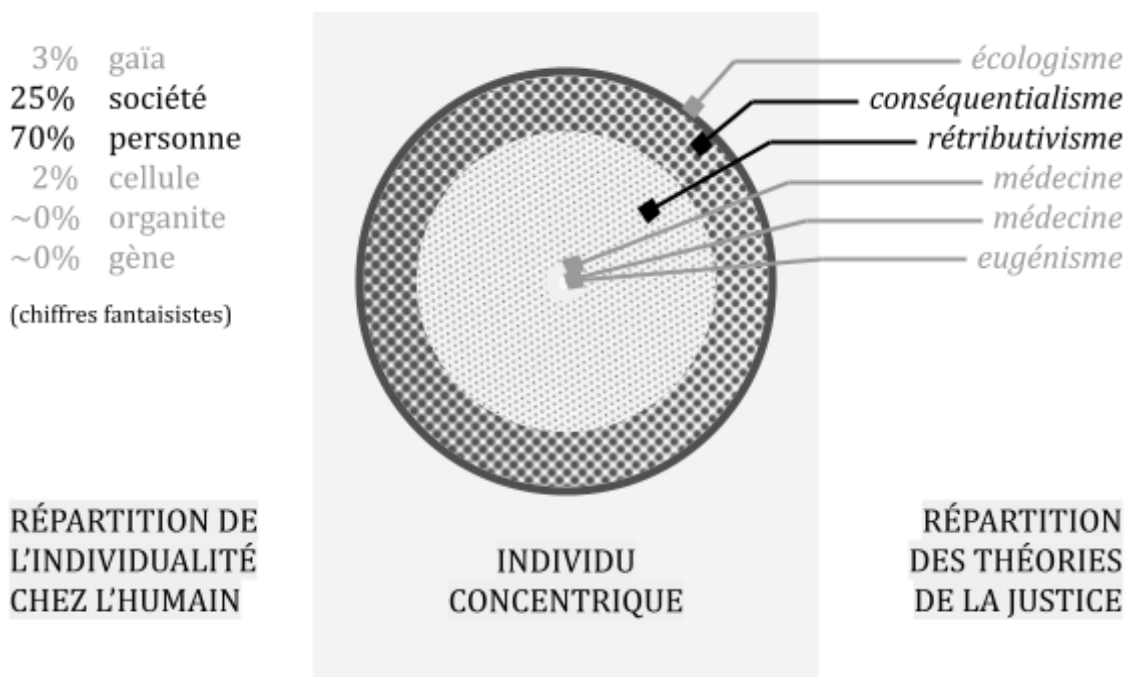


fig. iv : Schéma de la punition graduelle.

Or, à supposer que le corps social existe (en partie), alors on peut le léser. Le détournement de fonds publics et la fraude électorale, par exemple, sont des crimes dirigés spécifiquement contre le corps social. Une injure sexiste ou une attaque raciste impactent aussi le corps social, dans la mesure où elles cherchent à le scinder en sous-communautés et à les mettre en concurrence. Quant à la démagogie, qui renforce le ressentiment des intellectuels envers les incultes (et réciproquement)⁶⁵, elle relève elle

⁶³ LEVINE Donald N., « The Organism Metaphor in Sociology », *Social Research*, vol. 62, n° 2, 1995, pp. 239–265.

LEE Monica et SILVER Daniel, « Simmel’s Law of the Individual and the Ethics of the Relational Self », *Theory, Culture & Society*, vol. 29, n° 7–8, 2012, pp. 126–127.

⁶⁴ LEVINE Donald N., « The Organism Metaphor in Sociology », *art. cit.*, pp. 259–260.

Par ailleurs, Jeremy Bentham, fondateur de l’utilitarisme, s’était aussi nettement opposé à cette vision organismique—un peu ironiquement, compte tenu de notre discussion préalable.

⁶⁵ STEUDEMANN Michael J., « Rhetorical Education in Times of Demagoguery », *Rhetoric Society Quarterly* 49, 2019, pp. 297–314.

aussi d'une attaque contre le corps social. De manière générale, il n'est presque pas possible de léser un individu sans *aussi* léser le corps social. Car si les humains ont renoncé (par exemple) à un quart de leur individualité au profit de ce superindividu, alors tout crime fera toujours au moins deux victimes : un être humain pour $\frac{3}{4}$, et le corps social pour $\frac{1}{4}$. Or, qui dit deux victimes dit qu'il faut *additionner* deux punitions différentes, de différente nature (fig. iv).

Bien sûr, il y a de la variabilité dans ce principe. Un ermite en autarcie, par exemple, n'a presque rien cédé de son individualité ; il en est de même pour un banni ostracisé de la société. Par contraste, un idéologue dévoué corps et âme à une vraie cause patriotique (quoi que cela signifie concrètement) peut, à l'inverse, avoir cédé la majeure partie de son individualité au profit de la collectivité. Il en est de même pour les esclaves publics, du temps de la démocratie athénienne⁶⁶. Tout crime dirigé contre cet idéologue ou cet esclave public est donc un crime dirigé contre le corps social ; le criminel doit donc être puni presque entièrement par un raisonnement utilitariste. À l'inverse, le meurtre de l'ermite ou du banni lèse ce dernier *qua* personne, exclusivement. Le corps social n'est pas impacté ; le conséquentialisme n'a donc presque aucune place pour nous aider à former un jugement moral. En fait, dans ce cas de figure, notre version raisonnable du conséquentialisme échoue à justifier une punition qui va au-delà d'une tape sur la main⁶⁷.

*
**

Résumons rapidement où nous en sommes. Nous pouvons sauver l'utilitarisme et lui laisser une place dans notre raisonnement punitif : il punit un crime envers des victimes *qua* fractions d'une communauté. Plus sobrement, il considère que la communauté s'attaque elle-même. Il est parfaitement incapable de concevoir une punition pour un individu qui en attaque un autre ; et c'est normal de ce point de vue, puisque c'est une situation de pure sélection naturelle, de concurrence de tous avec tous. Que faire ?

⁶⁶ ISMARD Paulin, « Le simple corps de la cité : Les esclaves publics et la question de l'État grec », *Annales. Histoire, Sciences Sociales*, 2014, vol. 69, n° 3, pp. 723-751.

⁶⁷ L'utilitarisme classique et l'utilitarisme de la règle partagent ce problème, car ils mesurent tout à l'aune des conséquences pour l'ensemble du corps social. Quant au conséquentialisme agent-relative, il demande de jouer à la juge impartiale et ne peut donc pas se prononcer sans en savoir plus, comme nous l'avons déjà vu.

L'option la plus attirante, et de loin, c'est de combler ce creux moral par une seconde théorie de la punition, qui permet de punir un crime envers des êtres humains *qua* individus—ce qui permet de rendre justice à toute la victime, à la fois personne et communauté⁶⁸. Concrètement, une telle théorie se soucie uniquement d'*arbitrer* les relations interpersonnelles de manière impartiale. Si vous avez eu l'honneur ou la honte de croiser la route d'une juge, elle vous dira sans doute qu'elle se reconnaît plus dans ce métier d'arbitre que celui de système immunitaire de la société que lui assigne en douce l'utilitarisme de Harsanyi.

Or, il existe justement une deuxième famille de théories de la punition qui fonctionne plutôt sur ce mode-là. Il est temps de quitter le paradigme conséquentialiste et de mettre les théories rétributives à l'épreuve du problème des siamois.

4.5. Rawls comme une approximation limitée

Et Rawls dans tout ça ? C'est vrai, nous avons fait mine de l'oublier. Il s'agit pourtant d'une théorie conséquentialiste, mais opposée à l'utilitarisme. Pour éviter les problèmes de transition d'individualité induite, la solution rawlsienne consiste à remettre en question le principe de rationalité de von Neumann et Morgenstern⁶⁹. Plutôt que de faire une moyenne entre intérêts personnels, Rawls propose d'utiliser un nouveau principe de rationalité, qui ne concerne que les relations interindividuelles des humains au sein de la société ; spécifiquement, il veut maximiser les intérêts du plus mal loti, plutôt que les intérêts de la personne moyenne⁷⁰. Comme nous l'avons déjà dit, ce principe ne nous permet pas directement de résoudre le problème des siamois (section 3.3).

⁶⁸ Bien sûr, nous avons deux autres options. La première, c'est d'accepter que ceux qui quittent l'endogroupe perdent d'autant la protection de l'État. La loi et la justice ne concernent que les affaires d'intérêt public ; si la société n'a rien à gagner, matériellement, à rendre justice sur tel ou tel cas, elle ne doit pas le faire.

La deuxième option, c'est que l'État impose à tous ses administrés de renoncer à leur individualité. Cela revient à avaler tout rond le mythe du système de Harsanyi, et imposer une transition d'individualité terminale à tous ceux que touche la loi—*c'était son commandement ultime, et le plus essentiel*, pour citer Orwell.

⁶⁹ SCHULTZ Bart, « Persons, Selves, and Utilitarianism », *Ethics*, vol. 96, n° 4, 1986, p. 721.

⁷⁰ RAWLS John, *A Theory of Justice*, Oxford : Clarendon Press, 1972, xv+607 p.

Par contraste, comme nous l'avons vu, la solution clarkienne consiste à remettre en question la binarité de l'individualité⁷¹. Elle préfère répartir l'individualité dans des poupées russes, une solution qui semble découler naturellement des préceptes de l'évolution darwinienne. De ce point de vue, l'utilitarisme s'applique à un individu partiel seulement, mais pas sur l'ensemble des relations interpersonnelles. Toujours de ce point de vue-là, la théorie de la justice de Rawls est une approximation d'une transition d'individualité partielle. Pour le comprendre, on peut grossièrement mesurer la force des deux fonctions responsables de la transition d'individualité.

Chez Rawls, il y a un *mécanisme de démarcation* fort : la communauté qui délibère derrière le voile d'ignorance ne comprend que des humains. Tous les autres en sont exclus. Il faut noter que d'autres auteurs ont adapté ce système pour y inclure un second cercle qui comprend les animaux domestiques (poulets, cochons, chats, chiens, etc.), mais aussi les animaux liminaires en marge de nos villages et de nos villes (pigeons, renards, rats, etc.)⁷². Dans tous les cas, il y a une fonction de démarcation nette entre un (ou plusieurs) endogroupes et un exogroupe.

En revanche, chez Rawls, le *service d'ordre* est moyennement fort : chaque humain (ou chaque animal contractant) conserve la plus grande part de son individualisme. La raison est simple, car le seul but commun, c'est de maximiser le bien-être du plus mal loti. Rawls nous autorise à entrer en concurrence les uns contre les autres sur presque tous les plans, si cela améliore les conditions de vie du plus marginal. Cependant, *service d'ordre* il y a bel et bien ! Seul des anarchistes individualistes comme Stirner⁷³ ou Nozick⁷⁴, par exemple, cherchent à éviter toute aliénation de l'individu humain et cherchent donc à démanteler l'État. Or, Rawls n'adhère pas à cette idée : il vise plutôt, semble-t-il, un "juste" milieu propice à l'émulation humaine, à mi-chemin entre le collectivisme formique à la Harsanyi et l'individualisme farouche à la Stirner.

⁷¹ CLARKE Ellen, « A levels-of-selection approach to evolutionary individuality », *art. cit.*, p. 894.

⁷² LIBERTO Hallie, « Species Membership and the Veil of Ignorance: What Principles of Justice would the Representatives of all Animals Choose? », *Utilitas*, vol. 29, n° 3, 2017, pp. 299–320.

DONALDSON Sue et KYMLICKA Will, *Zoopolis: A Political Theory of Animal Rights*, Oxford : Univ. Press, 2014, 329 p.

⁷³ LEOPOLD David, « Max Stirner », in ZALTA Edward N. et NODELMAN Uri (dir.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, 2023.

⁷⁴ NOZICK Robert, *Anarchy, State, and Utopia*, New York : Basic Books, 1974, xvi+367 p.

En somme, Rawls nous commande de renoncer à la part d'individualité à laquelle l'humain moderne a déjà renoncé. Du point de vue de la biologie, la théorie de Rawls et son principe de rationalité propre à la société est une *approximation* heureuse d'une transition d'individualité partielle, qui postule le début d'existence et la part d'individualité du corps social. Or, Rawls a parfaitement raison quand il place l'art de marquer *précisément* les frontières entre individus au cœur du projet de la philosophie morale. Pour cette raison, malgré tous ses mérites indéniables, il faut rejeter l'approximation qu'il propose, qui risque de nous apporter plus de problèmes que de bénéfices pour des cas problématiques comme le problème des siamois.

Pour sauver Rawls malgré tout, il faudrait avaler quelques couleuvres. Certes, pourrait-il dire, l'individu biologique est une poupée russe, mais la personne morale ne l'est pas. Pourtant, pourrait-on rétorquer, une théorie de la justice sert à régenter le degré de coopération attendu entre humains ; or, le degré de coopération, c'est justement ce que la biologie mesure grâce à ce concept de transition évolutionnaire d'individualité. Même à dissocier parfaitement ces deux champs de recherche, et leurs notions respectives de l'individualité, cette fiction rawlsienne de la personne comme individu complet semble tout aussi fantastique que la fiction harsanyienne de la société comme individu complet. Ce faisant, l'attaque de Rawls contre Harsanyi perd une grande partie de sa légitimité, et se transforme en simple différence d'intuition.

5. LE RÉTRIBUTIVISME

5.1. Aperçu du rétributivisme classique

Comme pour le conséquentialisme, il existe plusieurs versions du *rétributivisme*. Toutes partagent un concept central, le *mérite*, selon le principe que chaque crime mérite une punition appropriée (une rétribution). Pour prendre une décision de justice, peu importent les conséquences de notre propre décision : tout ce qui compte, ce sont les motivations et les répercussions du crime passé, celui que l'on juge. Cette famille de doctrine est l'alternative la plus connue au conséquentialisme. Et il s'avère que le problème des siamois a déjà été utilisé, récemment, pour les attaquer.

Dans un article de 2018, la juriste Brittany Deitch a argumenté de manière assez convaincante que l'exemple des jumeaux siamois justifie de refuser *toutes* les versions du rétributivisme⁷⁵ ! Dans son attaque générale, elle met à l'épreuve les rétributivismes dits 'pur', 'déontologique', 'de seuil', 'négatif/faible', 'soucieux des victimes', 'mixte/hybride' et le rétributivisme dit 'conséquentialiste'. Dans la première partie de ce chapitre, je reprends de manière critique les stratégies argumentatives de Deitch. Dans chaque cas, nous partons du principe que les deux jumeaux représentent deux personnes différentes : un jumeau innocent qui ne mérite pas de prison ; et un jumeau coupable qui mérite dix ans de prison. Nous nous demandons quelle peine de prison $\{p\}$ est-ce qu'il faut donner à la paire.

La suite de ce chapitre consiste à construire une version du rétributivisme qui échappe aux critiques de Deitch—puis, de mettre à l'épreuve cette même stratégie et ses conséquences philosophiques sur la nature de la punition rétributive et le concept même de mérite. Nous pouvons commencer par les théories rétributives classiques.

*
**

Pour le *rétributivisme classique* ou *rétributivisme déontologique* (kantien, par exemple), il y a deux règles absolument obligatoires : (i) toujours punir tous les coupables, et (ii) ne jamais punir aucun innocent (ni punir avec disproportion)⁷⁶. On peut résumer cela comme "il faut toujours punir chaque personne au plus près de ce qu'elle mérite". Or, comme on l'a vu, c'est physiquement impossible dans le cas de jumeaux siamois. Pour prendre une décision, la juge doit violer une de ces deux règles absolument obligatoires—ce qui est impossible, par définition. Il n'existe aucune peine de prison que la juge puisse donner aux siamois. Formellement, la solution au problème est donc qu'il n'y en a pas. Nous la notons par un ensemble vide $\{\}$.

Pour le *rétributivisme de seuil* (aussi appelé *threshold deontology*), il est obligatoire de punir tous les coupables et aucun innocent... sauf situation extrême, comme un risque

⁷⁵ DEITCH Brittany, « Retributivist Theories' Conjoined Twins Problems », *University of Cincinnati Law Review*, vol. 87, n° 1, 2018, p. 139.

⁷⁶ KANT Emmanuel, *Métaphysique des mœurs*, RENAUT Alain (trad.), Paris : Flammarion, 2018 (1795), 2 vol.
MOORE Michael S., *Placing Blame: A Theory of the Criminal Law*, Oxford : Clarendon, 1997, pp. 154–158.

de guerre nucléaire⁷⁷. Même à largement baisser le seuil, nous ne sommes pas dans ce cas ici, donc le résultat est le même que plus haut : l'ensemble vide {}.

Pour le *rétributivisme négatif* (aussi appelé *rétributivisme faible*), le rétributivisme classique est trop rigide et sa première règle est injustifiée : punir n'est pas obligatoire⁷⁸. Si une communauté décide de ne jamais amender, ni ostraciser, ni écrouer personne, aucune faute morale n'a été commise. Le seul interdit moral, c'est de punir un innocent ou de punir un coupable trop sévèrement ; tous les moyens sont bons pour y parvenir. Quant au problème des siamois, la théorie est très claire : pour protéger l'innocent, il faut libérer sans condition les deux jumeaux, c'est-à-dire choisir la peine minimale. Cette peine, nous la notons {*min*}. Évidemment, cette lacune est exploitable par un jumeau perfide et, de crime en crime, cela revient à lui octroyer un permis de tuer. Intuitivement, c'est une conséquence qui semble assez immorale.

On pourrait prendre le parti inverse, et plutôt renoncer à la deuxième règle du rétributivisme classique. On peut appeler cette version le *rétributivisme fort*. Ainsi, punir les coupables au moins autant qu'ils le méritent est obligatoire, nonobstant les victimes collatérales. Peu de philosophes défendent cette approche, pour des raisons assez évidentes. En effet, l'objectif premier de la justice devient de punir l'ennemi public n°1 exactement autant qu'il le mérite (prison à vie) ; dans la mesure où la liberté d'un innocent entrave cet objectif, on peut le faire croupir en prison. Intuitivement, c'est une conséquence assez immorale. Quoi qu'il en soit, ici le problème des siamois est univoque : il faut condamner les deux jumeaux à la peine que le coupable mérite, c'est-à-dire la peine maximale. Nous la noterons {*max*}.

5.2. Théories hybrides conséquentialistes

Devant ces problèmes et d'autres, certains philosophes ont tenté de combiner le rétributivisme avec le conséquentialisme. Comme nous l'avons vu, cette idée n'est pas absurde ; mais le diable se cache dans les détails. Mitchell Berman, par exemple, a

⁷⁷ MOORE Michael S., « The Rationality of Threshold Deontology », in HURD Heidi M. (dir.), *Moral Puzzles and Legal Perplexities: Essays on the Influence of Larry Alexander*, Cambridge : Cambridge University Press, 2018, p. 371-387.

⁷⁸ WALEN Alec, « Retributive Justice », in ZALTA Edward N. et NODELMAN Uri (dir.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, 2023.

récemment défendu un *rétributivisme hybride*, où chacune des deux approches a son pré carré⁷⁹. Dans sa prise de décision, le juge doit décider si l'accusé est coupable ou non, et appliquer le *rétributivisme classique* : si l'accusé est réputé coupable, il faut le punir proportionnellement ; et s'il est réputé innocent, il faut le libérer. Berman utilise le *conséquentialisme* seulement pour justifier moralement l'existence d'erreurs judiciaires involontaires. Concrètement, le conséquentialisme permet de choisir le standard de preuve (cf. sections 2.2 et 2.3). Par exemple : pour chaque innocent condamné par erreur à la prison, on va blanchir par erreur cinq, dix ou vingt coupables. Selon comment on interprète le cas des siamois, l'approche de Berman se réduit au *rétributivisme classique* et se coince dans le même paradoxe {} ; ou elle se réduit à la réponse *conséquentialiste* {v}.

L'un des plus grands philosophes du droit du xx^e siècle, H.L.A. Hart (1907–1992), a proposé lui aussi une version du *rétributivisme hybride*⁸⁰. Pour comprendre l'idée, revenons au *rétributivisme négatif*. Plus tôt, nous avons dit qu'il était incomplet : il interdit au juge d'être trop sévère, mais pas d'être trop clément. Pour simplifier radicalement, on peut dire que H.L.A. Hart a proposé d'utiliser le *conséquentialisme* pour compléter la théorie. Concrètement, le conséquentialisme interdit au juge d'être trop clément : au minimum, il faut que la peine dissuade tout de même de futurs crimes. Il est interdit d'être plus sévère qu'un rétributiviste (le plafond), et il est interdit d'être plus clément qu'un conséquentialiste (le plancher) ; ces deux règles sont absolument obligatoires. Or, dans le cas des jumeaux siamois, le plancher est plus haut que le plafond ! Comme nous l'avons vu, le *rétributivisme négatif* demande l'amnistie générale ; et à l'inverse, le *conséquentialisme* demande très probablement une punition collective. La juge doit donc violer une, voire deux règles absolument obligatoires. En temps normal, la solution au problème est donc qu'il n'y en a pas : {}.

Certaines théories hybrides vont plus loin que Hart. Le *rétributivisme limitant* de Norval Morris, par exemple, affirme que le rétributivisme marque le plafond *et* le plancher⁸¹. Soyons de bonne foi et accordons à Morris deux bornes rétributives, les bornes {*min*} et {*max*}. Puis, pour pouvoir se décider plus précisément, Morris nous demande de

⁷⁹ BERMAN Mitchell N., « Punishment and Justification », *Ethics*, vol. 118, n° 2, 2008, pp. 258-290.

⁸⁰ HART H.L.A. et GARDNER John, *Punishment and Responsibility*, Oxford University Press, 2008, 277 p.

