

EXPLICATION MÉCANISTE ET CAUSALITÉ DESCENDANTE

MAX KISTLER¹

in Thierry Martin (éd.), *L'unité de la science*, Ed. Vuibert, 2009.

Introduction

Selon la conception classique de l'empirisme logique (Nagel, 1961), l'unification des théories scientifiques passe par la réduction : une théorie de haut niveau, par exemple psychologique, est réduite à une théorie de niveau inférieur, par exemple neurophysiologique, si l'ensemble des lois de la première peut être déduit de la seconde. Cependant, de nombreux cas d'unification théorique ne satisfont pas à cette contrainte. Une explication réductrice, par exemple d'un phénomène cognitif, est plus satisfaisante qu'une explication purement psychologique : le recours à des concepts neurophysiologiques ou chimiques permet de compléter l'explication psychologique. Cependant, le recours aux concepts de haut niveau reste indispensable. Lorsqu'on explique par exemple l'amnésie rétrograde, c'est-à-dire l'absence de mémoire à long terme portant sur un bref intervalle temporel précédant un traumatisme cérébral, le cadre de l'explication se situe sur le plan de la psychologie de la mémoire : on mentionne une régularité découverte par l'expérimentation psychologique, selon laquelle la fixation dans la mémoire à long terme est interrompue lorsque le traumatisme intervient entre 15 secondes et 20 minutes après l'épisode d'apprentissage. La découverte du mécanisme microscopique sous-jacent permet de compléter cette explication psychologique, sans pourtant pouvoir s'y substituer. Le mécanisme neurophysiologique sous-jacent à la mémoire à long terme permet notamment d'expliquer pourquoi un traumatisme, dans cet intervalle temporel, interrompt la mémorisation. Mais il est logiquement impossible d'expliquer un phénomène cognitif exclusivement au niveau neuronal ou moléculaire : il est nécessaire que le mécanisme, décrit en termes neuronaux et moléculaires, soit articulé aux concepts cognitifs, tels que le conditionnement, le stimulus conditionné, la mémoire à long terme etc. De telles explications témoignent de l'apparition de nouvelles « sciences inter-niveaux », telles que les neurosciences cognitives, qui intègrent plusieurs théories – la psychologie, les neurosciences et la biochimie – tout en les modifiant².

Le concept de mécanisme peut contribuer à faire évoluer la compréhension de l'unification théorique et à dépasser le débat entre les conceptions de la réduction en termes de lois-pont (Nagel, 1961), d'identification (Causey, 1977, Schaffner, 1993), ou d'élimination (P.S. Churchland, 1986, Bickle, 2003). En outre, l'analyse de l'explication mécaniste peut apporter un éclairage nouveau au débat sur la possibilité de processus causaux qui « traversent

¹ Département de philosophie, Université Pierre Mendès France, Grenoble
Maximilian.Kistler@upmf-grenoble.fr

² Le concept de théorie inter-niveau a été introduit par Darden et Maull (1977). Des auteurs comme Bechtel et McCauley l'ont utilisé pour forger une nouvelle conception de l'unification des théories scientifiques qui ne requiert plus la réduction des théories de haut niveau.

les niveaux » : certaines modifications des parties d'un système complexe modifient ses propriétés systémiques, et certaines modifications au niveau du système peuvent avoir des incidences causales au niveau de ses composantes. Nous nous concentrerons sur la question de savoir s'il existe des processus causaux « descendants », ou s'il s'agit là d'une apparence trompeuse, comme le soutiennent Craver et Bechtel (2006). Leur thèse s'accorde avec la conclusion d'un argument important de Kim (1998) selon lequel la causalité descendante soulève des difficultés conceptuelles insurmontables.

Le mécanisme de la fixation de la mémoire

Lorsqu'on explique le comportement d'un système complexe à partir de la notion de mécanisme, on montre comment l'articulation des parties du système assure la transition régulière d'une condition initiale à une condition terminale³. Prenons l'exemple du mécanisme de la fixation de la mémoire à long terme. On considère que le processus fondamental sous-jacent consiste à renforcer de manière durable l'efficacité synaptique : la potentialisation à long terme (LTP pour *long-term potentiation*) conduit à changer la taille des synapses, soit en augmentant la surface des synapses existantes soit en en créant de nouvelles.