⁸¹ FRASE Richard S., « Limiting Retributivism: The Consensus Model of Criminal Punishment », in TONRY Michael (éd.), *The Future of Imprisonment in the 21st Century*, Oxford Univ. Press, 2003, pp. 83-120.

s'appuyer sur l'utilitarisme. C'est une belle idée, mais pourquoi ne pas être un utilitariste tout court ? Le conséquentialisme est plus précis, certes, mais ses réponses dépendent des conditions empiriques, qui sont imprévisibles et capricieuses. Selon quoi, pour le bien-être commun, l'utilitariste peut vouloir sacrifier un bouc-émissaire innocent ou récompenser un criminel avéré plutôt que de le punir. D'où l'intérêt, d'après Morris, de combiner les deux approches. Cette théorie est peut-être une des plus consensuelles actuellement⁸². Quoi qu'il en soit, dans notre cas, le maximum et le minimum sont 0 et 10 ans de prison ; tout repose donc sur un raisonnement *utilitariste* $\{v\}$. Qui plus est, cette idée de borner l'utilitarisme semble sortir de nulle part, surtout s'il s'avère, comme nous l'avons défendu plus tôt, que rétributivisme et utilitarisme sont au chevet de victimes différentes (la personne et le corps social).

Mais la théorie la plus profondément hybridée est peut-être le *rétributivisme conséquentialiste*⁸³. Cette approche revient, en quelque sorte, à pirater l'utilitarisme : plutôt que de chercher à maximiser le bonheur moyen de tous, on lui ordonne de maximiser *le mérite moyen* des parties prenantes d'une affaire judiciaire. Dans cette version, par exemple, la juge est autorisée à gracier un criminel pour qu'il en dénonce deux autres, ou un pire criminel que lui ; de même, elle serait autorisée à punir un innocent pour en sauver deux autres, ou un plus pur innocent encore. Les problèmes commencent lorsqu'on cherche à formaliser cette approche et à l'appliquer au problème des siamois. En effet, comment calculer leur mérite moyen ? Dans sa critique, Deitch pense que ce *rétributivisme conséquentialiste* est un jeu à somme nulle⁸⁴. En effet, en punissant le coupable d'une peine p , on le rapproche de x de son mérite véritable. Ce faisant, on punit aussi l'innocent d'une peine p , ce qui l'éloigne de x de son mérite véritable. Collectivement, leur mérite moyen bouge de $x - x = 0$... c'est-à-dire pas du tout (fig. v). Au final, toutes les peines comprises entre $\{min\}$ et $\{max\}$ sont moralement équivalentes. Nous pouvons noter cette solution bornée $[min ; max]$.

Cette solution bornée nous laisse au point de départ du problème des siamois : notre première intuition, c'était bien d'exclure une réponse en-deçà du minimum et au-delà du maximum. Pire, elle nous interdit d'avoir une préférence morale supplémentaire. On

⁸² RISTROPH Alice, « Desert, Democracy, and Sentencing Reform Criminal Law », *Journal of Criminal Law and Criminology*, vol. 96, n° 4, 2005, p. 1302.

⁸³ CHRISTOPHER Russel L., « The Prosecutor's Dilemma: Bargains and Punishments », *Fordham law review*, vol. 72, n° 1, 2003, pp. 93-168.

⁸⁴ DEITCH Brittany, « Retributivist Theories' Conjoined Twins Problems », *art. cit.*, p. 139.

devrait alors choisir une peine au hasard, par liberté d'indifférence—par peur de finir comme l'âne de Buridan, mort de faim d'avoir trop longtemps hésité entre deux seaux d'avoine identiques⁸⁵. En quelque sorte, on doit interpréter $[min ; max]$ comme une distribution de probabilité uniforme continue. Dit encore autrement, c'est une loterie.

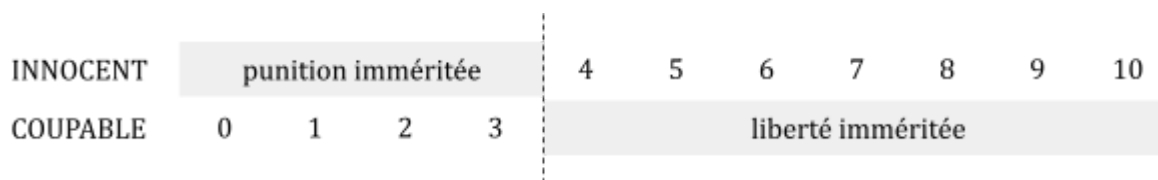


fig. v : Contre-argument du jeu à somme nulle.

Or, d'après les théories de la rationalité, il existe nécessairement une seule et unique punition $\{p\}$ rationnellement équivalente à la loterie $[min ; max]$. Vous pouvez vous en convaincre en analysant une loterie quelconque : par exemple, combien seriez-vous prêt-e à payer pour avoir une chance sur dix de gagner à l'euro-millions ? Si ce ticket était mis aux enchères, jusqu'à combien monterait la mise ? À partir de combien le ticket serait-il objectivement surévalué ? Parmi les théories de la rationalité, l'approche la plus orthodoxe est la *théorie de l'utilité espérée*, dont nous avons déjà parlé (*expected utility theory* de von Neumann et Morgenstern, section 3.2)⁸⁶. D'après cette dernière, un agent rationnel ne fait pas de différence entre un gain w et une loterie avec une espérance de gain w . En l'occurrence, l'espérance de notre loterie c'est tout simplement la moyenne⁸⁷ ; la solution $[min ; max]$ est donc rationnellement équivalente à la peine moyenne $\{moy\}$.

En fait, contrairement à la première analyse de Deitch, le *rétributivisme conséquentialiste* peut défendre la punition $[min ; max]$, mais aussi la punition $\{moy\}$. Mais le problème profond, c'est que ce raisonnement va dans les deux sens. Ainsi, une juge rationnelle pourrait tout autant transformer la punition $\{p\}$ quelconque, en une loterie $[0 ; 2 \cdot p]$. En fait, par définition ou presque, il n'y a aucune raison rationnelle de s'y opposer. Nous

⁸⁵ Cette expérience de pensée est nommée d'après Jean Buridan (~1300– ~1358), un philosophe scolastique qui a défendu que l'être humain avait une liberté d'indifférence. Mais c'est chez Aristote (384–322 AEC) qu'on trouve cet exemple spécifique (un humain qui hésite entre étancher sa faim ou sa soif).

ARISTOTE, *Traité du ciel*, DALIMIER Catherine et PELLEGRIN Pierre (trad.), Paris : Flammarion, 2004, 465 p.

⁸⁶ VON NEUMANN John et MORGENSTERN Oskar, *Theory of Games and Economic Behavior*, Princeton : Princeton University Press, 2004 (1944), 776 p.

⁸⁷ Comme il existe plusieurs types de moyennes, il faut préciser qu'il s'agit d'une moyenne arithmétique, qui prend la forme $(a + b)/2$.

pourrions, c'est vrai, postuler une nouvelle règle : « à choisir, la justice doit toujours être déterministe plutôt que stochastique, elle doit toujours appliquer une peine univoque plutôt qu'une loterie (toutes choses égales par ailleurs) ». Cette règle qui interdit la *justice stochastique* est vraiment très intuitive, c'est presque une évidence que personne ne prend la peine de préciser. Mais comment exactement pourrions-nous la justifier ? La seule idée qui vient en tête, c'est parce que des loteries de ce genre ne sont pas adaptées à des humains qui ne peuvent pas comprendre leur sens, ni intuire toutes ces équivalences rationnelles. Si j'ai eu de la chance ou de la malchance, vais-je encore comprendre que ça ne change rien au fait que je mérite ma peine malgré tout ? En somme, cela revient à postuler que la punition est avant tout une affaire de communication. En voulant aider le rétributivisme conséquentialiste, nous venons par inadvertance de rejeter tout un pan des théories du mérite au profit d'une autre approche, les *théories communicatives du mérite*.

5.3. Théories communicatives du mérite

Il existe une manière plus prometteuse de justifier l'existence de la rétribution et cette justification prend la forme de *théories communicatives du mérite*. C'est « peut-être la version la plus influente de ces dernières décennies »⁸⁸ en plus d'avoir une longue histoire, puisqu'Adam Smith (1723–1790) la défendait déjà dans sa *Théorie des sentiments moraux*, et Émile Durkheim (1858–1916) après lui dans *De la division du travail*⁸⁹. L'idée de cette approche est simple : *la punition est un message* qu'on peut traduire par « la société dénonce ou se dissocie de ce crime ». Bien sûr, la société dénonce plus fortement le meurtre que l'insulte, par exemple ; pour cette raison, elle punit plus sévèrement le premier. En bref, la peine la plus juste, c'est celle qui transmet au mieux le message de la société.

⁸⁸ « Punishment as Communication. Perhaps the most influential version of retributivism in recent decades seeks the meaning and justification of punishment as a deserved response to crime in its expressive or communicative character »

HOSKINS Zachary et DUFF Antony, « Legal Punishment », in ZALTA Edward N. et NODELMAN Uri (dir.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, 2024.

⁸⁹ ALLARD Aurelien, *Le mérite : signification, possibilité et valeur*, Thèse de doctorat, Paris 8, 2019, pp. 175 et 179.

Les théories dites 'dénonciatrices' défendent que le juge « s'adresse avant tout à la société dans son ensemble »⁹⁰. Chez Joel Feinberg, par exemple dans son essai de 1965 *The Expressive Function of Punishment*, la peine a quatre fonctions pour la communauté : se désolidariser du criminel, se désolidariser de ce type de crime, sacraliser la loi et, enfin, proclamer que tous les autres suspects sont innocents⁹¹. C'est ce dernier point qui pose un problème très évident dans le cas des siamois. Plus la juge veut proclamer l'innocence du jumeau innocent, plus les trois autres messages de la punition en pâtissent. La théorie de Feinberg ne nous dit pas ce qu'il se passe lorsqu'une fonction n'est pas réalisée : il semble que chaque fonction du message soit obligatoire. Dans ce cas, le problème des siamois n'a pas de réponse et on retombe sur la solution vide {}.

On peut sauver Feinberg et le rendre moins absolu, l'autoriser à affaiblir une fonction pour en renforcer une autre. Même s'il existait un compromis optimal, rien dans l'approche de Feinberg ne nous dit comment le trouver. Faudrait-il recourir à un argument conséquentialiste ? La théorie dénonciatrice deviendrait alors une simple théorie instrumentale inventée par un utilitariste de la règle (section 3.1). De manière générale, les siamois génèrent des tensions dans la théorie dénonciatrice, parce qu'elle ne nous dit pas comment communiquer clairement sur une punition collective ou une amnistie générale, ni justifier ce type de peine. Pour résoudre ce problème, il faudrait assortir la peine de prison (ou la libération) d'un discours explicatif ; c'est alors l'explication qui communique—et donc rend justice—à elle seule, plutôt que la peine de prison. À quoi sert encore une peine de prison dans ce cas ? À première vue, cela affaiblit considérablement l'intérêt d'une approche communicative du mérite⁹².

Dans la théorie de Nozick, par opposition, le message s'adresse spécifiquement au condamné⁹³. Dans une thèse de 2019 dédiée à la philosophie du mérite, Allard synthétise la théorie ainsi : « le but de la peine est de mettre en lien le criminel et les valeurs que ce criminel a bafouées, en lui montrant que les valeurs ont bien un impact

⁹⁰ *Ibid.*, p. 179.

⁹¹ FEINBERG Joel, « The Expressive Function of Punishment », *The Monist*, vol. 49, n° 3, The Open Court Publishing Co, Chicago, 1965, p. 397-423.

⁹² Nous verrons plus tard que, pour communiquer la punition, la peine de prison peut être remplacée par une peine de shifgrethor (section 9.2). Techniquement, on peut donc punir par le simple discours. Cependant, prison et shifgrethor ont un point commun : elles servent d'unité de compte universelle du mérite. Un discours explicatif *personnalisé* nous ramène à la situation de créativité punitive typique de la loi du talion, comme on a pu l'observer en Europe à l'époque des lois barbares (section 9.3).

⁹³ NOZICK Robert, *Anarchy, State, and Utopia*, New York : Basic Books, 1974, xvi+367 p.

dans sa vie »⁹⁴. Dans la version d'Anthony Duff, la manière d'y parvenir est plus spécifique que chez Nozick, puisque ce qui justifie la peine c'est de faire ressentir regret ou remors au condamné⁹⁵. Or, dans tous les cas, le siamois innocent n'a rien bafoué du tout ! Les théories de Nozick ou de Duff doivent délivrer deux messages antithétiques, à deux personnes différentes, par le même signifiant. Plus le message envers le coupable est intelligible, plus la communication avec le jumeau innocent en pâtit, et vice-versa. Cela ressemble à un jeu à somme nulle où tous les messages sont équivalents, ce qui correspond à la solution [*min* ; *max*].

Pour la *théorie expressive de la rétribution*, de Jean Hampton, le message s'adresse plutôt à la victime⁹⁶. Le but de la punition est de soigner un affront envers la valeur ou la dignité de la victime, qui a subi une blessure morale. Plus précisément, pour Hampton, le message implicite d'une action criminelle, c'est que la victime vaut moins que le criminel. Le message de la punition est que la victime—par la main de la justice—peut tout autant subjuguier le criminel, et donc que les deux sont bel et bien égaux⁹⁷. Malheureusement, la théorie de Hampton s'effondre aussi devant le problème des siamois. En effet, en punissant le jumeau coupable (ce qui soigne l'affront envers la victime), on punit le jumeau innocent, qui devient lui aussi une victime (de la justice, cette fois). On transvase petit à petit la blessure de la victime (ou des victimes) chez un autre innocent⁹⁸. Cela ressemble à un jeu à somme nulle où toutes les peines comprises entre [*min* ; *max*] sont équivalentes—c'est d'ailleurs aussi l'avis de Brittany Deitch⁹⁹.

Néanmoins, par le même argument de la loterie que nous avons utilisé plus tôt (section 5.2), nous pouvons transformer ces peines [*min* ; *max*] en une peine rationnellement équivalente, la punition {*moy*}. De plus, nous pouvons montrer que ces deux peines sont rhétoriquement inégales : on communique beaucoup plus efficacement le mérite des

⁹⁴ ALLARD Aurelien, *Le mérite : signification, possibilité et valeur*, op. cit., p. 179.

⁹⁵ DUFF R A, *Punishment, Communication, and Community*, New York : Oxford University Press, 2000, 272 p.

⁹⁶ HAMPTON Jean, « An Expressive Theory of Retribution », in CRAGG Wesley (éd.), *Retributivism and its Critics*, vol. 4., Stuttgart: Franz Steiner Verlag, 1992, pp. 12-13.

⁹⁷ *Idem*.

⁹⁸ Deitch remarque que le problème est encore plus insoluble dans le cas où les jumeaux seraient en procès l'un contre l'autre : on transvaserait la blessure de la victime chez elle-même ! Mais il est vrai que dans un tel cas, la situation semble insoluble, même pour notre sens moral intuitif ou pour le conséquentialisme. La théorie de Hampton ne donne pas de réponse, certes, mais ce n'est pas pour autant un défaut fatal.

⁹⁹ DEITCH Brittany, « Retributivist Theories' Conjoined Twins Problems », *University of Cincinnati Law Review*, vol. 87, n° 1, 2018, p. 139.

siamois par une peine univoque que par une loterie. Ainsi, par pragmatisme et contrairement à l'opinion de Deitch, les théories communicatives de Nozick, Duff et Hampton défendent la peine *{moy}*. Concrètement, si le coupable mérite une peine de dix ans de prison, il faut donner 5 ans de prison à chaque siamois.

Cependant, cette stratégie est un emplâtre sur une jambe de bois. En effet, tout cela ne se justifie que tant que les humains sont irrationnels. Si nous devions punir des très-rationnels siamois (chez Nozick et Duff) ou si la victime était un être de raison (chez Hampton) alors nous n'aurions plus aucune raison de favoriser une peine univoque plutôt qu'une loterie. Au contraire même ! Comme notre théorie parvient à une loterie, il serait étrange d'ajouter une transformation supplémentaire à la nature de la peine. Et ce genre de cas de figure ne serait pas impossiblement rare. Si Chang Bunker avait agressé un spécialiste des théories de la rationalité plutôt qu'un médecin, Jean Hampton devrait défendre l'installation d'une roue de la chance dans la salle d'audience. De même, si les frères Godina avaient assisté par hasard à une conférence de von Neumann ou de Morgenstern, il aurait fallu tirer leur punition d'un grand chapeau (probablement celui de Nozick ou de Duff venus assister au procès).

Je me moque gentiment, mais en réalité, les théories communicatives du mérite n'ont même pas besoin d'admettre ces conséquences un peu exotiques. Il existe une méthode détournée qui permet malgré tout de transformer la solution $[min ; max]$ en *{moy}*. Jusqu'à présent, nous avons tenté de tenir compte du mérite de bien des gens : le jumeau innocent et le jumeau coupable (rétributivisme classique et ses dérivés, rétributivisme conséquentialiste, théorie communicative de Nozick), la victime passée (théorie expressive de la rétribution de Hampton), puis aussi les victimes futures (rétributivisme limitant), etc., sans parvenir à nos fins. Pourtant, ce tour d'horizon n'est pas exhaustif. Qui reste-t-il ? Le dernier groupe social qui n'a pas été mentionné, ce sont les autres prisonniers. Le problème philosophique qui en résulte est très concret : est-ce que notre décision envers les siamois impacte les autres condamnés, et comment ? Et si nous tenons compte de ces derniers, est-ce que nous pouvons résoudre le problème des siamois ? Les trois chapitres qui suivent seront dédiés à cette question, d'abord par un argument empirique un peu bancal, puis par un argument ontologique basé sur la nature du mérite et les fonctions de la punition.

6. LA LOI DE CONSERVATION DE LA PUNITION

6.1. Mérite de la communauté des prisonniers

Dans un texte de la philosophe Alice Ristroph, une observation empirique a retenu mon attention¹⁰⁰. Appelons cela la *loi de conservation de la punition*. Dans nos pays occidentaux, les prisons sont constamment surchargées, au point où cela contraint de nombreux juges à donner des peines plus courtes¹⁰¹. Et quand les prisons ne sont pas surchargées, on observe que les juges font en sorte de les remplir¹⁰². Il s'avère que la capacité des établissements pénitentiaires (et plus largement, des 'ressources correctionnelles') est une contrainte importante à la sévérité des punitions.

Par exemple, si on détruit subitement la moitié des places de prison ou si on double subitement le nombre de criminels, alors plutôt que de purger 100% de sa peine "juste", chacun en purgera la moitié, faute de place. À l'inverse, si de la place se libère alors chaque prisonnier tendra à purger une peine plus longue. La conséquence de cette idée, c'est que quand un juge gracie arbitrairement le prisonnier Dupont, qui mérite dix ans de prison, en réalité ces dix ans de prison *ne disparaissent pas de la circulation* : sa place sera occupée par d'autres. Statistiquement, ce sont les autres prisonniers qui les purgent collectivement (ce qui représente, en Suisse, peut-être un jour de plus par tête de pipe en moyenne).

Par ailleurs, remarquons que notre raisonnement s'applique aussi à d'autres types de punition, comme les amendes. Pour s'assurer que la police est rentable, par exemple, il est courant que leur hiérarchie demande aux policiers un certain quota de contraventions à distribuer, quitte à ce que ces derniers redoublent d'efforts pour l'atteindre¹⁰³. À l'inverse, il arrive que l'administration soit surchargée par le nombre d'amendes à recouvrir, à tel point qu'une partie des contraventions finissent prescrites

¹⁰⁰ RISTROPH Alice, « Desert, Democracy, and Sentencing Reform Criminal Law », *Journal of Criminal Law and Criminology*, vol. 96, n° 4, 2005, p. 1295.

¹⁰¹ GAES Gerald G., « The Effects of Overcrowding in Prison », *Crime and Justice*, vol. 6, [University of Chicago Press, University of Chicago], 1985, p. 96.

¹⁰² BONNEAU Daniel D. et HALL Joshua C., « The impact of new prison construction on the likelihood of incarceration », *The American Journal of Economics and Sociology*, vol. 83, n° 1, 2024, p. 199–222.

¹⁰³ CARPENTER Dick M., SWEETLAND Kyle et McDONALD Jennifer, *The price of taxation by citation: Case studies of three Georgia cities that rely heavily on fines and fees*, Arlington, Va.: Institute for Justice, 2019, 56 p.

et ne sont jamais payées¹⁰⁴. Si un policier ou un fonctionnaire gracie arbitrairement le contrevenant Dupond, qui mérite 100.– d’amende, cette amende *ne disparaît pas de la circulation*. Statistiquement, d’autres contrevenants tirés au hasard la payeront à sa place, alors qu’ils auraient pû y échapper autrement. Comparé au cas de la prison, la présence explicite d’une loterie complexifie le raisonnement ; mais il suffit (à nouveau) de réfléchir en termes d’espérance de gain, et cela en revient au même sur le plan rationnel—en tout cas d’après von Neumann et Morgenstern.

Nuançons quelque peu, car ma description du fonctionnement des prisons et des amendes est un peu extrême. En réalité, si Dupont ou Dupond sont graciés, une partie de leur peine disparaîtra bel et bien de la circulation. Peut-être que seul un dixième, voire un centième de leur punition sera “redistribuée” aux autres condamnés. Mais toujours, quelque chose a bel et bien été conservé et redistribué ! Dans un État, la quantité de punition à distribuer n’est pas *parfaitement élastique* : gracier les Dupont—ou les punir trop sévèrement—aura toujours au moins une légère répercussion sur la communauté des autres condamnés.

Appliquons concrètement cette loi au problème des siamois. Le siamois coupable a commis un crime qui mérite dix ans de réclusion. Il a, en quelque sorte, une cellule réservée pour 10 ans ; c’est ce qu’il mérite. Si cette cellule est occupée pour moins de dix ans ou pour plus de dix ans, alors, *statistiquement*, le système judiciaire va changer légèrement la peine du reste des prisonniers pour profiter de la place qui se libère, même en partie.

6.2. Critique de cette loi de conservation

Voilà pour le mécanisme général. Ce dernier repose sur une série de suppositions douteuses que je vais maintenant détailler. Nous reviendrons plus tard sur ces dernières ; mais momentanément, pour les besoins du raisonnement, j’ai besoin que vous consentiez à suspendre votre incrédulité.

¹⁰⁴ BALBONI Julien, « Des amendes pénales impayées pour près d’un milliard d’euros en cinq ans », *L’Echo*, 08.10.2024, [www.lecho.be/economie-politique/belgique/general/des-amendes-penales-impayees-pour-pres-d-un-milliard-d-euros-en-cinq-ans/10567716.html], consulté le 15.12.2024]

Premièrement, supposons que la communauté des prisonniers purge, en moyenne, exactement la peine qu'elle mérite collectivement. Dans ce cas—le mérite des non-prisonniers inchangé par ailleurs—changer la peine de prison des condamnés, même très légèrement et très indirectement, est mal. Cette supposition mène à l'affirmation suivante : la cellule du siamois coupable, réservée pour dix ans, doit être occupée pendant 10 ans (le mérite des non-prisonniers inchangé par ailleurs).

Mais est-ce que le mérite des non-prisonniers est “inchangé par ailleurs” ? Oui, au prix d'une deuxième supposition : nous devons adopter une théorie rétributive du type “ensemble borné” [*min* ; *max*], comme pour les théories communicatives du mérite, ou le rétributivisme conséquentialiste. Dans ce cas, le mérite collectif des siamois ne change *virtuellement* pas, quelle que soit la peine de prison qu'ils subissent (sauf au-delà de 10 ans de prison). Cette supposition nous permet de simplifier l'affirmation précédente : la cellule du siamois coupable doit être occupée pour dix ans.

Mais qu'est-ce que cela veut dire “occuper une cellule pour dix ans” ? Troisièmement, supposons qu'en prison, les siamois occupent la place de deux condamnés non-siamois, ni plus ni moins. Or, occuper une cellule pendant dix ans, cela revient à occuper deux places de prison pendant cinq ans. Cette dernière supposition nous amène à notre conclusion : la peine juste à donner aux siamois, c'est une peine univoque de cinq ans de prison. Comme elle correspond à la moyenne des peines méritées par chaque jumeau, nous pouvons la noter {*moy*}.

Grâce à toutes ces suppositions, nous pouvons compléter les *théories communicatives du mérite* et le *rétributivisme conséquentialiste*. Nous obtenons la solution {*moy*}, une réponse univoque et qui ne viole pas trop fortement notre sens moral, le tout sans supposer que les humains sont irrationnels. Hourra ! mais pas vraiment. Vous pouvez cesser de suspendre votre incrédulité, car tout notre raisonnement repose sur trois prémisses douteuses.

La première prémisse douteuse, c'est que tous les prisonniers purgent *exactement* la peine qu'ils méritent. Sans cette supposition, notre résultat s'effondre très rapidement. Même avec une *conservation de la punition* parfaite, il suffit que la communauté des prisonniers purge collectivement 10 ans de trop (ou de pas assez) seulement, pour que la peine optimale à donner aux siamois change radicalement et qu'il faille les libérer (ou les condamner à dix ans de prison). Et avec une élasticité un peu forte, l'effet est

démultiplié. Or, il est très probable qu'en moyenne les juges soient au moins un peu biaisés. Les seules peines justes plausibles sont donc plutôt $\{min\}$ ou $\{max\}$.

La deuxième prémisse douteuse, c'est que les siamois réservent la place de deux condamnés, plutôt qu'un seul ou que zéro. Bien entendu, c'est empiriquement faux : du point de vue du système carcéral, les deux siamois prennent la place physique d'une personne avec des besoins spécifiques, comme il y en a beaucoup d'autres. Par conséquent, au pire la peine juste devrait être de dix ans de prison.

La troisième prémisse douteuse consiste à violer la *guillotine de Hume*, et donc à affirmer que savoir ce qui *est* suffit pour savoir ce qui *doit être*¹⁰⁵. Pour faire le pont entre faits observables et préceptes moraux, d'après Hume, il faut toujours ajouter un argument spécial. Certes, en pratique la punition des États n'est jamais parfaitement élastique, mais il est peut-être moralement nécessaire qu'elle le devienne. Pire, même si elle est inélastique en pratique, il est peut-être moralement nécessaire d'agir comme si elle était parfaitement élastique, par fiction philosophique. Dans cette situation, la communauté des prisonniers et son mérite deviennent hors-sujet.

6.3. Au juste... qu'est-ce que le mérite ?

Nous venons de lister trois critiques dévastatrices pour la loi de conservation de la punition. Pour espérer échapper à ces critiques, nous devons démontrer trois choses. Premièrement, la communauté des prisonniers purge toujours exactement la peine qu'elle mérite, *par définition*. Deuxièmement, la place physique qu'occupent les siamois importe peu : tout ce qui importe c'est qu'ils soient exactement *deux* à faire l'expérience de la punition. Troisièmement, le simple acte de punir génère *nécessairement* les deux conséquences ci-dessus. Certes, peut-être que gracier les Dupondt ne change rien à la vie concrète des autres condamnés. Pour autant, cela *change le sens de leur punition* et la rend plus sévère qu'avant.

En somme, pour sauver la loi de conservation de la punition, qui était une observation contingente, nous la transformons en une propriété nécessaire. Pour construire cette nouvelle version, nous aurons besoin d'un argument un peu plus touffu et abstrait. Nous

¹⁰⁵ HUME David, *Traité de la nature humaine*, III, I, I, SALTEL Philippe (trad.), Paris : Flammarion, 1999.

allons devoir nous intéresser à l'ontologie de la punition, mais surtout à l'ontologie de... la monnaie. J'espère que cette question vous passionne, puisqu'elle va nous occuper pour les prochains chapitres.

*
**

Ce qui devrait vous surprendre néanmoins, c'est que jusqu'à présent je n'ai pas défini la "punition" ni le "mérite". À cela, deux bonnes raisons. La première raison, c'est que cela nous importait peu pour l'instant : le problème des siamois posait problème aux théories rétributives, quel que soit le sens exact de ces mots. La deuxième bonne raison, c'est que ces concepts restent assez flous et changeants, selon les théories et les auteurs ; mesurer le mérite est même « une sorte de mystère métaphysique » selon les mots de la philosophe Alice Ristroph¹⁰⁶. Je vais maintenant tenter d'en éclaircir le sens, en essayant de rester aussi proche que possible de la réalité empirique.

Pour dissoudre ce mystère, je pense qu'il est pertinent de s'intéresser aux travaux des métaphysiciens sur un autre phénomène social : la monnaie. L'idée n'est pas si absurde. Pensons à quel point la métaphore monétaire est prégnante quand on parle de punition et de mérite : un tel doit "payer pour ses crimes", un autre a "remboursé sa dette à la société", telle action est "cher payée", alors que tel autre "s'en tire à bon compte". Au-delà du sens commun, j'ai deux raisons plus profondes de m'intéresser à la philosophie de la monnaie. La première, c'est que les questions centrales sont très proches : "qu'est-ce qu'une juste peine" est analogue à "qu'est-ce qu'un juste prix". L'échange entre De même, le phénomène d'inflation et de déflation monétaire a quelques analogies avec le phénomène de conservation de la punition que l'on vient d'analyser. La deuxième, ce sont les travaux de Georg Simmel (1858–1918), dont la *Philosophie de la monnaie* est souvent encensée par les spécialistes de ce sujet de niche¹⁰⁷. Or, la théorie monétaire de Simmel, une fois appliquée au mérite, démontre exactement ce que nous voulons. Chez

¹⁰⁶« Exactly how much punishment an offender deserves is something of a metaphysical mystery, or so it has appeared to be in the past. »

RISTROPH Alice, « Desert, Democracy, and Sentencing Reform Criminal Law », *art. cit.*, p. 1293.

¹⁰⁷ SIMMEL Georg, *The Philosophy of Money*, BOTTOMORE Tom et FRISBY David (trad.), London : Routledge, 2011 (1900), 537 p.

NAIN Aditya et JUNG p. G., *Understanding Money: Philosophical Frameworks of Monetary Value*, Taylor & Francis, 2021, 133 p.

ZELMANOVITZ Leonidas, *The Ontology and Function of Money: The Philosophical Fundamentals of Monetary Institutions*, Lexington Books, 2015, 471 p.

Simmel, par définition et par essence, détruire la monnaie de M. Dupond nous rend collectivement plus riches, alors même que nous n'avons pas plus de papier-monnaie dans les mains. De même, par définition et par essence, effacer la peine de M. Dupond rend plus sévère la punition de tous les autres, alors même qu'ils ne sortent pas plus tard de prison. Voilà donc notre programme : construire une *théorie monétaire de la punition*.

7. LA PHILOSOPHIE DE LA MONNAIE DE SIMMEL

— *I mean, it's one banana Michael. What could it cost, ten dollars ?*

7.1. Substantialisme : la monnaie comme commodité

Pour comprendre la *Philosophie de la monnaie*¹⁰⁸ de Georg Simmel (1858–1918), il faut d'abord comprendre l'opposition entre deux très anciennes traditions philosophiques qu'on appelle le *substantialisme* et le *relativisme*. Ces dernières ne sont pas d'accord sur le type d'objets utilisables comme monnaie. D'après les relativistes, la monnaie fonctionne par pure convention sociale arbitraire. Diogène de Sinope (~413 à ~323 AEC), par exemple, aurait proposé de remplacer l'argent par des osselets¹⁰⁹. Par contraste, d'après les substantialistes, l'objet-monnaie doit aussi avoir une valeur économique en tant qu'objet de consommation. Aristote (384–322 AEC), par exemple, pense que « la valeur de la monnaie est stable, durable et digne de crédit, *comme si elle lui était inhérente* »¹¹⁰. Simmel, lui, se range dans le camp de Diogène.

Pour comprendre pourquoi Simmel donne raison à Diogène contre Aristote, il faut reformuler le problème philosophique un peu plus concrètement. En pratique, ces

¹⁰⁸ SIMMEL Georg, *The Philosophy of Money, op. cit.*, 537 p.

¹⁰⁹ *Ibid.*, p. 18.

¹¹⁰ HELMER Étienne, « De quoi la monnaie est-elle la mesure ? Aristote, Diogène et Platon ou la convention monétaire dans l'horizon du politique », *Cahiers d'économie Politique*, vol. n° 72, n° 1, 2017, p. 12.

théories cherchent à répondre à la question compliquée suivante : « pourquoi est-ce qu'une pomme vaut 0.50 CHF, plutôt que 0.- ou 100.- CHF ? ».

Pour y faire face, la théorie *substantialiste* commence par s'attaquer à une question simpliste : pourquoi est-ce qu'une pomme vaut autant qu'une pomme, plutôt que 0 ou 100 pommes¹¹¹ ? Selon ces théories, c'est parce qu'une pomme a à peu près les mêmes nutriments et la même saveur qu'une autre (valeur-utilité), ou parce qu'une pomme demande autant de travail à récolter qu'une autre (valeur-travail), ou parce qu'une pomme est aussi rare qu'une autre (offre-demande). En somme, on analyse l'objet-pomme et sa trajectoire causale dans le passé (travail) et dans le futur (utilité), et le rapport entre les deux (offre-demande). Par conséquent, *troquer* une pomme contre une pomme est juste, alors qu'une contre zéro ou cent est injuste. Voilà pour la question simpliste.

Fort de cette nouvelle sagesse, la théorie substantialiste *remplace une pomme par une pièce de 50 cts* et applique le même type de raisonnement. La pièce a de la valeur parce que les États en demandent pour les impôts (valeur-utilité), ou parce que cela coûte de faire un prêt bancaire et de frapper monnaie et parce qu'empêcher fraude et contrefaçon demande un contrôle constant (valeur-travail), ou parce que les États contrôlent la masse monétaire et l'inflation (offre-demande). La pièce s'échange contre la pomme car elles recèlent la même utilité, le même travail, car l'offre de pomme correspond à la demande de monnaie. Par conséquent, *troquer* une pomme contre une pièce de 50 ct. de CHF est juste, alors que la troquer contre 0.- ou 100.- CHF est injuste.

En bref, d'après les *substantialistes*, la valeur de la monnaie a la même origine que les autres commodités ; et une transaction monétaire est de même nature qu'un troc entre deux commodités. Ces deux principes s'impliquent d'ailleurs l'un et l'autre. Or, Simmel les rejette tous deux : la "vraie" monnaie n'est pas une commodité, et acheter ou vendre n'est pas du "vrai" troc.

¹¹¹ C'est, par exemple, le principe à la base de la théorie monétaire de John Locke (1632-1704) : une pomme vaut une pomme, et deux pommes valent deux fois plus qu'une pomme. « Hence it is evident, that an equal quantity of Silver is always of equal value to an equal quantity of Silver. »

LOCKE John, « Considérations sur les conséquences de la diminution de l'intérêt et de l'augmentation de la valeur de l'argent », in *Écrits monétaires*, BRIOZZO Florence (trad.), Paris : Classiques Garnier, 2011 (1696), 463 p.

7.2. Simmel : la monnaie comme symbole abstrait

Selon Simmel, qu'est-ce que la "vraie" monnaie, donc ? C'est la solution à un problème d'histoire philosophique (ce qui est différent d'une question historique). Simmel et la plupart des philosophes de la monnaie commencent par postuler une fiction historique : avant la monnaie, seul existait le troc. Puis, l'invention de la monnaie fait reculer le troc et le remplace, de plus en plus, par une économie dite "monétaire". La monnaie idéale, c'est la monnaie telle qu'elle fonctionne dès que le troc disparaît entièrement, dès que toutes les interactions sociales deviennent des interactions économiques qui sont explicitement monnayées. C'est peut-être un monde terrible, mais c'est dans ce monde fictif et simpliste que les propriétés de la monnaie seront les plus faciles à analyser. Là-bas, à quoi ressemble la monnaie ?

Dans ce monde caricatural, Simmel peut donc poser une question simpliste : pourquoi est-ce que, durant un laps de temps, la quantité totale Q_{tot} de biens économiques à échanger, vaut la masse monétaire M_{tot} qui circule en un an ? La réponse est assez simple à fournir : parce que c'est une tautologie, une vérité analytique. Puisqu'il n'y a pas de troc dans ce monde, par définition, la monnaie est l'intermédiaire de tous les échanges économiques. L'ensemble de toute la valeur économique échangée (valeur-utilité, valeur-travail, offre-demande, etc.) *a la même valeur* que l'ensemble de toute la monnaie qui circule¹¹² : $Q_{\text{tot}} = M_{\text{tot}}$.

On pourrait s'amuser à postuler que, par conséquent, toute cette monnaie qui circule a exactement la même utilité que toute l'économie réunie, ou a demandé autant de

¹¹² Attention, monnaie "qui circule", et non pas "monnaie". Si un seul billet de banque sert à faire mille transactions successives, on compte mille fois le même billet ; c'est tout le concept de *vélocité de la monnaie* en économie. De la même manière, si Balthazar Picsou stocke dans son coffre *cinq milliards de quadruplimilliards de décuplimilliards de multiplimilliards de fantastimilliards de centrifugilliards de dollars et seize cents*, sans qu'un seul centime sorte de son coffre (il n'utilise pas sa fortune et n'est pas taxé dessus), alors tout se passe exactement comme si cette somme monstrueuse n'existait pas.

En fait, Simmel reprend assez clairement la théorie monétaire du philosophe sceptique écossais David Hume (1711-1776), qu'il développe dans ses *Discours politiques*. On peut ainsi lire aux paragraphes §12 et §13 de *De la monnaie* :

« It seems a maxim almost self-evident, that the prices of every thing depend on the proportion between commodities and money, and that any considerable alteration on either has the same effect, either of heightening or lowering the price. »

« If the coin be locked up in chests, it is the same thing with regard to prices, as if it were annihilated; if the commodities be hoarded in magazines and granaries, a like effect follows. »

HUME David, « De la monnaie », « De l'intérêt », « De la balance commerciale », in *Discours politiques*, Édimbourg : Kincaid & Donaldson, 1752, pp. 281-307.

besogne que tout le reste de l'économie. C'est le postulat substantialiste—qui est empiriquement assez étrange, on s'en rend vite compte. À l'inverse, on pourrait s'amuser à postuler que la monnaie n'a aucune valeur économique, elle n'a que de la valeur *nominale* : intrinsèquement, un Vreneli vaut 20.– CHF et c'est tout. On peut postuler qu'il n'y a rien de plus que cette définition-là. La monnaie n'est alors qu'un chiffre qui *symbolise et quantifie* une valeur économique, sans que les nombres dans un compte en banque aient eux-mêmes une valeur dûe à leur substance. C'est le postulat de Simmel. Explorons ses conséquences.

Fort de cette nouvelle sagesse, Simmel *remplace Q_{tot} par une pomme et remplace M_{tot} par une pièce de 50 ct.* Plus exactement, il sait qu'une pomme représente un billardième (par exemple) des biens échangés Q_{tot} ; et qu'une pièce de 50 centimes représente, elle aussi, un billardième de la masse monétaire en circulation M_{tot} (toujours sur le même laps de temps). Tant que la proportionnalité est conservée, la transaction sera juste. Par conséquent, une pièce de 50 cts. de CHF symbolise *bien* la valeur de la pomme, alors que 0.– ou 100.– CHF la symbolisent *mal*. Ou plus exactement, la pièce symbolise la valeur *relative* de la pomme, par rapport à l'ensemble de l'économie.

Le calcul se tient, mais les prémisses semblent douteuses : lorsque vous cherchez à connaître le prix d'une pomme, vous ne commencez pas par estimer la somme totale de l'économie mondiale, même au doigt mouillé. Mais l'idée de Simmel a du mérite, car on peut remarquer que vous commencez par estimer le coût relatif d'une pomme dans votre budget. En l'occurrence, par rapport à votre budget pour votre repas de midi... qui est une proportion de votre budget de la semaine, qui est une proportion de votre budget annuel—et ce budget annuel, c'est-à-dire votre salaire et vos revenus, il est souvent lui aussi une proportion du budget d'une entreprise ou d'un État, qui est, de proche en proche, une simple proportion de tous les échanges économiques. Or, chacun de ces budgets est censé réaliser une tâche précise : ils correspondent chacun à une quantité de valeur économique fixe (utilité, somme de travail, etc.). Autrement dit, nous ne passons pas directement de Q_{tot} à une pomme et de M_{tot} à 50 cts. En revanche, collectivement, nous appliquons le raisonnement simmelien de proche en proche, de sous-ensemble en sous-ensemble—ce qui est cognitivement plus accessible. Mais cela ne change rien à l'observation de Simmel : la pièce de 50 centimes symbolise la valeur *relative* de la pomme par rapport à l'ensemble de l'économie ; ainsi, la pièce n'a pas besoin d'avoir de valeur économique, *elle peut se contenter d'être un symbole.*

7.3. La loi de conservation de la valeur monétaire

L'alternative que décrit Simmel ici est lourde de répercussions philosophiques. En effet, entre le *monde de commodités* des substantialistes et le *monde de prix* des relativistes, la monnaie n'a pas la même fonction, ne nécessite pas les mêmes croyances, ni la même ontologie. Nous développerons ces conséquences par la suite. Pour l'instant, revenons à nos moutons : le problème des siamois et, en particulier, la loi de conservation de la punition.

La conséquence immédiate de la proportionnalité que décrit Simmel, c'est que si on divise arbitrairement les avoirs monétaires de chacun par deux, alors la monnaie restante double de valeur d'échange ; et vice-versa si on double la masse monétaire. De même, si je détruis l'argent de M. Dupond, qui représente $k\%$ de la masse monétaire, alors la valeur d'échange du reste de la monnaie mondiale augmente à raison de $\frac{k}{100-k}$ $\cong k\%$.

En pratique cependant, la fortune de M. Dupond est une goutte d'eau dans l'océan et les flux monétaires ne sont de loin pas assez homogènes pour que l'effet soit décelable. Qui plus est, la destruction et la création monétaire influencent en retour la taille de l'économie, ce qui fait qu'ils ne provoquent souvent ni déflation, ni inflation. Mais conceptuellement, il y a bel et bien un impact sur le reste de la communauté des utilisateurs de monnaie. Cette idée, elle est exactement analogue à celle que nous cherchions à construire pour le problème des siamois.

Plus largement, les doctrines substantialistes et relativistes décrivent deux types de valeur différentes : la valeur substantielle et la valeur nominale. On peut admettre que la monnaie recèle un peu des deux en même temps ! Si nous revenons au cas de M. Dupond, détruire la monnaie de ce dernier fait disparaître la valeur substantielle, mais conserve la valeur nominale (exactement comme si cette valeur nominale était distribuée entre toutes les autres pièces de monnaie). Cet impact est particulièrement peu décelable, il est vrai ; mais statistiquement, effet il y a.

On pourrait cependant affirmer que Simmel se trompe totalement, et que la monnaie n'a *que* de la valeur substantielle, et *aucune* valeur symbolique. On pourrait même admettre qu'elle a de la valeur symbolique, mais que cela représente un dysfonctionnement qu'il faut ignorer pour le genre de dilemme moral qui nous intéresse. Ce qu'il nous manque,

c'est donc une preuve que l'euro, la livre ou le dollar ont besoin d'avoir une valeur relative pour pouvoir *fonctionner* comme monnaie.

7.4. Transition fonctionnelle de la monnaie

Le débat entre substantialisme et relativisme est profondément lié à un autre débat : à quoi sert la monnaie ? Depuis le travail théorique de l'économiste William Stanley Jevons (1835–1882) et son *Money and the Mechanism of Exchange*¹¹³, en première analyse, on attribue quatre fonctions distinctes à la monnaie : *réserve de valeur*, *intermédiaire des échanges*, *norme de paiement différé* et *unité de compte*.

Dans les théories substantialistes, la monnaie a deux fonctions possibles : une fonction de *réserve de valeur* ou une fonction d'*intermédiaire des échanges*. Concrètement, si je veux de l'argent, est-ce pour l'enterrer dans un coffre au trésor ou pour l'échanger sur-le-champ contre quelque chose d'autre ? Pour le mercantiliste John Locke (1632–1704), par exemple, s'enrichir implique d'échanger son surplus de travail contre des trésors—la monnaie étant la meilleure *réserve de valeur* à accumuler¹¹⁴. Pour le libéral Adam Smith (1732–1790), par contraste, s'enrichir passe par la division du travail au sein de la communauté¹¹⁵—un tel système nécessite un *intermédiaire des échanges*, la monnaie, qui circule constamment et sans entraves¹¹⁶. De manière générale, ce débat porte sur la *vélocité* de la monnaie, qui mesure combien de fois un billet change de mains, en moyenne, en un temps donné.