Nous devons nous contenter ici d'une esquisse très simplifiée de l'état actuel des connaissances sur ce mécanisme. Les potentiels d'action par lesquels l'information est transmise le long des neurones, sont relayés, lors du passage d'un neurone à l'autre, par un mécanisme chimique, esquissé dans la fig. 1.

Lorsqu'un potentiel d'action suffisamment intense parvient à la terminaison d'un neurone qui est séparé du neurone adjacent par l'espace intersynaptique, un neurotransmetteur est libéré par le neurone présynaptique. En l'occurrence, il s'agit de glutamate. Les molécules de glutamate se diffusent dans l'espace intersynaptique. La membrane de la cellule postsynaptique contient différentes sortes de récepteurs de glutamate : ce sont de grosses protéines qui traversent la membrane. Elles ont tendance à former des liaisons avec les molécules de glutamate. La réception d'une molécule de glutamate sur la membrane postsynaptique peut avoir différentes conséquences. Au cas où le récepteur est de type AMPA⁴, un nouveau potentiel d'action peut être engendré pour se propager le long du neurone postsynaptique, permettant ainsi la transmission nerveuse. Mais la membrane postsynaptique contient aussi une seconde sorte de récepteurs de glutamate, appelées NMDA⁵. Si la cellule postsynaptique est fortement stimulée, par une grande quantité de glutamate et une forte dépolarisation, ces récepteurs changent de configuration quand elles se lient à une molécule de glutamate : elles ouvrent un pore à travers la membrane qui est perméable aux ions de Ca²⁺. Ces ions entrent massivement dans la cellule postsynaptique où elles déclenchent une chaîne de réactions biochimiques qui débouche sur la synthèse de nouvelles protéines qui provoquent la croissance de dendrites, ce qui conduit à augmenter la surface de la synapse ou à la création de nouvelles synapses.

La LTP est un mécanisme qui conduit au renforcement de la réponse de la cellule postsynaptique (à droite dans la fig. 1) à une stimulation par un potentiel d'action en provenance de la cellule présynaptique (à gauche, dans la fig. 1). L'explication de ce mécanisme utilise nécessairement des concepts appartenant à différents niveaux de description. Pour commencer

³ Voir (Machamer, Darden et Craver, 2000, p. 3).

⁴ « AMPA » est l'abréviation de l'acide amino-hydroxy-méthyl isoxazole propionique.

⁵ « NMDA » est l'abréviation du N-méthyl-D-aspartate.

avec le niveau le plus élevé qui est celui des capacités comportementales d'un animal, on ne peut comprendre le mécanisme sans parler de sa fonction : il s'agit de forger des souvenirs aptes à modifier ses réactions comportementales. Un tel apprentissage est schématisé dans la fig. 2. Dans un protocole expérimental classique, on observe comment des rats apprennent à s'orienter dans un labyrinthe. Un « labyrinthe d'eau de Morris » est inondé ; les animaux s'y déplacent à la nage. Lorsqu'un rat normal découvre une plateforme immergée et donc invisible sur laquelle il a pied, il mémorise très vite son emplacement. En revanche, un rat dont le mécanisme de la LTP est inopérant est incapable de le mémoriser.

L'analyse du mécanisme de l'apprentissage permet de comprendre à la fois la formation du souvenir chez l'animal sain et l'amnésie de l'animal souffrant d'une pathologie : on complète la description du changement des capacités comportementales de l'animal (dans la fig. 2, le plus haut niveau) par des informations concernant la modification de structures à des niveaux inférieurs, c'est-à-dire au niveau de certaines composantes de l'animal. L'apprentissage est notamment rendu possible par la création, dans l'hippocampe de l'animal, d'une structure neuronale qui joue le rôle d'une « carte cognitive », dans notre exemple la carte du labyrinthe (dans la fig. 2, le second niveau). Afin de comprendre la genèse de cette carte cognitive dans l'hippocampe, il faut faire appel à des processus qui se déroulent à un niveau encore inférieur : celui des composantes du réseau neuronal de l'hippocampe. Il s'agit de modifications des synapses entre les différents neurones de la voie latérale de Schaffer (dans la fig. 2, le troisième niveau). Enfin, on ne peut comprendre la modification des synapses sans faire appel à des processus qui se déroulent au niveau des composantes moléculaires des synapses, notamment les canaux ioniques des récepteurs NMDA (dans la fig. 2, le niveau le plus bas).