Au sein des théories relativistes, la monnaie a deux autres fonctions possibles : une fonction de *norme de paiement différé* ou une fonction d'*unité de compte*. D'après les

¹¹³ JEVONS William, *Money and the Mechanism of Exchange*. New York: D. Appleton and Company, 1875, 342 p.

¹¹⁴ Parmi les défenseurs de cette fonction de réserve de valeur, Piteau cite les économistes David Cass (1937–2008) et Karl Shell (né en 1938).
PITEAU Michel, « Monnaie de compte et système de paiements chez James Steuart : Quel rôle pour la stabilité bancaire ? », *Revue économique*, vol. 53, n° 2, 2002, p. 246.

¹¹⁵ Plus exactement, chacun s'enrichit en se professionnalisant dans un métier spécifique, ce qui produit plus de richesse au global. Néanmoins, un tel système pose le problème de la double coïncidence des besoins : le menuisier ne veut pas forcément être payé en farine par le meunier. Ce problème se résout par l'introduction d'un *intermédiaire des échanges* que tout le monde acceptera (avant tout pour l'échanger ensuite à quelqu'un d'autre).

¹¹⁶ Parmi les défenseurs de cette fonction d'intermédiaire des échanges, Piteau cite l'économiste américain Irving Fisher (1867–1947).
Idem.

chartalistes comme Georg Friedrich Knapp (1842–1926), la monnaie est une unité de mesure des dettes, qui n’a pas de valeur en soi—on parle de *norme de paiement différé*¹¹⁷. Par contraste, pour les tenants de ce qu’on appellerait aujourd’hui la théorie quantitative de la monnaie, comme David Hume (1711–1776), la monnaie sert à mesurer la valeur des autres biens économiques—on parle alors d’*unité de compte*. Or, un emprunt et un achat se distinguent dans leur rapport à la création et à la destruction monétaire. Rembourser un emprunt détruit la monnaie, qui existe pour être re-transformée en trésor symbolique inerte (c’est sa fonction) ; alors que faire un achat fait circuler la monnaie, qui existe pour être constamment ré-échangée et jamais détruite (c’est sa fonction). Entre ces deux doctrines de politique économique, on retrouve donc un débat analogue sur la vitesse monétaire.

En résumé, chaque devise représente un compromis en termes de vitesse et de conventionnalité, ce qui est aussi un compromis fonctionnel (fig. vi). Mais pourquoi est-ce que toutes les devises actuelles suivent, à peu près, le même compromis fonctionnel ? Pourquoi est-ce qu’on ne trouve pas des monnaies fortement spécialisées en réserve de valeur, et d’autres en unité de compte¹¹⁸ ? Pour les économistes et les philosophes, c’est un signe qu’il existe un compromis optimal, qui fait office d’attracteur : malgré le chaos et la complexité de l’économie, malgré les intérêts croisés et contradictoires, la plupart des monnaies dérivent vers une de ces quatre fonctions plutôt que vers les autres (fig. vi). Par ailleurs, plus elles s’éloignent de cet attracteur, moins elles seront adoptées et plus leur fonctionnement risque de devenir chaotique. Cet attracteur, nous l’appellerons par la suite la *fonction essentielle* de la monnaie.

¹¹⁷ D’après les chartalistes, formellement, l’État ne “paye” pas le travail des fonctionnaires et des militaires à sa solde, il s’endette auprès d’eux. En réalité, la monnaie est un gage symbolique qui prouve qu’à la fin du mois, l’État “en doit une” à ses employés. Puis, ces gages circulent (ou non) dans l’économie et lorsque quelqu’un (peu importe qui) “en doit une” aux pouvoirs publics—sous forme de taxes, impôts, amendes—il peut rendre ces gages à l’État plutôt que de payer en nature.

Cette tradition est marquée par James Steuart (1707–1780), Léon Walras (1834–1910) et John Maynard Keynes (1883–1946). Parmi les théoriciens contemporains, Piteau cite également Charles Kindleberger (1910–2003), Aglietta et Delaplace.

Idem.

¹¹⁸ Au niveau macroéconomique, on observe toutefois quelques légères spécialisations fonctionnelles. Les économistes actuels distinguent par exemple la monnaie scripturale et la “monnaie centrale”, une norme de paiement différé spécialisée que les banques (et personne d’autre) empruntent à une banque centrale. Cependant, le taux de change entre monnaie scripturale et monnaie centrale est fixe, plutôt que flottant. Voici donc de meilleurs exemples : le dollar est le moyen d’échange le plus courant pour de nombreuses matières premières, et il sert d’unité de compte mondiale pour les économistes. Par contraste, l’or a maintenu sa réputation de réserve de valeur en temps de crise, tandis que le franc suisse a pu être un moyen d’échange privilégié pour commercer avec des états parias comme l’Allemagne nazie ou l’Afrique du Sud sous l’apartheid.

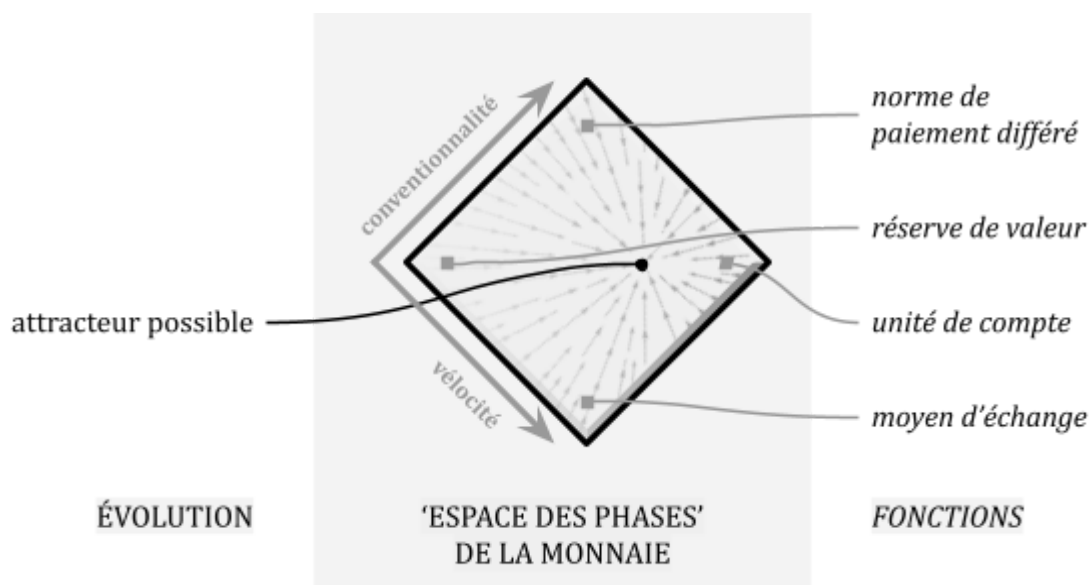


fig. vi : Schéma des fonctions de la monnaie.

La principale originalité de Simmel, c'est de dire que cette fonction essentielle change avec le temps. Ce changement graduel se fait dans une direction précise, malgré de très nombreuses oscillations : les conceptions substantialistes sont progressivement grignotées par la conception symboliste. Quant à notre système économique actuel, il se trouve quelque part entre-deux (fig. vi). Plus exactement, Simmel veut montrer que la monnaie est née comme un *moyen d'échange*, mais elle devient de plus en plus une *unité de compte*. Cela présuppose que la vélocité de la monnaie a toujours été plutôt élevée. Ce parti-pris, Simmel ne le justifie pas explicitement ; il part de la conviction que la monnaie est née avec le troc, et que Smith a raison contre Locke. Mais à première vue, l'approche de Simmel s'adapte à n'importe quelle transition, dans n'importe quel ordre, voire à un processus cyclique ! Pour comprendre pourquoi Simmel opte pour *cette* transition fonctionnelle parmi toutes celles possibles, il faut s'intéresser un peu plus aux implications ontologiques et psychologiques de cette transition fonctionnelle.

7.5. Transition psychologique de la monnaie

Dans quel sens évolue la monnaie ? Simmel revient à l'idée centrale du relativisme : la monnaie est une convention sociale. Cela veut dire que pour faire du troc ou utiliser la monnaie, les hommes ont besoin, au minimum, d'une *théorie mentale naïve*

de l'origine de la valeur. La transition fonctionnelle de la monnaie est donc, aussi, une transition de modèle mental.

Pour analyser cette transition psychologique, Simmel va se baser largement sur Immanuel Kant¹¹⁹ (1724–1804). Ainsi, il va distinguer deux manières d'utiliser (et de penser) la monnaie : le *moyen* et le *but*¹²⁰. En particulier, cette distinction kantienne correspond très bien à la distinction entre la fonction de *moyen d'échange* et celle de *réserve de valeur*. Dans un monde de troc (comme chez Smith), je veux me débarrasser le plus vite possible de ma monnaie, afin d'acheter une pomme. La monnaie est donc un simple moyen. À l'inverse, dans un monde de trésors (comme chez Locke), je veux conserver le plus possible ma monnaie. Elle est donc une fin en soi¹²¹. Simmel apporte ensuite une couche de nuance à cette description. En effet, les deux heuristiques (la monnaie comme moyen et comme fin) coexistent nécessairement ; c'est seulement leur rapport de force qui change avec l'évolution culturelle. Pour Simmel, par ailleurs, il est impossible qu'une pure monnaie soit atteinte, car elle cesserait de fonctionner¹²² ; là aussi, il semble s'appuyer sur des idées de Kant sur la psychologie humaine.

Comme nous le voyons, ce mécanisme permet de décrire une transition de *vélocité* de la monnaie. Mais elle ne permet pas vraiment d'expliquer une transition de *conventionnalité*, qui donnerait raison au relativisme par rapport au substantialisme ! Simmel semble conscient de cette lacune. Il va ainsi s'évertuer à montrer que la fonction d'unité de compte implique—dans notre psyché en tout cas—la fonction de réserve de valeur. Or, dans le graphe que nous venons de construire (fig. vi), ces deux fonctions s'opposent en tout point ! Cette tension au sein de la *Philosophie de la monnaie* est particulièrement problématique. Mais on peut y remédier. Pour ce faire, il faut dépasser Kant et ajouter une nouvelle variable au modèle psychologique de Simmel, grâce à deux concepts plus récents : le *même* et la *transition évolutionnaire d'individualité*.

¹¹⁹ Kant a également écrit sur la monnaie ; mais peu de littérature a été publiée à ce sujet.

JAMES Aaron, « Kant, Innes, and the Copernican Turn in Monetary Theory », in TINGUELY Joseph J. (éd.), *The Palgrave Handbook of Philosophy and Money*, Basingstoke : Palgrave Macmillan, 2024

¹²⁰ SIMMEL Georg, *The Philosophy of Money*, *op. cit.*, pp. 229–230.

¹²¹ La théorie de Simmel se tient. Elle décrit très bien, par exemple, la psyché de Balthazar Picsou—canard le plus riche du monde et avare parfait—dont la motivation principale est d'accumuler le plus d'or et de dollars possible dans un coffre au trésor gargantuesque, sans rien dépenser.

¹²² « Money performs its services best when it is not simply money, that is when it does not merely represent the value of things in pure abstraction. »

SIMMEL Georg, *The Philosophy of Money*, *op. cit.*, p. 163.

8. UNE TRANSITION D'INDIVIDUALITÉ MONÉTAIRE

8.1. Une transition ontique de la monnaie

La théorie de Georg Simmel décrit, nous l'avons vu, une transition fonctionnelle de la monnaie, ce qui nécessite une transition psychologique. Mais en plus de tout cela, elle décrit une transition *ontique*. Comme exemple de transition ontique dans la nature, on peut citer l'eau liquide qui change de phase et se cristallise sous forme de glace ; ou bien l'uranium 238 qui devient, à la fin de la chaîne de désintégration, du plomb 206. De la même manière, à l'échelle d'un millénaire, la monnaie tend elle aussi à changer d'ontologie. Pour décrire cette transition, Simmel forge un nouveau concept, la « spiritualisation croissante de la monnaie »¹²³. D'après Simmel, ce qu'il se passe, c'est que la monnaie passe graduellement d'une commodité à une idée¹²⁴. Plus exactement, elle passe d'une pure substance à une pure fonction—quoique Simmel précise que la valeur substantielle est elle aussi fonctionnelle, au fond¹²⁵. Cela a des implications ontiques, dont Simmel lui-même est bien conscient¹²⁶.

Néanmoins, la “spiritualisation” n'est pas vraiment un mécanisme explicatif, ni un moteur de l'évolution culturelle. C'est plutôt ce que Kant appelle une condition de possibilité : il faut que le cerveau humain soit *capable* de conceptualiser une monnaie purement symbolique pour pouvoir l'utiliser. Mais cela n'implique pas qu'une monnaie symbolique finit, tôt ou tard, par gagner en puissance ! La “spiritualisation” n'implique pas la “croissance”. On peut imaginer une “spiritualisation partielle”, où la monnaie ne gagne plus en symbolisme au-delà d'une certaine limite. On peut tout autant imaginer

¹²³ *Ibid.*, p. 198.

¹²⁴ Pour Simmel, la seule façon de générer une transition ontologique, c'est par un acte mental (qu'il soit conscient ou inconscient), qui rend les choses commensurables (comparables). En fait, tout acte mental consiste à des petites transitions ontologiques de ce genre, à petite échelle.

Disons, par exemple, que j'aie deux pommes dans ma cuisine. J'y vois *deux* pommes. Une personne qui n'aurait pas voulu mettre de l'unité dans la diversité, aurait plutôt dit “j'ai pommi et pommu dans ma cuisine”, par exemple. Ainsi, pour Simmel, la monnaie d'un monde de prix c'est la version ultime de cet acte mental-là : on considère que toutes les choses sont comparables, car tout, absolument tout a un prix.

¹²⁵ *Ibid.*, p. 167.

¹²⁶ Simmel compare notamment son approche à celle de Baruch Spinoza (1632–1677), célèbre pour avoir décrit le réel comme un monde de matière (une étendue) *et* un monde d'idées, lesquels changent de manière identique en parallèle l'un de l'autre ; les deux mondes n'étant, sur le plan ontologique, qu'une seule et même chose (monisme).

une “spiritualisation cyclique”, où la monnaie change régulièrement de fonction économique : *intermédiaire des échanges* → *unité de compte* → *norme de paiement différé* → *réserve de valeur* → *intermédiaire des échanges* → *etc.* Cette idée n’est pas saugrenue : dans son *Dette : 5000 ans d’histoire*¹²⁷, l’anthropologue David Graeber (1961–2020) a défendu l’idée que la monnaie oscille comme un yo-yo entre une fonction d’intermédiaire des échanges (monnaie des marchands, essentialisée dans son rapport à un support métallique) et de norme de paiement différé (monnaie virtuelle de crédit, gérée politiquement)¹²⁸. On peut même imaginer une “spiritualisation chaotique”, où la monnaie change de degré de conventionnalité et de vélocité au petit bonheur la chance¹²⁹. En fait, formellement, on peut imaginer à peu près n’importe quel champ de vecteurs au sein de l’espace des phases que nous avons construit plus tôt (fig. vii).

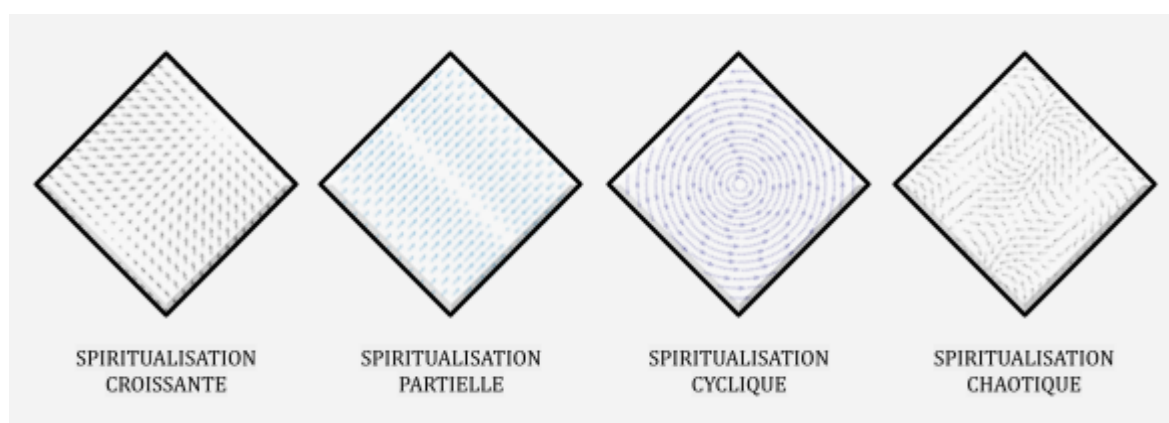


fig. vii : Trois alternatives à la “spiritualisation croissante” de la monnaie.

La question reste donc presque entière : pourquoi la monnaie change-t-elle de fonction ? À mon sens, la clé du problème se trouve dans la notion de *transition évolutionnaire d’individualité*. Pour le comprendre, il faut revenir à la distinction entre substantialisme et relativisme que décrivait Simmel. Dans le *monde de commodités* des substantialistes, la pièce de 50 centimes a une valeur inhérente, qui est la même que la valeur inhérente de la pomme. Pour analyser la valeur, ils conçoivent les pièces de monnaie comme des atomes, si l’on veut. Dans un *monde de prix*, en revanche, la pièce

¹²⁷ GRAEBER David, *Debt : The First 5000 Years*, New York : Melville House Publishing, 2011, 624 p.

¹²⁸ BADAIRE Quentin, « Gilles Deleuze face aux hétérodoxies monétaires régulationnistes, post-keynésiennes et marxistes : jalons pour un dialogue transdisciplinaire sur la monnaie et la dette », p. 3, in *Institutionnalismes monétaires francophones : bilan, perspectives et regards internationaux*, Lyon (France), 1–3 Juin 2016.

¹²⁹ On peut penser, par exemple, aux *mouvements browniens*, lorsque de petits objets qui flottent sur l’eau ou dans l’air semblent bouger tout seuls. Le phénomène est dû à l’accumulation de chocs chaotiques entre ces objets et les molécules du gaz ou du liquide.

de 50 centimes symbolise la valeur relative de la pomme par rapport à l'ensemble de l'économie. Simmel conçoit la masse monétaire comme une matière unitaire, dont les pièces de monnaies sont des portions. Autrement dit, l'un décrit une collection d'individus, l'autre décrit un individu à décomposer.

Ce qui se joue-là, sur le plan ontologique, c'est une transition d'individualité. Les objets-monnaies ont renoncé à une partie de leur individualité et de leur agentivité propre, et la masse monétaire est déjà en partie devenue un agent individuel. En somme, je propose d'appliquer le concept de *transition évolutionnaire d'individualité* à la monnaie. Pour ce faire, j'aimerais réutiliser la théorie d'Ellen Clarke (section 4.1). Il faut donc éliminer une objection assez coriace, de prime abord : la monnaie n'est pas vivante. Je pense qu'on peut décomposer cette objection en deux. La monnaie n'est pas biologique et elle n'a pas d'agentivité.

Premièrement, la monnaie n'est pas biologique, mais ce n'est pas grave. Quand Ellen Clarke définit le *levels problem*, elle ne présuppose pas que l'individu est un organisme biologique. Elle présuppose seulement qu'il est soumis à la sélection darwinienne¹³⁰ (section 4.1). De même pour nous : pour que l'analyse de Clarke marche, notre individu a juste besoin d'être soumis à une sélection culturelle. L'évolution culturelle et l'évolution biologique ont en effet beaucoup de mécanismes analogues.

Deuxièmement, la monnaie n'a pas d'agentivité, mais ce n'est pas grave non plus. Nous avons déjà défendu plus tôt la pertinence d'une "heuristique de personnification" au sein de la biologie (section 3.4). Or, cette heuristique a aussi été appliquée à l'évolution des cultures pour devenir la *mémétique*, une doctrine que l'on doit au philosophe Daniel Dennett (1942–2024). Elle consiste à traiter les objets culturels comme des agents—appelés *mèmes*—qui vivent en parasitant les esprits humains et qui se répliquent par contagion sociale, et qui sont soumis à sélection culturelle. Enfin, il faut mentionner que Simmel n'a rien contre la personnification : dans son propre texte, il

¹³⁰ CLARKE Ellen, « A levels-of-selection approach to evolutionary individuality », *Biology & Philosophy*, vol. 31, n° 6, 2016, p. 893–911.

accepte explicitement de personnifier certaines formes de cognition, par heuristique uniquement¹³¹.

8.2. La transition évolutionnaire d'individualité de la monnaie

Nous avons le feu vert pour tenter d'appliquer la théorie d'Ellen Clarke à la transition fonctionnelle et ontologique de la monnaie. Peut-on parler de *transition évolutionnaire d'individualité* ? Pouvons-nous identifier des *mécanismes de démarcation* et des *services d'ordre* au sein de la monnaie ?

Pour le bien de l'analyse, simplifions notre problème. Comme nous l'avons vu, en biologie, l'individualité se répartit en poupée russe. Chez Clarke, nous pouvons distinguer trois situations qui nécessitent trois types d'analyse différents : un *life*, un *levels* et un *gaia problem*. Le *life problem* concerne les individus minimaux, atomiques. C'est le cas d'un monde de troc substantialiste. À l'inverse, le *gaia problem* concerne le superindividu maximal, fusion de tous les individus au monde. C'est le cas d'un monde de prix simmelien¹³². Simplifions donc notre problème et partons du principe que la monnaie transitionne directement des individus atomiques à un superindividu maximal, sans individus intermédiaires : il n'y a qu'un seul *levels problem*. Ce que la théorie de Clarke prédit dans ce cas-limite, c'est que la monnaie va progressivement perdre ses *mécanismes de démarcation* et le remplacer par des *services d'ordre*. Autrement dit, Clarke prédit que la monnaie va changer de fonction. Sa première fonction doit aussi servir de mécanisme de démarcation entre devises ; sa seconde fonction doit aussi agir comme un service d'ordre au sein de la devise. Développons cette idée.

¹³¹ Simmel suggère l'idée que la monnaie existe dans l'inconscient—or, l'inconscient est un agent invisible caché au sein du cerveau, une boîte-noire fourre-tout de la psychologie dont l'existence n'a jamais été démontrée. C'est d'ailleurs ce que souligne Simmel. Pour cette raison, il le présente comme une heuristique utile.

SIMMEL Georg, *The Philosophy of Money, op. cit.*, pp. 157–158, p. 230.

¹³² Il est intéressant de noter que Simmel rapproche la monnaie idéale de la notion de Dieu. En effet, l'essence de Dieu, écrit-il, c'est que toutes les diversités et les contradictions du monde parviennent à une unité en lui. Il est difficile de ne pas y voir un parallèle avec Gaïa, l'organisme ultime qui les unifie tous et marque la fin de la dissension et de la concurrence au sein du vivant.

Ibid., p.237.

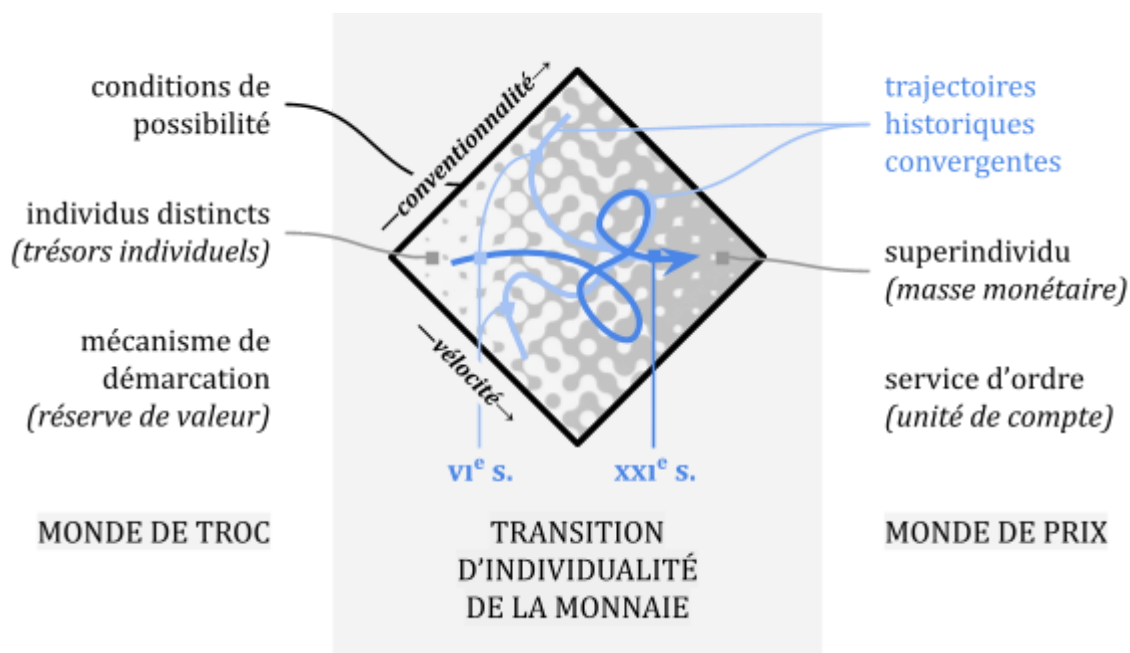


fig. viii : Schéma de la transition d'individualité monétaire.

Un *mécanisme de démarcation* vise à favoriser la sélection culturelle de toutes les monnaies contre toutes. Il force un objet-monnaie à conserver sa valeur, alors même que la valeur d'échange d'autres objets-monnaies s'effondre ou explose. Cette fonction monétaire a déjà un nom : une "réserve de valeur" (fig. viii). À l'inverse, un *service d'ordre* vise à empêcher toute sélection culturelle entre objets-monnaies. Il force un objet-monnaie à changer de valeur lorsque la valeur de tout autre objet-monnaie change. Ce phénomène de conservation de la valeur monétaire, Simmel la décrit en long et en large ! Cette fonction monétaire a déjà un nom : une "unité de compte" (fig. viii).

La philosophie de la monnaie de Simmel décrit donc une transition évolutionnaire d'individualité au sens de Clarke. Grâce à cette notion, nous pouvons converger vers les thèses principales de Simmel. Tout d'abord, comme la transition d'individualité ne peut se faire que sur un seul axe, cela nous donne une bonne raison de croire au modèle de "croissance linéaire" que Simmel défend, par contraste avec un modèle "cyclique" ou un modèle "brownien" (c'est-à-dire chaotique) (section 8.1). Ensuite, une transition d'individualité est toujours très légèrement entamée et jamais parfaitement terminée (section 4.2). Nous avons donc une bonne raison de croire que l'euro, la livre ou le dollar ont au moins une part de valeur relative.

Par ailleurs, ces deux fonctions de réserve de valeur et d'unité de compte s'opposent à la fois en terme de vélocité et de degré de conventionnalité. Cela a deux implications. La

première, c'est que les fonctions d'*intermédiaire des échanges* et de *norme de paiement différé* sont plus instables que les autres, au moins au long terme. La deuxième implication, c'est que nous pouvons enfin relier les différents mécanismes psychologiques que décrit Simmel, comme la distinction entre fin et moyen (section 7.5), et la spiritualisation de la monnaie (section 8.1). En dernière analyse, ces mécanismes ont une fonction commune. Si la monnaie gagne en vélocité et/ou gagne en symbolisme abstrait, c'est grosso modo à cause des mêmes pressions de sélection culturelle.

Avant de revenir définitivement à nos moutons siamois, je me permet de solidier cette théorie que je propose ici, sur le plan empirique et théorique. Sur le plan empirique, on peut en effet vérifier que cette transition d'individualité de la monnaie a bien eu lieu historiquement, par exemple dans l'histoire du territoire suisse. Par ailleurs, on peut montrer que progresser dans la transition d'individualité est une stratégie gagnante pour survivre aux pressions de sélection. Sur le plan théorique, il s'avère que l'ontologie de cette 'transition d'individualité' de la monnaie peut être interprétée d'un grand nombre de manières, sans que cela pose réellement problème.

8.3. Histoire monétaire de la Suisse

Pour obtenir ce résultat, nous sommes partis du principe que la monnaie ne traverse qu'une seule et unique transition évolutionnaire d'individualité, sans étapes intermédiaires. Relaxons maintenant cette hypothèse. Ce que nous venons de montrer, c'est que les deux fonctions de la monnaie—réserve de valeur et unité de compte—agissent l'une contre l'autre. La *réserve de valeur* résiste aux transitions évolutionnaires d'individualité, tandis que l'*unité de compte* les encourage. Dans les faits, rien n'interdit plusieurs transitions d'individualité successives, ou qui se recourent dans le temps. En fait, alors que l'histoire du vivant est marquée par un nombre assez restreint de transition d'individualité successives, l'histoire monétaire est une succession continue et chaotique de réformes qui servent à fusionner des masses monétaires disparates entre elles.

Pour le comprendre en pratique, regardons la monnaie du point de vue d'un cambiste, dont le métier consiste à *troquer* des devises les unes contre les autres. Son but est de

thésauriser les devises qui seront, pense-t-il, bientôt en plus forte demande ; et de se débarrasser des autres avant que la demande ne chute. Pour le cambiste, la monnaie n'est pas une *unité de compte*, elle n'a qu'une fonction de *réserve de valeur*. Dans un monde de cambiste, les monnaies ont une valeur économique, mais elles n'ont pas de valeur nominale. Par extension, il traite les différentes masses monétaires (\$, £, €, etc.) comme des atomes, des individus en concurrence les uns contre les autres. Pour mesurer la progression d'une transition évolutionnaire d'individualité, il suffit en quelque sorte de mesurer la viabilité, voire l'ubiquité du travail de cambiste ! Plus largement, plus nous sommes tôt dans une ETI, plus les marchands sont des cambistes malgré eux ; et lorsque la transition est accomplie, les derniers cambistes spécialisés doivent mettre la clé sous la porte.

Pour l'illustrer maintenant, prenons l'histoire monétaire de la Suisse médiévale et moderne. Au v^e siècle, à cause du retrait des troupes romaines de la région, l'économie monétaire s'effondre lentement et on revient au troc dans l'usage quotidien. Il faut attendre six siècles pour que la monnaie redevienne un outil de la vie courante (xi^e s.), et deux de plus pour que l'économie monétaire soit à nouveau florissante (xiii^e s.). Aux xvi^e et xvii^e siècles, sur le territoire suisse, entre quinze et vingt seigneurs frappent leurs propres devises (aux noms parfois identiques, mais aux valeurs différentes). Plus largement, il y a une « prépondérance des pièces italiennes, espagnoles, françaises et allemandes dans la masse en circulation (jusqu'à 60% du total vers 1600) »¹³³. En guise de politique monétaire, les divers gouvernements de Suisse publient des listes officielles de cours de change, qu'ils mettent à jour plus ou moins périodiquement¹³⁴. Dans une telle situation, chacun est son propre cambiste, qu'il le veuille ou non : « tous ceux qui étaient amenés à voyager à travers la Suisse (marchands grands et petits, entrepreneurs militaires, diplomates, baillis) devaient s'initier aux divers systèmes en vigueur »¹³⁵. Il faut attendre 1850 pour que l'État fédéral nouvellement créé parvienne à imposer une monnaie unique, le franc suisse. Suite à cette réforme, en une année, « les cantons retirèrent [...] presque 66 millions d'anciennes pièces, se répartissant entre 319 variétés (frappe et valeur) »¹³⁶. En fin de compte, toutes ces masses monétaires locales

¹³³ PETER Markus, SCHMUTZ Daniel et DREGEN Bernard, « Monnaie », MARTIN Pierre-G (trad.), *Dictionnaire historique de la Suisse (HLS DHS DSS LIS)*, 18.02.2014,

¹³⁴ *Idem*.

¹³⁵ *Idem*.

¹³⁶ PETER Markus, SCHMUTZ Daniel et DREGEN Bernard, « Monnaie », *art. cit.*, 2014

fusionnent : la transition évolutionnaire d'individualité monétaire est complète, ce qu'on déduit facilement par l'observation qu'il n'y a plus la moindre spéculation possible à échanger des francs contre d'autres francs.

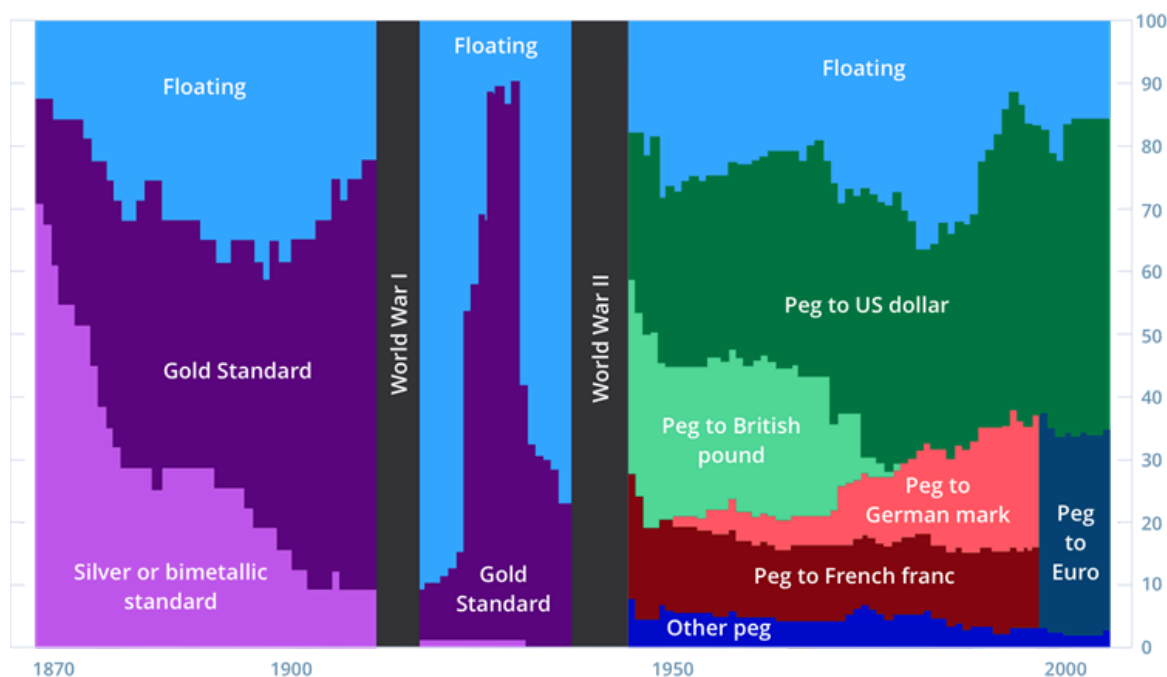


fig. ix : Évolution des masses monétaires en concurrence (xx^e s.).
(pourcentage de pays dont les devises sont indexées les unes aux autres)¹³⁷

Par après, on retrouve le même processus de transition évolutionnaire d'individualité à l'échelle européenne et mondiale. En 1865, la Suisse, la France, la Belgique et l'Italie définissent une sorte de monnaie commune (union monétaire latine), basée sur le cours de l'or (et au début, de l'argent)¹³⁸. L'union vacille à chaque crise d'envergure (guerre de 1870 et de 1914-18) et s'effondre dans les années 1920 ; puis, c'est au tour de l'étalon-or de vaciller (crise de 1929-36, guerre de 1939-45) et de s'effondrer ; en 1952, le dollar prend la place de l'étalon-or (accords de Bretton-Woods), mais ce système mondial s'effondre lui aussi avec la crise économique des années 1970¹³⁹. Le franc suisse flotte depuis. Mais la Suisse fait figure d'exception ici. Depuis le début du xxi^e siècle, près de la moitié des pays européens ont adopté une monnaie unique, l'euro. Et de manière générale, la grande majorité des pays du monde utilisent une devise à taux de change

¹³⁷ Changement de régimes monétaires à travers l'histoire. basé sur FEENSTRA Robert C. et TAYLOR Alan M., *International macroeconomics*, Londres : Macmillan, 2014 (2007), p. [n.c.], basé sur MEISSNER Christopher M. et OOMES Nienke, « Why do countries peg the way they peg? The determinants of anchor currency choice », *Journal of International Money and Finance*, vol. 28, n° 3, 2009, p. 522-547.

¹³⁸ PETER Markus, SCHMUTZ Daniel et DREGEN Bernard, « Monnaie », *art. cit.*, 2014

¹³⁹ *Idem.*

fixe (nonobstant les dévaluations), en général définie à partir du dollar ou d'une autre monnaie occidentale (fig. ix).

D'après Simmel, le moteur de ce changement tient à l'évolution de la coopération entre les humains, qui devient plus libre. Par là, il veut dire que chacun peut de plus en plus constamment changer de partenaires de coopération¹⁴⁰. Or, la transition d'individualité monétaire reproduit et facilite ce phénomène : par exemple, le citoyen ne dépend pas d'individus particuliers mais uniquement de leurs services—et grâce à la monnaie, ces services peuvent être achetés à n'importe quel concurrent (double coïncidence des besoins)¹⁴¹. En somme, pour Simmel, la monnaie représente des forces sociales abstraites. Une autre hypothèse évolutionnaire, c'est que la monnaie est une formidable *unité de mesure* de la valeur, qui nous rend un léger service cognitif : il est plus simple de faire du commerce et d'estimer la valeur économique des choses quand elles ont un prix précisément quantifiable. Par extension, les vendeurs qui proposent le même produit au même prix deviennent interchangeables—ce qui converge avec la thèse de Simmel.

L'explication la plus crédible cependant, comme notre résumé historique le laisse sous-entendre, tient sans doute à l'importance décisive de la politique monétaire dans la survie d'un État ou d'une communauté, en particulier en période de crise ou de croissance extrême¹⁴². Le cambiste, au fond, ne profite pas seulement de la concurrence entre devises ; il la renforce. En effet, le bon spéculateur n'anticipe pas l'évolution d'une monnaie, il anticipe l'évolution des croyances de la foule. Or, si la foule est constituée de cambistes qui s'anticipent les uns les autres, cela forme une boucle de rétroaction qui décuple l'effet des moindres sursauts. Si la moitié des utilisateurs du *thaler*, par exemple, sont des cambistes par nature (parce qu'ils ont dans les poches une demi-douzaine d'autres devises à troquer), l'effet spéculatif est dangereux. En revanche, en fusionnant le *thaler* avec la monnaie voisine, le *batz*, les cambistes *thaler-batz* disparaissent : cela diminue déjà une partie de cette influence spéculative. Or, plus le commerce de monsieur Tout-le-monde relève du commerce international, et plus

¹⁴⁰ Par ailleurs, Simmel avance que ce gain de liberté n'implique ni plus ni moins de bien-être ou de qualité de vie. Contrairement aux vœux pieux des libéraux, la concurrence de marché n'apporterait pas grand-chose aux hommes. La liberté dont Simmel parle est en fait la *liberté d'indifférence* décrite par le philosophe scolastique Jean Buridan (~1300– ~1358).

SIMMEL Georg, *The Philosophy of Money*, BOTTOMORE Tom et FRISBY David (trad.), London : Routledge, 2011 (1900), pp. 300–301.

¹⁴¹ *Ibid.*, p. 301.

¹⁴² Une économiste aura certainement une meilleure manière de l'expliquer encore.

monsieur Tout-le-monde renforce la force de frappe des cambistes, ce qui renforce l'effet spéculatif. Autrement dit, le moteur de la transition évolutionnaire d'individualité de la monnaie tient à l'évolution de la coopération entre des humains de plus en plus lointains et nombreux... ce qui en fait une coopération avec une toujours plus grande liberté de choix, comme l'expliquait Simmel.

8.4. Dernières remarques sur l'ontologie de la monnaie

Mais comment la monnaie peut-elle être en partie une étendue de matière et en partie une multitude d'atomes ? N'est-ce pas contradictoire ? Oui et non. Il serait plus juste de dire que la monnaie réalise un bouquet de fonctions monétaires en simultanée. Chaque fonction monétaire génère son propre type de valeur, qui implique sa propre ontologie. Schématiquement :

mécanisme de démarcation ∞ *réserve de valeur* ∞ *valeur substantielle* ∞ *atomisme* ;
service d'ordre ∞ *unité de compte* ∞ *valeur nominale* ∞ *matière unitaire*.

Un compromis fonctionnel produit donc, naturellement, un compromis ontologique ! D'ailleurs, tout cela se formalise très bien. Plutôt que d'utiliser une logique à deux valeurs de vérité (vrai/faux, chaud/froid, atomes/matière), on peut utiliser une *fuzzy logic*, une logique floue graduelle¹⁴³ (vrai...faux, chaud...froid, atomes...matière).

Peut-être que cette cohabitation ontologique reste contre-intuitive pour certains. Dans ce cas, imaginez simplement que deux monnaies de nature différente, chacune avec sa fonction, coexistent sous la même appellation et s'incarnent dans les mêmes objets, mais que l'une remplace progressivement l'autre. Cela importe peu ici. Il faut se souvenir que la monnaie est une construction sociale. Les théories monétaires des philosophes prétendent être un reflet de la "nature" de la monnaie ; mais la nature de la monnaie ne se réduit pas aux devises (pièces, billets, *digits* sur un compte, etc.), mais s'étend aussi aux convictions et aux pratiques des utilisateurs et des institutions... qui sont d'ailleurs influencées, en retour, par les théories des penseurs de l'économie¹⁴⁴. Les philosophes de la monnaie sont simplement occupés à formaliser et orienter une "opinion moyenne".

¹⁴³ KOSKO Bart, ISAKA Satoru, « Fuzzy logic », *Scientific American* 269 (1), 1993, pp. 76–81.

¹⁴⁴ MĀKI Uskali, « Reflections on the Ontology of Money », *Journal of Social Ontology*, vol. 6, n° 2, 2021, p. 245–263.

Ou plutôt, ils essayent de formaliser (et améliorer) la théorie naïve qui correspond le mieux aux pratiques qu'ils observent.

C'est pour cette raison que j'insère le concept de *transition évolutionnaire d'individualité* au cœur de la théorie de Simmel—bien qu'il ne l'ait jamais exprimé en ces termes, il est vrai. Cela s'explique probablement surtout par un hasard de contexte : les concepts de la *Philosophie de la monnaie* doivent plus à Kant qu'à Darwin ; d'autre part, Simmel a des convictions vitalistes¹⁴⁵. Malgré tout, je pense que cette reformulation que je propose permet une belle clarification de l'idée et de l'intuition sous-jacente décrite dans *Philosophie des Geldes*. Dans tous les cas, c'est sous cette forme-là que je compte appliquer la théorie de la monnaie simmélienne au problème du mérite.

*

**

Cette parenthèse métaphysique et économique est peut-être difficile à digérer. L'idée centrale que je veux retenir de cette petite aventure de deux chapitres, c'est que l'histoire d'un phénomène social aussi complexe que la monnaie peut se résumer par une idée assez simple : une *transition évolutionnaire d'individualité*. Progressivement, la monnaie commence par être une collection d'objets distincts sans liens les uns aux autres, troqués les uns contre les autres, valeur économique contre valeur économique ; puis elle devient une masse monétaire qui supprime les objets distincts, où on équivaut une *masse* de valeur nominale symbolique contre une *masse* de valeur économique.

Or, la punition suit elle aussi le cheminement de la monnaie et son histoire peut être schématisée de la même manière. C'est en tout cas ainsi que je propose d'analyser les théories communicatives du mérite. De cette manière, nous pourrions enfin prouver l'existence de la *loi de conservation de la punition*.

¹⁴⁵ LEE Monica et SILVER Daniel, « Simmel's Law of the Individual and the Ethics of the Relational Self », *Theory, Culture & Society*, vol. 29, n° 7-8, 2012, pp. 124-145.

9. UNE THÉORIE MONÉTAIRE DE LA PUNITION

9.1. Substantialisme : la valeur de la punition

Toutes les théories rétributives cherchent à répondre, entre autres, à la question compliquée suivante : pourquoi est-ce qu'un viol mérite 10 ans de réclusion, plutôt que la liberté ou la prison à vie ?

Pour répondre à cette question compliquée, nous pouvons commencer par analyser un modèle simpliste de la punition, la loi du talion : œil pour œil, dent pour dent... pomme pour pomme. Toute faute s'échange contre une punition de même nature. À partir de cette rétribution simpliste, les théoriciens classiques peuvent construire une théorie de la punition. Selon ces approches, la rétribution est équivalente car les dommages (souffrance psychologique, isolement social, pertes matérielles) sont identiques. Il s'agit d'une approche substantialiste de la punition. Grâce à cette analyse, on peut répondre à la question compliquée : les impacts négatifs d'un viol pour la victime et la société en général *valent*, en terme de conséquences, les impacts négatifs d'une peine de réclusion de dix ans de prison pour le coupable. Par ailleurs, par la suite, on peut moduler la loi du talion, pour tout un tas de raisons plus ou moins justifiées : par principe on sera plus clément, pour être moralement supérieur au criminel, pour permettre la réinsertion ; ou au contraire, on sera plus sévère par principe, pour être moralement supérieur au criminel, pour décourager la récidive, pour faire un exemple, etc.

Cependant, on peut prendre le problème par un autre bout. Imaginons un monde où la vengeance n'existe pas, où toute moquerie, toute exclusion sociale, toute vengeance interpersonnelle, aussi légère soit-elle, est traitée par un tribunal total. Imaginez une surveillance de masse assistée par IA ou bien une sorte de juge des enfers. Dans ce monde fictionnel, nous pouvons poser une question simpliste : pourquoi est-ce que, durant un laps de temps, la quantité totale Q_{tot} de méfaits est parfaitement compensée par la masse punitive P_{tot} qui est purgée chaque année ? La réponse est assez simple à fournir : parce que c'est une tautologie, une vérité analytique. Dans ce monde, par définition, toute faute est compensée justement. L'ensemble de tous les dommages mondiaux (psychologiques, sociaux, à la propriété, etc.) *a la même valeur d'échange* que l'ensemble de toute la punition qui est purgée : $Q_{\text{tot}} = P_{\text{tot}}$.

Comment expliquer cette égalité ? Le substantialiste pourra postuler que les juges causent autant de tort aux auteurs de trouble que ces derniers en ont causé. Or, d'une part, c'est empiriquement douteux (notre monde actuel ne fonctionne vraisemblablement plus ainsi, par exemple). D'autre part, comme Simmel l'a montré plus tôt pour la monnaie, *l'approche substantialiste n'est pas métaphysiquement nécessaire*. On pourrait le démontrer abstraitement à nouveau mais, pour aiguiller notre intuition, un exemple vaut mieux que dix démonstrations.

9.2. Le shifgrethor : la punition comme symbole abstrait

Imaginons la planète Géthen. Sur cette dernière, les tribunaux n'ont qu'un seul outil à disposition : distribuer du *shifgrethor* aux coupables¹⁴⁶. Un violeur, par exemple, est couramment puni de 10 shifgrethors, ajoutés dans son compte judiciaire. Surprenamment, avoir un faible ou grand shifgrethor ne cause pas de tort substantiel : ni spoliation, ni ostracisme, aucune conséquence indirecte. Malgré tout, chaque habitant de Géthen prend son shifgrethor très au sérieux, avec le même zèle qu'un financier qui scrute le cours de la bourse et la même aversion qu'un croyant terrifié par le jugement dernier. Par exemple, quand un géthennien trouve qu'un versement de shifgrethor est nettement trop sévère ou trop clément, il ressent une rage sourde et indignée devant ce déni de justice. De manière générale, les géthenniens sont persuadés qu'il faut classer le shifgrethor parmi les émotions négatives. Si vous refusez de croire que cela soit possible, vous avez le droit d'imaginer que leur conviction est causée par un lavage de cerveau, qu'elle est encodée génétiquement, etc.—peu importe ! L'important c'est que sur Géthen le shifgrethor n'est qu'un chiffre qui *communique, symbolise et quantifie* une injustice, et qu'il n'y a rien besoin de plus pour rendre la justice rétributive. Autrement dit, *le shifgrethor est l'aboutissement complet des théories communicatives du mérite*.

¹⁴⁶ Géthen est le nom de la planète du roman *La main gauche de la nuit* de Ursula K. Le Guin. Je vole aussi le mot *shifgrethor* à cette histoire. Par contre, mon shifgrethor n'a presque rien à voir avec celui de LeGuin : dans le roman, pour des raisons narratives, le shifgrethor est une notion d'honneur et de justice informelle dont le sens est difficile d'accès, et qui n'est (certainement) pas utilisée dans les tribunaux. Il veut dire quelque chose comme « l'inverse du rang social, la capacité à maintenir l'égalité dans n'importe quelle relation, tout en respectant l'autre en tant que personne »—en tout cas d'après l'analyse de George Slusser, spécialiste de science-fiction, lequel est cité dans WHITE Donna, « Dancing with Dragons: Ursula K. Le Guin and the Critics », *Faculty Book Publications*, 1999. pp. 56-60.

LE GUIN Ursula K., *The Left Hand of Darkness*, New York : Ace Books, 1969, 286 p.

Par ailleurs, il faut souligner que l'expérience de la douleur (la douleur comme qualia, pas le problème physique sous-jacent) est, elle aussi, une pure communication du même type que le shifgrethor.

Mais alors, d'où vient la valeur du shifgrethor ? Pourquoi est-ce qu'un viol se punit par dix shifgrethors plutôt que par zéro ou un million ? L'approche simmelienne l'explique très facilement. Je sais qu'un viol représente un millionième (par exemple) des torts causés Q_{tot} sur Géthen ; et que 10 shifgrethors représentent, eux aussi, un millionième de la masse punitive exécutée P_{tot} (toujours sur le même laps de temps). Tant que cette proportionnalité est conservée, le jugement sera juste. À l'inverse, si je gracie ou si je punis arbitrairement, donc si je détruis ou si je crée arbitrairement du shifgrethor dans le système, la proportionnalité est modifiée—et cela suffit à rendre mon jugement injuste. En effet, le shifgrethor est une unité de mesure : si cette unité de mesure est instable, elle ne communique plus grand-chose de clair. Or, si punir c'est communiquer et si mal punir est injuste, alors mal communiquer est injuste.

Cette doctrine a trois conséquences, les trois conséquences promises pour résoudre le problème des siamois. Premièrement, la communauté des condamnés purge *exactement* la peine qu'elle mérite car, par définition, $Q_{\text{tot}} = P_{\text{tot}}$. Deuxièmement, les siamois sont exactement deux personnes car ils sont deux individus juridiques—pour appliquer le problème, on suppose qu'il y a une sorte de bug informatique qui verse le shifgrethor à double chez l'un et l'autre jumeau. Troisièmement, sur Géthen la punition est parfaitement élastique, là aussi par définition car $Q_{\text{tot}} = P_{\text{tot}}$. Qui plus est, ce système est moralement désirable. En effet, le shifgrethor ne fait qu'appliquer les préceptes des théories communicatives du mérite : celui qui trouve ces dernières moralement pertinentes doit nécessairement aussi apprécier le shifgrethor géthennien.

Mais ce qui est vrai pour Géthen ne l'est pas forcément pour la Terre. Les leçons du shifgrethor sont-elles applicables au problème bien réel et concret des jumeaux siamois ? C'est-à-dire : théoriquement, peut-on construire un *continuum* entre un monde du talion (punition purement substantialiste) et un monde de shifgrethor (punition purement symbolique) ? Et empiriquement, peut-on placer notre système judiciaire bien réel quelque part entre ces deux extrêmes ? Commençons par l'argument empirique.

9.3. Histoire de la rétribution en Suisse

Nous avons des raisons de penser qu'il existe bel et bien un mouvement historique qui va d'un monde du talion à un monde de shifgrethor. Pour l'illustrer, prenons l'histoire des peines et de la criminalité sur le territoire suisse, de l'époque médiévale à l'époque moderne. Le plus simple consiste à commencer cette histoire avec la fin du droit romain, qui laisse sa place aux "lois barbares" (VI^e-IX^e siècles), lesquelles arbitrent principalement les conflits entre clans, plutôt qu'entre personnes. En théorie, ces conflits peuvent se résoudre en consultant « des catalogues d'amendes subtilement graduées »¹⁴⁷—une forme de justice finalement pas si lointaine de celle d'un monde de shifgrethor. En pratique, cependant, beaucoup de clans résolvent leurs conflits par la *guerre privée*, une alternative au tribunal qui est courante en Suisse tout au long du Moyen-Âge (malgré les efforts du clergé, puis des gouvernements, pour l'endiguer)¹⁴⁸.

Pour arbitrer les conflits entre particuliers, en revanche, le Haut Moyen-Âge privilégie « le principe de la loi du talion »¹⁴⁹, à base d'une grande variété de châtiments corporels douloureux, de mutilations variées et de peines de mort créatives, qui permettent ainsi de couvrir toute la gravité des crimes possibles. Par ailleurs, ces peines dépendent du rang social des personnes impliquées et diffèrent d'une région à l'autre (ce qui constitue autant de petites masses punitives clairsemées). Par contraste, les peines "quantifiables" comme l'amende, la prison ou le bannissement, ont une utilité très circonscrite. Au Moyen-Âge, l'amende est reversée à la victime, spécifiquement pour laver son honneur (en cas d'injure, de coups et blessures)¹⁵⁰. Le bannissement, lui, est une perte d'honneur qui sert généralement « de moyen de pression pour obtenir le paiement de dettes ou d'amendes »¹⁵¹. Quant à la prison, au vu de sa brutalité, elle est considérée comme un type de châtiment corporel¹⁵².

¹⁴⁷ GAUCH Caroline, « Peines », PIGUET Florence (trad.), *Dictionnaire historique de la Suisse (HLS DHS DSS LIS)*, 26.11.2013.

¹⁴⁸ STADLER Hans, « Guerre privée », MARTIN Pierre-G. (trad.), *Dictionnaire historique de la Suisse (HLS DHS DSS LIS)*, 23.10.2006.

¹⁴⁹ GAUCH Caroline, « Peines », *art. cit.*, 2013.

¹⁵⁰ GAUCH Caroline, « Peines », *art. cit.*, 2013.

¹⁵¹ DUBLER Anne-Marie, « Bannissement », SANDOZ DUTOIT Anne (trad.), *Dictionnaire historique de la Suisse (HLS DHS DSS LIS)*, 11.11.2008.

¹⁵² GAUCH Caroline, « Peines », *art. cit.*, 2013.

Avec l'émergence de l'État moderne (xvi^e-xviii^e siècles), les punitions deviennent plus communicationnelles et plus quantifiables. Le tribunal devient public, ce qui entraîne « une extension et une aggravation des peines infamantes comme le pilori, la tournoire ou le carcan »¹⁵³. D'un autre côté, les punitions quantifiables en unité de temps ou de monnaie gagnent en importance. Les travaux forcés d'intérêt général (appelés "sonnettes") tendent à remplacer les punitions corporelles¹⁵⁴ ; les condamnations à mort diminuent nettement, parfois substituées par le ramage forcé sur les galères¹⁵⁵ ; de même, on observe que les amendes—désormais reversées à l'État—servent à punir une plus en plus grande variété de petits larcins¹⁵⁶. Par contraste, c'est aussi à cette époque que se développent les fameuses chasses aux sorcières, qui touchent toute l'Europe du xv^e au xvii^e siècle, mais qui sont inventées en Suisse et y restent populaires très longtemps, surtout dans les « territoires où l'exercice de la haute justice était très morcelé » et durant les périodes de disette, qui exacerbent les conflits de voisinage¹⁵⁷.

Il faut attendre le milieu du xix^e siècle pour que la plupart des cantons fassent « de l'emprisonnement la sanction principale »¹⁵⁸. Les châtiments corporels sont définitivement abolis en 1874, et la peine de mort en 1942. Enfin, dans le code pénal du début du xxi^e siècle, le jour-amende, le travail forcé, la prison et l'expulsion du territoire se comptent tous dans la même unité : le temps. Ces masses punitives sont de plus en plus unifiées dans un même système comptable national, voire international avec le renforcement continu de la *Cour européenne des droits de l'homme* depuis 1959. L'exception majeure à ce système comptable, c'est le contraste entre la peine *ferme* et la peine *avec sursis*, qui gagne en importance. Mais d'un autre côté, cette dernière mutation nous rapproche toujours plus d'un monde de shifgrethor. En effet, en dernière analyse, le shifgrethor se réduit à des peines en éternel sursis.

¹⁵³ BURGHARTZ Susanna et HELMES-CONZETT Cornelius, « Criminalité », GAILLARD Ursula (trad), *Dictionnaire historique de la Suisse (HLS DHS DSS LIS)*, 23.10.2006.

¹⁵⁴ GAUCH Caroline, « Peines », *art. cit.*, 2013.

¹⁵⁵ *Idem.*

BURGHARTZ Susanna et HELMES-CONZETT Cornelius, « Criminalité », *art. cit.*, 2006.

¹⁵⁶ GAUCH Caroline, « Peines », *art. cit.*, 2013.

BURGHARTZ Susanna et HELMES-CONZETT Cornelius, « Criminalité », *art. cit.*, 2006.

¹⁵⁷ PFISTER Ulrich, « Sorcellerie », GAILLARD Ursula (trad.), *Dictionnaire historique de la Suisse (HLS DHS DSS LIS)*, 16.10.2014.

¹⁵⁸ *Idem.*

Ainsi, aujourd'hui, le commun des mortels a une conception plutôt relativiste des punitions. Pour soutenir cette idée, la philosophe Alice Ristroph cite des résultats consensuels en psychologie sociale¹⁵⁹. Par exemple, si on demande aux gens combien d'années de détention mérite un voleur de voiture (analyse substantialiste), les réponses varient radicalement. En revanche, si on leur demande « qui mérite une plus grande peine de prison, un voleur de voiture ou un meurtrier » (analyse relativiste), il y a plus ou moins consensus. L'analyse empirique de Ristroph est un bon indice que notre monde contemporain se trouve quelque part entre un monde du talion et un monde de shifgrethor. Ce mouvement historique n'est pas une coïncidence. En effet, nous avons les moyens de montrer qu'il s'agit d'une *transition évolutionnaire d'individualité* de la punition.

Petite parenthèse : Ristroph elle-même croit, par erreur, que ce caractère relativiste de la punition affaiblit les théories rétributives du mérite. Or, comme nous venons de le voir, les théories communicatives du mérite peuvent résoudre le problème des siamois *justement* parce que, dans leur version la plus épurée, elles aboutissent au shifgrethor de la planète Géthen, un monde où la punition a une valeur *relative* à l'ensemble des punitions et aucune valeur substantielle¹⁶⁰ !

9.4. La transition évolutionnaire d'individualité de la rétribution

Comme plus tôt à propos de la monnaie, cette théorie du shifgrethor a des conséquences ontologiques intéressantes. On peut analyser ces dernières avec le formalisme de Georg Simmel et celui d'Ellen Clarke, qui s'adaptent parfaitement à cet objet social. Lorsqu'on s'intéresse aux origines de la valeur des punitions, on réalise vite qu'il se joue une *transition d'individualité* : dans un monde du talion, les punitions sont des atomes ; alors que dans un monde de shifgrethor, la masse punitive est une matière unitaire, dont chaque peine est une portion (fig. x).

¹⁵⁹ RISTROPH Alice, « Desert, Democracy, and Sentencing Reform Criminal Law », *Journal of Criminal Law and Criminology*, vol. 96, n° 4, 2005, p. 1294.

¹⁶⁰ Si on veut être précis, le problème des théories substantialistes, c'est qu'elles ne permettent pas d'interdire formellement la *justice stochastique*, un système qui organise des loteries rationnellement équivalentes à la juste punition (section 5.2). Pour refuser cette équivalence, il faut que la punition soit un acte de communication symbolique, que la valeur de la punition ne soit pas substantielle. En bref, l'analyse simmelienne de la punition est une force, pas une faiblesse.

Ainsi, selon cette analyse, un “jour de prison” est une étape de cette transition d’individualité, qui fonctionne en partie comme une punition atomique, et en partie comme une portion d’une masse punitive. Comme on l’a vu plus tôt à propos de la théorie simmélienne de la monnaie, la coexistence de doctrines aux ontologies antagonistes (atomes *versus* matière unitaire) n’est pas forcément répréhensible sur le plan philosophique. Dans un monde intermédiaire entre talion et shifgrethor, on s’attendrait même plutôt à y voir adoptée une ontologie intermédiaire de la punition.

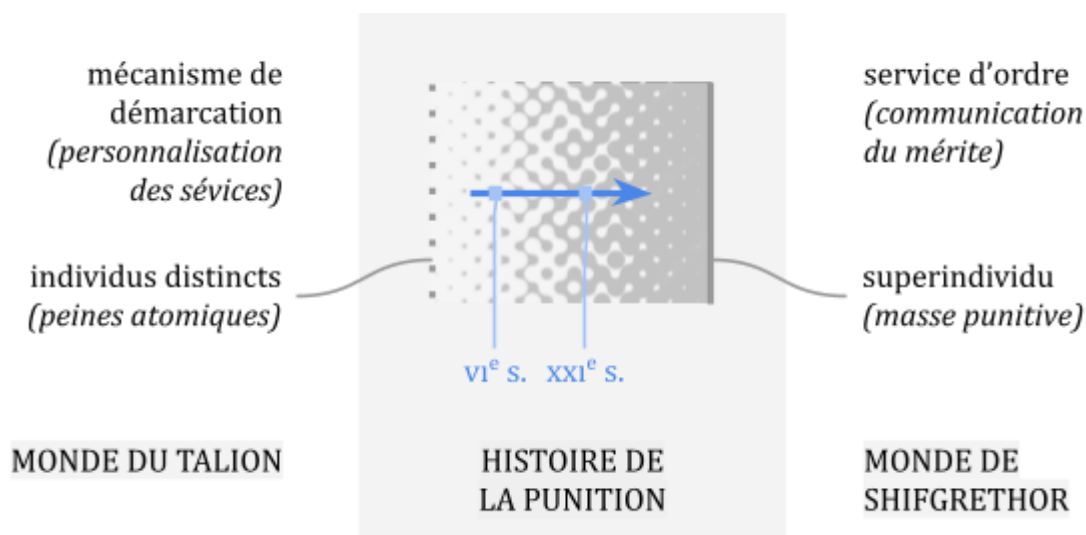


fig. x : Schéma de la transition d’individualité de la punition.

Par ailleurs, il s’agit sûrement d’une transition *évolutionnaire*. Quel est son moteur ? Difficile à dire. En guise de première ébauche, inspirons-nous de l’analyse de Simmel à propos de l’évolution de la monnaie¹⁶¹, car il est plus que probable que la punition, elle aussi, représente des forces sociales abstraites. Est-ce qu’un monde de shifgrethor nous rend plus libre, dans le sens qu’on peut changer plus facilement de partenaire de coopération ? Oui, probablement. En effet, un monde de shifgrethor est l’apogée des théories communicatives du mérite : non seulement chacun sait exactement ce qu’il mérite, mais cette information est publique, accessible à n’importe quel inconnu. La valeur morale de chaque partenaire est réduite à son *shifgrethor*—tout comme la valeur économique de chaque service est encodée dans le prix. En somme, ici aussi, plus l’être humain a d’occasions d’exercer sa *liberté d’indifférence*, plus il est libre.

¹⁶¹ SIMMEL Georg, *The Philosophy of Money*, BOTTOMORE Tom et FRISBY David (trad.), London : Routledge, 2011 (1900), p. 301.

Plus tôt, avec la théorie simmelienne de la monnaie, nous avons vu que pour prouver formellement cette transition d'individualité, il fallait s'en remettre à l'analyse de l'individualité biologique de Clarke. Nous avons démontré que la monnaie change de fonction (réserve de valeur → unité de compte) et que chez Clarke, ces deux fonctions monétaires ont aussi une fonction individualisante (mécanisme de démarcation → service d'ordre) et que cela suffit à démontrer la transition d'individualité. Développons la même idée pour la punition.

Dans un monde du talion avec punition substantielle, la fonction principale d'une punition est d'être objectivement juste—et donc une punition d'une valeur stable. Si ma peine de dix coups de fouet est juste, elle le restera, même si demain le tribunal décide de gracier tous les futurs accusés ou de faire de la pendaison la peine minimale. Cela signifie que chaque acte de punition est jugé séparément ; si un acte punitif est recopié ou abandonné par les autres juges, c'est uniquement pour ses qualités propres. Autrement dit, cette première fonction de la punition favorise la sélection culturelle de toutes les punitions contre toutes : c'est donc un *mécanisme de démarcation*.

À l'inverse, dans un monde de shifgrethor, la fonction principale de la punition est d'être une "unité de compte" du mérite relatif. Cette fonction-là force une punition à changer de valeur lorsque la gravité relative de n'importe quelle autre punition change. C'est justement ce phénomène de conservation de la punition collective que nous utilisons pour résoudre le problème des siamois. Si un acte punitif exprimé en shifgrethor est recopié ou abandonné par les autres juges, ils recopient ou abandonnent *toutes* les autres punitions du même geste. Autrement dit, cette seconde fonction de la punition empêche toute sélection culturelle entre différentes peines : c'est donc un *service d'ordre*.

Si vous êtes attentifs, vous aurez peut-être remarqué une circularité vicieuse dans ce raisonnement. En effet, le monde du talion est parfaitement stable : il ne contient aucune aspérité qui pourrait générer une nouvelle fonction à la punition—c'est d'ailleurs la fonction d'un mécanisme de démarcation d'empêcher cela. Mais Ellen Clarke a déjà discuté de cette objection (section 4.2) ! Pour la résoudre, on suppose qu'un phénomène dû au hasard (une innovation culturelle) génère soudain un très léger service d'ordre. Par exemple, si on décide de toujours utiliser l'épée pour exécuter les nobles (quel qu'ait été leur *modus operandi*) alors cela rend leurs crimes et leurs mérites

commensurables, directement comparables : si la valeur de la punition change pour l'un de ces condamnés noble, elle bouge aussi pour tous les autres. À partir de là, la sélection culturelle peut agir sur les deux niveaux d'individualité et renforcer, de génération en génération, ce service d'ordre au détriment du mécanisme de démarcation.

Ce changement graduel de fonction de la punition décrit donc une *transition évolutionnaire d'individualité* au sens exact de Clarke. Voilà en somme la réponse que nous pouvons apporter au « mystère métaphysique » de la mesure du mérite dont parlait Alice Ristroph¹⁶². Par ailleurs, comme nous l'avons dit, cela permet de résoudre le problème des siamois. La seule manière de refuser la *loi de conservation de la punition*, c'est de vouloir vivre dans un monde de talion, et refuser tout système juridique mixte qui traite également la punition comme une peine symbolique. Or, à notre sens, un pur monde du talion ne peut pas exister—la transition vers un monde de shifgrethor est toujours nécessairement entamée, même légèrement. Même comme fiction philosophique, le pur substantialisme a le défaut d'autoriser la justice stochastique (section 5.2). Êtes-vous prêts à organiser des lotos dans les tribunaux ? Il s'agit probablement du prix à payer pour s'opposer à la loi de conservation.

Ainsi, puisque toute punition est composite, à la fois substantialiste et communicative, le rétributivisme nous ordonne de condamner les deux siamois à la peine moyenne {*moy*}, le mérite d'autrui égal par ailleurs.

¹⁶² « Exactly how much punishment an offender deserves is something of a metaphysical mystery, or so it has appeared to be in the past. »

RISTROPH Alice, « Desert, Democracy, and Sentencing Reform », *art. cit.*, p. 1293.

CONCLUSION

Solution au problème des siamois

Ces longs raisonnements successifs donnent radicalement raison à Rawls (tel que paraphrasé par Bart) : « les frontières entre individus ont plus qu'une importance indirecte et définissent en fait l'objet même de la préoccupation morale »¹⁶³ (section 3.3). Le problème des siamois est intéressant pour un philosophe, car il empêche les théories de la punition d'appliquer une définition naïve de l'individualité. Tout au long de ce travail, nous avons contrasté cette définition naïve avec une définition tirée de la philosophie de la biologie, qui s'appuie sur le concept de *transition évolutive* d'individualité, un phénomène bien établi (section 4.1). La grande force de ce concept, c'est qu'il est fonctionnaliste, ce qui nous permet de l'appliquer au-delà du monde biologique, comme à des sociétés et des institutions humaines, ainsi qu'à des objets culturels fonctionnels (monnaie, punition, etc.).

La seule famille de théories de la punition qui bute sur le problème des siamois, c'est la *loi du talion* et ses dérivées (section 5.1). D'après cette loi, une punition est un troc entre deux torts équivalents, celui que le coupable a causé à la victime, et celui que la justice impose au coupable. Dans le cas des siamois, la juge n'a presque aucune restriction : punir le coupable comme un pur innocent, ou l'innocent comme un pur coupable revient au même [*min;max*]. Les jumeaux siamois poussent la loi du talion au paradoxe ou, du moins, à organiser une loterie (justice stochastique) (section 5.2). Or, la raison profonde de cet échec s'explique parce que cette théorie affirme à la fois l'individualité des personnes (individualisme) et l'individualité des punitions (substantialisme) (section 9.4).

Pour passer outre le problème des siamois, une autre forme de théorie rétributive, les *théories communicatives du mérite*, s'appuient sur l'expérience de pensée d'un *monde de shifgrethor* (section 9.2). Dans cette expérience de pensée, la valeur de la punition est

¹⁶³ « As distinct systems of ends, persons cannot simply be cashed out as just so many containers for valuable experiences, since the boundaries between individuals are of more than merely derivative importance and indeed define the very object of moral concern. »

SCHULTZ Bart, « Persons, Selves, and Utilitarianism », *art. cit.*, p. 72.

nominale/symbolique, plutôt que substantialiste (tout comme la monnaie par rapport au troc). Ce monde nie l'individualité des punitions, qui ne sont qu'une proportion de toute la masse punitive exécutée par la justice (section 9.4). Plus exactement, d'après notre analyse, pour maintenir la stabilité 'monétaire' de la punition, il ne faut ni en détruire ni en ajouter par rapport à ce que la communauté (et donc les siamois) mérite au total (section 6.2). Cela nous permet de construire une *loi de conservation de la punition* (section 6.1). Pour ne pas léser la communauté des autres condamnés, pour n'importe quel crime, la peine habituelle du coupable est répartie entre les deux siamois—ce qui revient à faire une moyenne entre les mérites des deux siamois (section 6.2).

Pour passer outre ce même problème des siamois, l'utilitarisme—une toute autre famille de théories de la punition—s'appuie sur l'expérience de pensée de l'*Observatrice impartiale* (section 3.2). Cela amène à nier l'individualité des siamois, qui ne sont qu'une portion d'un monumental superindividu. En réalité, le siamois criminel est une simple partie de la communauté qui s'est attaquée elle-même (section 4.4). D'après notre analyse, le *standard de preuve*, qui compare la préférence des institutions entre deux types d'erreur judiciaire (libérer un coupable vs condamner un innocent) permet de calculer assez précisément la punition encourue par les siamois. Pour un crime grave, nous avons estimé que la peine habituelle du coupable serait divisée par un facteur compris entre trois et quinze (section 2.2).

Utilitarisme, rétributivisme et loi du talion ne se contredisent pas, mais se complètent. En effet, la *transition évolutionnaire d'individualité* admet la coexistence des individus et de leur superindividu. Plus exactement, elle défend que cette transition graduelle se quantifie. Par exemple, si on mesure que plus de la moitié de la *transition d'individualité humaine* a eu lieu, alors la punition se répartit à 40% chez la personne (rétributivisme) et à 60% chez la communauté (utilitarisme). Puis, si on mesure que la moitié de la *transition d'individualité de la punition* a eu lieu, alors la punition de la personne se répartit à 20% dans une punition isolée (loi du talion) et à 20% dans la masse punitive (théorie communicative du mérite). En pratique, cependant, une punition ne peut pas être décomposée dans sa composante substantielle et symbolique. Par chance, les préférences de la théorie communicative du mérite, plus spécifiques, s'imposent sans que cela ne contredise la loi du talion. En revanche, une punition quantifiable (comme

une peine de prison) peut être décomposée selon ses victimes (la personne et la communauté).

Histoire de faire les choses proprement, donnons-nous la peine de mathématiser le calcul de la punition. Soient p , i et c les punitions méritées, respectivement, par le duo, par le jumeau innocent et par le jumeau coupable ; soit $0 < t < 1$, la mesure de l'avancement de la transition évolutionnaire d'individualité, de la personne vers la communauté ; et soit $(1 : k)$, le standard de preuve défendu par l'utilitarisme. Dans ce cas, on peut construire l'égalité suivante :

$$p = i + \left(\frac{1-t}{2} + \frac{t}{k+1} \right) (c - i)$$

Reprenons l'expérience de pensée d'Anne-Claire que nous avons présentée en introduction. Anne mérite dix ans de prison ; sa sœur en mérite zéro. Nous pouvons asséner avec une certitude assez élevée qu'il serait très injuste que la juge libère Anne ou condamne Claire à dix ans de réclusion. La punition juste tourne plutôt autour de deux à quatre ans de prison. Par exemple, si le mérite de l'innocente est $i = 0$ [année], le mérite du coupable est $c = 10$ [année], et si la transition d'individualité vaut $t = 0.6$ et le standard de preuve est de $k = 5$, alors la juge doit condamner Anne-Claire à une peine de trois ans d'emprisonnement ; en effet, $p = 0 + (0.2 + 0.1) * 10 = 3$.

Limites de cette solution

Bien évidemment, notre réponse ne manque pas de défauts. Notre raisonnement s'appuie sur deux prémisses clés. La première, c'est qu'il faut rendre justice à la victime en entier, dans toute son individualité. Il faut donc rendre justice à la personne *et* à la communauté bafouée par le crime. Il doit être possible d'attaquer cette idée de manière convaincante, soit en niant qu'une communauté ait une quelconque valeur morale directe, soit en niant qu'une personne hors-communauté ait une quelconque valeur morale directe. Cela serait sans doute aussi l'occasion de remettre les deux Johns dans la course¹⁶⁴.

¹⁶⁴ RAWLS John, *A Theory of Justice*, Oxford : Clarendon Press, 1972, xv+607 p.

HARSANYI John C., « Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility », *Journal of Political Economy*, vol. 63, n° 4, University of Chicago Press, 1955, p. 309-321.

La seconde prémisse clé découle d'une observation concrète. Vous aurez remarqué que les théories communicatives du mérite et l'utilitarisme préconisent des punitions différentes. Ici, nous sommes partis du principe qu'elle ne se contredisent pas, mais qu'elles se complètent.. En somme, elles ont chacune leur pré carré. Or, il est possible d'attaquer cette *hypothèse du pré carré*. Il suffit de montrer que les deux théories sont au chevet de la même victime (la communauté) et qu'elles pratiquent les deux une punition à la valeur nominale/symbolique (monde de shifgrethor). Pour arriver à cette conclusion, il faudrait sans doute montrer qu'il n'y a qu'un seul processus évolutionnaire, une seule transition d'individualité pour la punition et pour les humains.

Cet angle d'attaque est très pertinent ! Pour comprendre l'importance du problème, il faut se déplacer un instant à l'échelle de la personne—qu'on la conçoive comme une suite de tranches de vie (section 3.3) ou un amas de molécules (section 3.4). Car à l'échelle de la personne, le shifgrethor est déjà en place ! Il prend la forme de "l'effet que ça fait" d'avoir ou d'aller mal, c'est-à-dire la souffrance comme qualia. En effet, la souffrance n'est pas un mal substantiel : elle communique, symbolise et quantifie un problème interne à l'animal, qui est le vrai mal substantiel. Par exemple, d'après Georg Simmel, sa théorie relativiste s'applique parfaitement à la mesure de la souffrance et du bonheur ; c'est d'ailleurs le premier exemple qu'il utilise pour illustrer et rendre sa théorie monétaire plus intuitive pour le lecteur. Nous avons donc des raisons de croire que *la souffrance sert d'unité de compte au sein de la personne*. Cette hypothèse a trois répercussions très intéressantes.

La première répercussion intéressante, c'est que les débats en philosophie de la punition sont analogues à des débats en philosophie de l'esprit. Pour savoir si les théories communicatives du mérite et l'utilitarisme sont complémentaires ou contradictoires, nous devons en même temps répondre à une des questions les plus épineuses de la philosophie de l'esprit : sous quelles conditions historiques la conscience peut-elle apparaître ? C'est-à-dire, à quoi sert la conscience ?

La deuxième répercussion prend la forme d'une critique de l'utilitarisme. Comme on le sait, l'utilitarisme classique utilise la souffrance comme unité de compte du bien et du mal. Or, cette unité est conçue pour mesurer l'état interne de la personne, et non pas l'état interne de la société (ou plutôt, de l'ensemble des êtres conscients, la noosphère). Par contraste, les descendants de l'utilitarisme de Bentham n'ont eu de cesse de

modifier et de peaufiner cette unité de mesure. Cette famille de théories de la justice finira-t-elle par défendre une unité de mesure de l'état interne de la société, de la noosphère ou de la biosphère ? Cela revient-il à traiter de plus en plus la société, la noosphère ou la biosphère comme un être individuel *et* conscient ?

La troisième répercussion prend la forme d'une critique des théories communicatives du mérite. Comme on le sait, les théories communicatives du mérite ne savent pas exactement à qui s'adresse ladite communication : le criminel, la victime, la communauté ? Pour résoudre le problème des siamois, nous avons avancé que la punition s'adresse notamment à la communauté des condamnés : le mérite se mesure toujours par rapport au mérite collectif. Comme nous l'avons vu, de plus, le mérite est une sorte de mystère métaphysique, dont on ignore ce qu'il mesure exactement. Il se pourrait que cette famille de théorie finisse par définir le mérite comme une unité de mesure de la santé de la société. Cela revient-il à traiter de plus en plus la société comme un être conscient *et* individuel ?

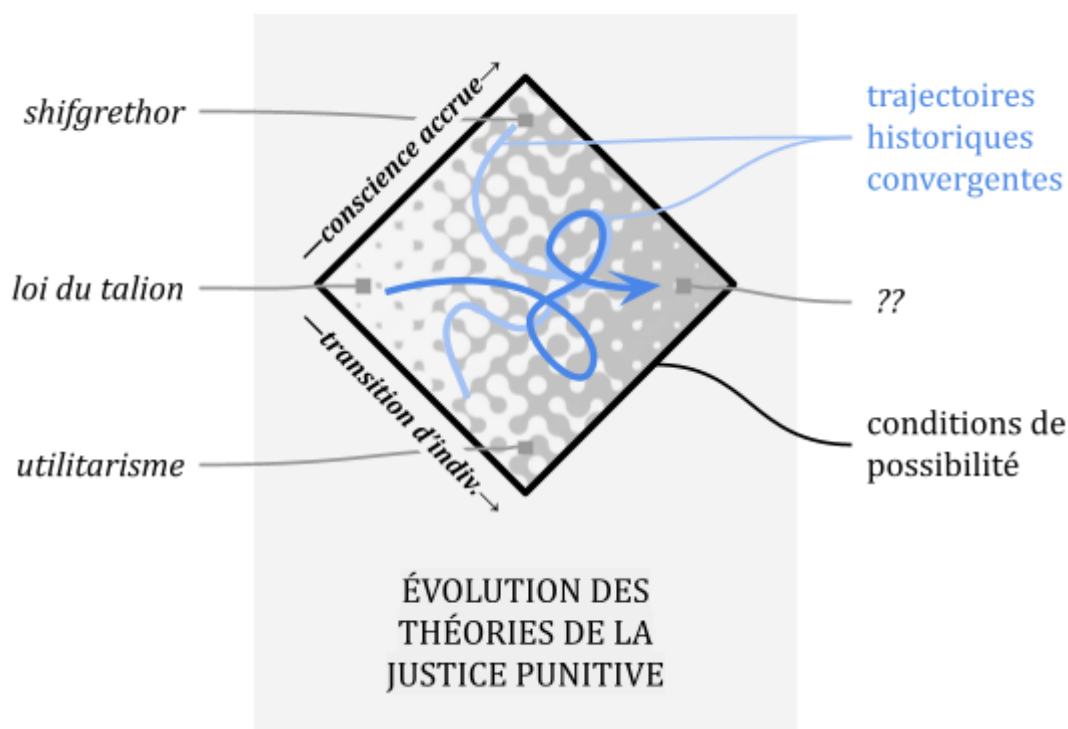


fig. xi : Schéma de la transition d'individualité de la justice punitive.

En résumé, l'hypothèse du pré carré part du principe que le graphe ci-dessus ne tient pas la route (fig. xi). Ce graphe, vous en reconnaissez sûrement la forme. Il ressemble beaucoup à celui que j'ai utilisé pour résumer les désaccords entre théories monétaires,

qui prenait la forme d'un espace des phases avec un axe de vélocité et un axe de conventionnalité de la monnaie (section 7.4, fig. vi ; section 8.2, fig. viii). Un tel graphe implique l'existence de quatre fonctionnements antagonistes de la justice, dont nous pouvons déjà placer les trois premières. À quoi ressemble cette hypothétique quatrième fonction de la punition ? Peut-elle exister ? Tout ce que nous savons pour l'instant, c'est qu'elle doit servir de service d'ordre dans une tout autant hypothétique transition évolutionnaire d'individualité de la justice humaine. Cette perspective, dans tous les cas, mérite d'être explorée.

Solution au problème des boucliers humains

Le problème des siamois dans le droit pénal est, comme nous l'avons vu, analogue au problème des boucliers humains dans le droit humanitaire (section 1.3). Pour résoudre le problème des boucliers humains, nous devons donc étudier les transitions évolutionnaires d'individualité au sein de ce droit. Or, ce dernier est encore très jeune, né à la toute fin du xix^e siècle seulement ; il est trop tôt pour prédire avec confiance son évolution. Peut-être que la dialectique en son sein, entre les agents de bonne foi et les agents perfides, formera une course aux armements interminable qui ne pourra pas se résoudre. Ou peut-être qu'il prendra la voie d'une transition d'individualité, comme les autres systèmes pénaux. Explorons cette possibilité.

La particularité de la guerre, c'est qu'elle oppose non pas des individus, mais des groupes (*esprit de corps*), voire des peuples (*effort de guerre*). C'est en temps de guerre que la métaphore organismique des communautés humaines est la plus pertinente et que le rationnement, les punitions collectives et la chasse aux ennemis intérieurs atteignent leur paroxysme. Ces phénomènes, symptômes d'une "union sacrée", sont en fait tous des exemples de *services d'ordres*. Toutes les personnes, cependant, ne renonce pas à leur individualité au même degré. Les militaires en uniforme et les résistants la sacrifient presque entièrement, leur être entier est dévoué à l'effort de guerre ; par contraste, des réfugiés en fuite à l'étranger ne renoncent à aucune individualité et leur effort de guerre est négligeable. Des travailleurs de l'armement ou des boucliers humains fournissent un effort de guerre qui se situe entre ces deux extrêmes. Ce sont ces multiples degrés de transition d'individualité que cherche à mesurer le droit de la

guerre—il s’agit là de son premier principe, que nous avons déjà rencontré dans le problème des boucliers humains (section 1.1).

Le deuxième principe qu’ajoute le droit humanitaire, c’est l’interdiction du génocide. Pour y parvenir, le droit distingue trois niveaux d’avancement dans cette *transition d’individualité* : un niveau avancé, qui permet de soutenir un effort de guerre ; un niveau normal, qui permet seulement de maintenir un effort de paix ; et un niveau minimal, où plus aucune institution ne subsiste. De ce point de vue, un but de guerre juste, c’est de faire rebrousser chemin à la *transition d’individualité* de l’ennemi jusqu’à un niveau ‘normal’ où tout effort de guerre est impossible. En revanche, il est injuste de poursuivre ce processus jusqu’à l’anéantissement complet de la société ennemie et de toutes ses institutions, jusqu’à ce que l’ennemi ne soit même plus un peuple à proprement parler.

De ce point de vue-là, le droit humanitaire obéit en réalité encore largement à la *loi du talion*. En effet, en cas de crime de guerre, qui est responsable ? Jusqu’à présent d’après ce système de droit, seules des personnes peuvent être responsables. Par conséquent, pour punir un crime de guerre ou un crime contre l’humanité, la cour pénale internationale de la Haye ampute effectivement un collectif de certains individus ciblés (comme on ampute la main du voleur).

Dans un monde de *shifgrethor*, par contraste, le droit humanitaire punit avant tout le collectif *qua* individu. Il vise plutôt à quantifier et réclamer des réparations de guerre pour chaque acte contraire au droit. Comme au temps des *lois barbares* (section 9.3), le droit constitue un catalogue d’amendes savamment graduées. Par ailleurs, ces amendes ne sont pas reversées aux victimes (pensez au cas de bouclier humain, ce qui reviendrait à ce qu’une partie d’un peuple se reverse l’amende à une autre part de lui-même), mais à la communauté des États (l’ONU). Par ailleurs, ces amendes doivent se payer en *shifgrethor*—l’unité de compte de l’honneur—et non pas en monnaie, en commodités, ni en hommes.

Le droit de la guerre peut difficilement se revendiquer de l’*utilitarisme*, ni plus généralement du conséquentialisme. En effet, la guerre n’est rien d’autre qu’un *mécanisme de démarcation* entre groupes le plus total possible¹⁶⁵. En pratique, il est vrai

¹⁶⁵ À ce titre, pour parler de l’homme en temps de guerre, il faudrait hélas souvent parler de *homo genocidensis*. À mon très humble avis, la ‘mystérieuse’ disparition systématique des autres espèces d’hominidés s’explique probablement plus par les tendances génocidaires des groupes d’*homo sapiens*, que par toute autre différence anatomique ou cognitive.

que des nations ennemies continuent à collaborer sur de nombreux aspects, rien que dans la mesure où leurs économies restent en partie intriquées au sein du marché mondial. Ainsi, l'effort de guerre d'un camp contribue toujours légèrement à l'effort de guerre adverse (et vice versa)—mais le but de la guerre reste de minimiser ce phénomène.

Pour toutes ces raisons, nous avons de bonnes raisons de croire que le problème des boucliers humains n'est pas un problème fondamental du droit humanitaire, dans le sens où il peut se modifier lui-même. Nous pensons donc que les thèses défendues par Judith Butler et Vasuki Nesiah sont des signaux annonciateurs des réformes des droits humains qui viendront tôt ou tard¹⁶⁶. On le voit en particulier quand elles critiquent le droit de la guerre et son obsession à distinguer civils et militaires, et quand elles suggèrent qu'il devrait plutôt s'ancrer sur le *droit à l'autodétermination des peuples* contre la colonisation—ce qui n'est qu'une manière détournée de demander au droit de la guerre de réfléchir (d'abord) en termes de peuples plutôt que de personnes. Notre manière d'arriver à cette même conclusion est, clairement, beaucoup moins romantique ; mais elle a l'avantage d'avoir une assise théorique particulièrement générale, la *transition évolutionnaire d'individualité*, qui s'applique à tous les types d'objets vivants et culturels.

*

**

En guise de mot de la fin, revenons à des questions normatives : vaut-il mieux une *transition évolutionnaire d'individualité* complète, partielle, ou pas de transition d'individualité du tout ? Vaut-il mieux une théorie morale qui colle aux transitions d'individualités réelles, ou peut-on se permettre une fiction philosophique en toute connaissance de cause¹⁶⁷ ? Plus largement, si tous les chemins de l'histoire devaient mener à des transitions d'individualité, vaut-il mieux presser le pas ou faire marche arrière ? Ces questions restent ouvertes ; mais il faut bien commencer par le descriptif avant d'oser proposer un principe capable de sauter vers un argument normatif.

¹⁶⁶ BUTLER Judith, « Human shields », *London Review of International Law*, vol. 3, n° 2, 2015, p. 223–243.

NESIAH Vasuki, « Human shields, human heresies », *International Politics Reviews*, vol. 10, n° 1, 2022, pp. 36–42.

¹⁶⁷ La difficulté qu'il y a à donner une réponse claire, c'est qu'elle aura des répercussions sur toutes les situations où des objets résistent à des transitions d'individualité (ce qui inclut par exemple le cancer, deuxième cause de mortalité dans les pays riches). La seule manière d'éviter pour sûr la contradiction et les critères *ad hoc*, c'est de proposer des principes éthiques qui suivent au plus près l'avancement de chaque transition, voire qui les anticipent légèrement. Peut-on faire mieux ?

BIBLIOGRAPHIE

Justice (théorie)

- ALLARD Aurelien, *Le mérite : signification, possibilité et valeur*, Thèse de doctorat, Paris 8, 2019, 319 p., URL : theses.fr/2019PA080012, consulté le 15.12.2024.
- BENTHAM Jeremy, *An Introduction to the Principles of Morals and Legislation*, Londres : W. Pickering, 1780–1789, 2 vol.
- BERMAN Mitchell N., « Punishment and Justification », *Ethics*, vol. 118, n° 2, 2008, p. 258–290, DOI : [10.1086/527424](https://doi.org/10.1086/527424).
- BLACKSTONE William, *Commentaries on the laws of England*, Oxford : Clarendon Press, 1766, 4 vol.
- CUDD Ann et SEENA Eftekhari, « Contractarianism », in ZALTA Edward N. et NODELMAN Uri (dir.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, 2021, URL : plato.stanford.edu/entries/contractarianism/, consulté le 15.12.2024.
- DARDENNE Emilie, « From Jeremy Bentham to Peter Singer », *Revue d'études benthamiennes*, n° 7, 2010, DOI : [10.4000/etudes-benthamiennes.204](https://doi.org/10.4000/etudes-benthamiennes.204).
- DEITCH Brittany, « Retributivist Theories' Conjoined Twins Problems », *University of Cincinnati Law Review*, vol. 87, n° 1, 2018, pp. 139–169.
- DONALDSON Sue et KYMLICKA Will, *Zoopolis: A Political Theory of Animal Rights*, Oxford : Univ. Press, 2014, 329 p.
- DUFF R A, *Punishment, Communication, and Community*, New York : Oxford University Press, 2000, 272 p.
- FEINBERG Joel, « the Expressive Function of Punishment », *The Monist*, vol. 49, n° 3, 1965, pp. 397–423, DOI : [10.5840/monist196549326](https://doi.org/10.5840/monist196549326).
- FRASE Richard S., « Limiting Retributivism: The Consensus Model of Criminal Punishment », in TONRY Michael (éd.), *The Future of Imprisonment in the 21st Century*, Oxford : Oxford University Press, 2003, pp. 83–120, DOI : [10.1093/oso/9780195161632.003.0004](https://doi.org/10.1093/oso/9780195161632.003.0004).

- HARSANYI John C., « Cardinal Utility in Welfare Economics and in the Theory of Risk-taking », *Journal of Political Economy*, vol. 61, n° 5, 1953, pp. 434–435.
- , « Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility », *Journal of Political Economy*, vol. 63, n° 4, 1955, pp. 309–321.
- HAMPTON Jean, « An Expressive Theory of Retribution », in CRAGG Wesley (éd.), *Retributivism and its Critics*, Stuttgart: Franz Steiner Verlag, 1992, pp. 4–[n.c.].
- HART H.L.A. et GARDNER John, *Punishment and Responsibility*, Oxford University Press, 2008, 277 p., DOI : [10.1093/acprof:oso/9780199534777.001.0001](https://doi.org/10.1093/acprof:oso/9780199534777.001.0001).
- HUME David, *Traité de la nature humaine*, tome 3, *La morale*, SALTEL Philippe (trad.), Paris : Flammarion, 1999, 286 p.
- HOSKINS Zachary et DUFF Antony, « Legal Punishment », in ZALTA Edward N. et NODELMAN Uri (dir.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, 2024, URL : plato.stanford.edu/archives/spr2024/entries/legal-punishment/, consulté le 15.12.2024.
- KANT Emmanuel, *Métaphysique des mœurs*, RENAULT Alain (trad.), Paris : Flammarion, 1994 (1795).
- LEE Monica et SILVER Daniel, « Simmel's Law of the Individual and the Ethics of the Relational Self », *Theory, Culture & Society*, vol. 29, n° 7–8, 2012, pp. 124–145, DOI : [10.1177/0263276411435569](https://doi.org/10.1177/0263276411435569).
- LEOPOLD David, « Max Stirner », in ZALTA Edward N. et NODELMAN Uri (dir.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, 2023, URL : plato.stanford.edu/archives/win2023/entries/max-stirner/, consulté le 15.12.2024.
- LIBERTO Hallie, « Species Membership and the Veil of Ignorance: What Principles of Justice would the Representatives of all Animals Choose? », *Utilitas*, vol. 29, n° 3, 2017, pp. 299–320, DOI : [10.1017/S0953820816000315](https://doi.org/10.1017/S0953820816000315).
- MOORE Michael, *Placing Blame: A Theory of the Criminal Law*, Oxford : Clarendon, 1997, pp. 154–158.
- , « The Rationality of Threshold Deontology », in HURD Heidi M. (dir.), *Moral Puzzles and Legal Perplexities: Essays on the Influence of Larry Alexander*, Cambridge : Cambridge University Press, 2018, p. 371–387, DOI : [10.1017/9781108227025.022](https://doi.org/10.1017/9781108227025.022).

- NOZICK Robert, *Anarchy, State, and Utopia*, New York : Basic Books, 1974, xvi+367 p.
- RAWLS John, *A Theory of Justice*, Oxford : Clarendon Press, 1972, xv+607 p.
- RISTROPH Alice, « Desert, Democracy, and Sentencing Reform Criminal Law », *Journal of Criminal Law and Criminology*, vol. 96, n° 4, 2005, pp. 1293–1352.
- RIZZOLLI Matteo et SARACENO Margherita, « Better that ten guilty persons escape: punishment costs explain the standard of evidence », *Public Choice*, vol. 155, n° 3–4, Springer, 2013, pp. 395–411.
- CHRISTOPHER Russel L., « The Prosecutor’s Dilemma: Bargains and Punishments », *Fordham law review*, vol. 72, n° 1, 2003, pp. 93-168.
- SCHULTZ Bart, « Persons, Selves, and Utilitarianism », *Ethics*, vol. 96, n° 4, 1986, pp. 721–745.
- SINNOTT-ARMSTRONG Walter, « Consequentialism », in ZALTA Edward N. et NODELMAN Uri (dir.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, 2023, URL : plato.stanford.edu/archives/win2023/entries/consequentialism/, consulté le 15.12.2024.
- WALEN Alec, « Retributive Justice », in ZALTA Edward N. et NODELMAN Uri (dir.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, 2023, URL : plato.stanford.edu/archives/win2023/entries/justice-retributive/, consulté le 15.12.2024.
- WEYMARK John A., « A reconsideration of the Harsanyi–Sen debate on utilitarianism », in ROEMER John E. et ELSTER Jon (dir.), *Interpersonal Comparisons of Well-Being*, Cambridge : Cambridge University Press, 1991, pp. 255–320, DOI : [10.1017/CB09781139172387.009](https://doi.org/10.1017/CB09781139172387.009).

Justice (appliquée)

- BARGU Banu, « Human shields », *Contemporary Political Theory*, vol. 12, n° 4, 2013, pp. 277–295, DOI : [10.1057/cpt.2013.1](https://doi.org/10.1057/cpt.2013.1).
- BONNEAU Daniel D. et HALL Joshua C., « The impact of new prison construction on the likelihood of incarceration », *The American Journal of Economics and Sociology*, vol. 83, n° 1, 2024, pp. 199–222, DOI : [10.1111/ajes.12529](https://doi.org/10.1111/ajes.12529).

- BURGHARTZ, Susanna; HELMES-CONZETT Cornélius, « Criminalité », GAILLARD Ursula (trad.), *Dictionnaire historique de la Suisse (HLS DHS DSS LIS)*, 26.08.2022, URL : hls-dhs-dss.ch/fr/articles/016556/2022-08-26/, consulté le 14.12.2024.
- BUTLER Judith, « Human shields », *London Review of International Law*, vol. 3, n° 2, 2015, pp. 223–243, DOI : [10.1093/lril/lrv011](https://doi.org/10.1093/lril/lrv011).
- CARPENTER Dick M., SWEETLAND Kyle et McDONALD Jennifer, *The price of taxation by citation: Case studies of three Georgia cities that rely heavily on fines and fees*, Arlington, Va.: Institute for Justice, 2019, 56 p., URL : arks.princeton.edu/ark:/88435/dsp016395wb23h, consulté le 15.12.2024.
- DAVIS Colleen, « Complicity, Crime and Conjoined Twins », *Alternative Law Journal*, vol. 42, n° 1, 2017, pp. 18–23.
- DUBLER Anne-Marie, « Bannissement », SANDOZ DUTOIT Anne (trad.), *Dictionnaire historique de la Suisse (HLS DHS DSS LIS)*, 11.11.2008, URL : hls-dhs-dss.ch/fr/articles/027301/2008-11-11/, consulté le 15.12.2024.
- GAES Gerald G., « The Effects of Overcrowding in Prison », *Crime and Justice*, vol. 6, 1985, pp. 95–146.
- GAUCH Caroline, « Peines », PIGUET Florence (trad.), *Dictionnaire historique de la Suisse (HLS DHS DSS LIS)*, 26.11.2013, URL : hls-dhs-dss.ch/articles/048699/2013-11-26/, consulté le 15.12.2024.
- GIVELBER Daniel, « Lost Innocence: Speculation and Data about the Acquitted », *American Criminal Law Review*, vol. 42, n° 4, Fall 2005, pp. 1167–1200.
- GORDON Neve et PERUGINI Nicola, « The politics of human shielding: On the resignification of space and the constitution of civilians as shields in liberal wars », *Environment and Planning D: Society and Space*, vol. 34, n° 1, 2016, pp. 168–187.
- GROSS Samuel R. et O'BRIEN Barbara, « Frequency and Predictors of False Conviction: Why We Know So Little, and New Data on Capital Cases », *Journal of Empirical Legal Studies*, vol. 5, n° 4, 2008, pp. 927–962, DOI : [10.1111/j.1740-1461.2008.00146.x](https://doi.org/10.1111/j.1740-1461.2008.00146.x).
- GILDEA Robert, « Resistance, Reprisals and Community in Occupied France », *Transactions of the Royal Historical Society*, vol. 13, 2003, pp. 163–185.
- LIEM Marieke *et alii*, « Homicide clearance in Western Europe », *European Journal of Criminology*, vol. 16, n° 1, SAGE Publications, 2019, pp. 81–101, DOI : [10.1177/1477370818764840](https://doi.org/10.1177/1477370818764840).

NESIAH Vasuki, « Human shields, human heresies », *International Politics Reviews*, vol. 10, n° 1, 2022, pp. 36–42, DOI : [10.1057/s41312-022-00138-0](https://doi.org/10.1057/s41312-022-00138-0).

PFISTER Ulrich, « Sorcellerie », GAILLARD Ursula (trad.), *Dictionnaire historique de la Suisse (HLS DHS DSS LIS)*, 16.10.2014, URL : hls-dhs-dss.ch/fr/articles/011450/2014-10-16/, consulté le 15.12.2024.

SCHMITT Michael N., « Human shields in international humanitarian law », *The Columbia journal of transnational law*, vol. 47, n° 2, 2009, pp. 292–338, DOI : [10.1163/9789047427049_003](https://doi.org/10.1163/9789047427049_003).

STADLER Hans, « Guerre privée », MARTIN Pierre-G. (trad.), *Dictionnaire historique de la Suisse (HLS DHS DSS LIS)*, 23.10.2006, URL : hls-dhs-dss.ch/articles/008606/2006-10-23/, consulté le 15.12.2024.

Biologie

BAILLON Ludovic et BASLER Konrad, « Reflections on cell competition », *Seminars in Cell & Developmental Biology*, vol. 32, 2014, pp. 137–144, DOI : [10.1016/j.semcdb.2014.04.034](https://doi.org/10.1016/j.semcdb.2014.04.034).

BEER Randall D., « Autopoiesis and Cognition in the Game of Life », *Artificial Life*, vol. 10, n° 3, 2004, pp. 309–326, DOI : [10.1162/1064546041255539](https://doi.org/10.1162/1064546041255539).

CLARKE Ellen, « The Multiple Realizability of Biological Individuals », *The Journal of Philosophy*, vol. 110, n° 8, 2013, pp. 413–435, DOI : [10.5840/jphil2013110817](https://doi.org/10.5840/jphil2013110817).

———, « A levels-of-selection approach to evolutionary individuality », *Biology & Philosophy*, vol. 31, n° 6, 2016, pp. 893–911, DOI : [10.1007/s10539-016-9540-4](https://doi.org/10.1007/s10539-016-9540-4).

DAWKINS Richard, *The Selfish Gene*, Oxford University Press, 2016, 497 p.

DENNETT Daniel C., *Freedom Evolves*, London : Penguin Books, 2004, xiii+347 p.

HAMILTON W. D., « The genetical evolution of social behaviour. I », *Journal of Theoretical Biology*, vol. 7, n° 1, 1964, pp. 1–16, DOI : [10.1016/0022-5193\(64\)90038-4](https://doi.org/10.1016/0022-5193(64)90038-4).

KOZIELSKA M. *et alii*, « Segregation distortion and the evolution of sex-determining mechanisms », *Heredity*, vol. 104, n° 1, 2010, pp. 100–112, DOI : [10.1038/hdy.2009.104](https://doi.org/10.1038/hdy.2009.104).

LEVINE Donald N., « The Organism Metaphor in Sociology », *Social Research*, vol. 62, n° 2, The New School, 1995, pp. 239–265.

NELSON Paul et MASEL Joanna, « Intercellular competition and the inevitability of multicellular aging », *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, n° 49, 2017, pp. 12982–12987, DOI : [10.1073/pnas.1618854114](https://doi.org/10.1073/pnas.1618854114).

OKASHA Samir, « Social Justice, Genomic Justice and the Veil of Ignorance: Harsanyi Meets Mendel », *Economics and Philosophy*, vol. 28, n° 1, 2012, pp. 43–71, DOI : [10.1017/S0266267112000119](https://doi.org/10.1017/S0266267112000119).

SOBER Elliott et WILSON David Sloan, *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Cambridge, Mass. : Harvard Univ. Press, 2003, 394 p.

WEST Stuart A. *et alii*, « Major evolutionary transitions in individuality », *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, n° 33, National Academy of Sciences, 2015, pp. 10112–10119, DOI : [10.1073/pnas.1421402112](https://doi.org/10.1073/pnas.1421402112).

Monnaie et économie

BADAIRE Quentin, « Gilles Deleuze face aux hétérodoxies monétaires régulationnistes, post-keynésiennes et marxistes : jalons pour un dialogue transdisciplinaire sur la monnaie et la dette », 32 p., in *Institutionnalismes monétaires francophones : bilan, perspectives et regards internationaux*, Lyon (France), 1–3 Juin 2016, tiré de https://imf2016.sciencesconf.org/110145/Texte_pour_colloque [Institutionnalisme monetaires Badaire Quentin .pdf](#) (consulté le 01.01.2025).

FEENSTRA Robert C. et TAYLOR Alan M., *International macroeconomics*, Londres : Macmillan, 2014 (2007), 867 p.

GRAEBER David, *Debt : The First 5000 Years*, New York : Melville House Publishing, 2011, 624 p.

HELMER Étienne, « De quoi la monnaie est-elle la mesure ? Aristote, Diogène et Platon ou la convention monétaire dans l'horizon du politique », *Cahiers d'économie Politique*, vol. n° 72, no 1, 2017, pp. 7–26.

HUME David, « De la monnaie », « De l'intérêt », « De la balance commerciale », in *Discours politiques*, Édimbourg : Kincaid & Donaldson, 1752, pp. 281–307, tiré de <https://davidhume.org/texts/pld/> (consulté le 01.01.2025).

- JAMES Aaron, « Kant, Innes, and the Copernican Turn in Monetary Theory », in TINGUELY Joseph J. (éd.), *The Palgrave Handbook of Philosophy and Money*, Basingstoke : Palgrave Macmillan, 2024, DOI : [10.1007/978-3-031-54140-7_12](https://doi.org/10.1007/978-3-031-54140-7_12).
- JEVONS William, *Money and the Mechanism of Exchange*. New York: D. Appleton and Company, 1875, 342 p.
- LOCKE John, « Considérations sur les conséquences de la diminution de l'intérêt et de l'augmentation de la valeur de l'argent », in *Écrits monétaires*, BRIOZZO Florence (trad.), Paris : Classiques Garnier, 2011 (1696), 463 p.
- MÄKI Uskali, « Reflections on the Ontology of Money », *Journal of Social Ontology*, vol. 6, n° 2, De Gruyter, 2020, pp. 245–263, DOI : [10.1515/jso-2020-0063](https://doi.org/10.1515/jso-2020-0063).
- MEISSNER Christopher M. et OOMES Nienke, « Why do countries peg the way they peg? The determinants of anchor currency choice », *Journal of International Money and Finance*, vol. 28, n° 3, 2009, pp. 522–547, DOI : [10.1016/j.jimonfin.2008.08.007](https://doi.org/10.1016/j.jimonfin.2008.08.007).
- NAIN Aditya et JUNG Pradesh G., *Understanding Money: Philosophical Frameworks of Monetary Value*, Taylor & Francis, 2021, 133 p.
- VON NEUMANN John et MORGENSTERN Oskar, *Theory of Games and Economic Behavior*, Princeton : Princeton University Press, 2004 (1944), 776 p., DOI : [10.1515/9781400829460](https://doi.org/10.1515/9781400829460).
- PETER Markus, SCHMUTZ Daniel et DREGEN Bernard, « Monnaie », MARTIN Pierre-G (trad.), *Dictionnaire historique de la Suisse (HLS DHS DSS LIS)*, 18.02.2014, URL : hls-dhs-dss.ch/fr/articles/047970/2014-02-18/, consulté le 12.12.2024.
- PETTIGREW Richard, « Risk, rationality and expected utility theory », *Canadian Journal of Philosophy*, vol. 45, n° 5–6, Routledge, 2015, pp. 798–826, DOI : [10.1080/00455091.2015.1119610](https://doi.org/10.1080/00455091.2015.1119610).
- PITEAU Michel, « Monnaie de compte et système de paiements chez James Steuart: Quel rôle pour la stabilité bancaire ? », *Revue économique*, vol. 53, n° 2, 2002, pp. 245–271. DOI : doi.org/10.3917/reco.532.0245.
- SIMMEL Georg, *The Philosophy of Money*, BOTTOMORE Tom et FRISBY David (trad.), London : Routledge, 2011 (1900), 537 p.
- ZELMANOVITZ Leonidas, *The Ontology and Function of Money: The Philosophical Fundamentals of Monetary Institutions*, Lexington Books, 2015, 471 p.

Divers

- ARISTOTE, *Traité du ciel*, DALIMIER Catherine et PELLEGRIN Pierre (trad.), Paris : Flammarion, 2004 (350 AEC), 465 p.
- BALBONI Julien, « Des amendes pénales impayées pour près d'un milliard d'euros en cinq ans », *L'Echo*, 08.10.2024, URL : www.lecho.be/economie-politique/belgique/general/des-amendes-penales-impayees-pour-pres-d-un-milliard-d-euros-en-cinq-ans/10567716.html, consulté le 15.12.2024.
- BOSTROM Nick, « Ethical Issues in Advanced Artificial Intelligence », in SMIT Iva *et alii* (dir.), *Cognitive, Emotive, and Ethical Aspects of Decision Making in Humans and in Ai*, Windsor, Ont. : International Institute for Advanced Studies in Systems Research and Cybernetics, 2005 (2002), pp. 12–17, URL : nickbostrom.com/ethics/ai, consulté le 15.12.2024.
- CADY Marlene, « The pure joy of being alive », *People Weekly*, vol. 32, 03.07.1989, pp. 64–71.
- MCCARTHY Rory, « Israeli troops open fire on women outside mosque », *The Guardian*, 03.11.2006, URL : www.theguardian.com/world/2006/nov/04/israel, consulté le 15.12.2024.
- DREGER Alice Domurat, *One of Us: Conjoined Twins and the Future of Normal*, Cambridge, Mass. : Harvard University Press, 2004, 198 p.
- LE GUIN Ursula K., *The Left Hand of Darkness*, New York : Ace Books, 1969, 286 p.
- ISMARD Paulin. « Le simple corps de la cité : Les esclaves publics et la question de l'État grec », *Annales. Histoire, Sciences Sociales*, 2014, vol. 69, n° 3, pp. 723-751, DOI : 10.3917/anna.693.0723, URL : shs.cairn.info/revue-Annales-2014-3-page-723, consulté le 15.12.2024.
- KOSKO Bart et ISAKA Satoru, « Fuzzy Logic », *Scientific American*, vol. 269, n° 1, 1993, pp. 76–81.
- STEUDEMAN Michael J., « Rhetorical Education in Times of Demagoguery », *Rhetoric Society Quarterly* 49, 2019, pp. 297–314, DOI : [10.1080/02773945.2019.1610642](https://doi.org/10.1080/02773945.2019.1610642).
- WHITE Donna, « Dancing with Dragons: Ursula K. Le Guin and the Critics », *Faculty Book Publications*, 1999, URL : orc.library.atu.edu/atu_faculty_books/48, consulté le 15.12.2024.