La question des relations causales traversant les niveaux

L'analyse des explications mécanistes nous met en mesure d'aborder à nouveaux frais la question largement débattue de la possibilité de relations causales qui « traversent les niveaux ». Selon certains, nous ne pouvons être des agents libres que si nous pouvons agir sur le monde physique en vertu de nos décisions, et plus généralement en vertu de nos états et processus mentaux⁶. La possibilité d'une telle causalité « descendante » appartient au cœur de la doctrine de l'émergence, développé au 19^e et au début du 20^e siècle⁷, selon laquelle 1) certains systèmes complexes possèdent des propriétés dites systémiques qui sont qualitativement différentes des propriétés de leurs parties et 2) ces propriétés peuvent causalement modifier des états de choses au niveau des composantes des systèmes. Cependant, l'existence de telles relations causales descendantes a été remise en cause par un argument influent dû à Kim (1998)⁸. En effet, Kim soutient que la causalité descendante est incompatible avec deux principes métaphysiques plausibles. 1) Selon le « Principe de la clôture causale du monde physique », pour chaque événement physique e (voir fig. 3) qui a lieu à l'instant t , il existe, à chaque instant antérieur t^* , une cause complète physique c . Si l'on envisage la possibilité qu'un événement physique e soit le produit d'une cause C de niveau

⁶ Davidson (1963) montre qu'il n'est justifié de dire qu'une personne a agi pour une raison donnée que si la représentation de cette raison a aussi été la cause de son action. Voir Kistler, 2006b.

⁷ L'ouvrage de Broad, *Mind and its Place in Nature* (Broad, 1925) est souvent considéré comme l'œuvre majeure de ce courant.

⁸ Pour une analyse plus détaillée de l'argument de Kim, voir Kistler, 2005 et 2006a.

mental à l'instant t^* , ce principe implique que l'événement e ait aussi une cause physique complète à ce même instant t^* , représentée par « c » dans la fig. 3.

2) Le second principe dit que la surdétermination causale est exceptionnelle : il exclut que tous les événements physiques e qui seraient causés par des événements mentaux C soient « surdéterminés », au sens où ils auraient, outre une cause mentale complète C , une autre cause physique complète c contemporaine de C . En réalité, sauf peut-être dans des cas exceptionnels, e n'a qu'une seule cause complète dont le premier principe affirme qu'elle est physique. Il est donc impossible que la cause mentale C agisse de manière descendante sur l'événement physique e .

La présentation des résultats de la recherche scientifique sur les mécanismes s'accompagne souvent de l'affirmation que la causalité peut traverser les niveaux de complexité. « L'analyse biologique de l'apprentissage, dit Kandel, requiert la mise en évidence d'une relation causale entre des molécules spécifiques et l'apprentissage » (Kandel, 2000, p. 1268). Kandel reconnaît en particulier l'existence de la causalité descendante : « L'apprentissage produit des changements dans l'efficacité des connexions neuronales » (Kandel, 2000, p. 1275), et la psychothérapie serait impossible sans la causalité descendante : « dans la mesure où l'intervention sociale fonctionne [...], [par exemple] par la psychothérapie, [...] il faut qu'elle fonctionne en agissant sur le cerveau » (*ibid.*).

L'analyse expérimentale des mécanismes semble effectivement mettre en évidence des relations causales qui coupent à travers les niveaux de composition : dans des expériences « ascendantes » (ou *bottom-up*), on intervient expérimentalement sur une composante d'un mécanisme, pour en observer les conséquences au niveau des propriétés systémiques. Un type important d'expériences de ce genre est la technique du *knockout* génétique. Pour étudier par exemple la contribution du récepteur NMDA (dans la région pertinente de l'hippocampe) sur le mécanisme de l'apprentissage par LTP, on crée une population d'animaux transgéniques dont le matériel génétique a été modifié de sorte à enlever (ou à rendre inopérant) le gène responsable de la production de la protéine NMDA. On observe que la capacité d'apprentissage d'animaux dont l'hippocampe est dépourvu de récepteurs NMDA est diminuée. L'hypothèse selon laquelle ces récepteurs sont des constituants nécessaires d'un mécanisme de l'apprentissage permet d'expliquer ce phénomène.

Nous devons nous limiter ici, faute de place, à considérer les expériences qui semblent mettre en évidence la causalité descendante. Dans une telle expérience, on intervient expérimentalement sur une propriété systémique, pour en observer les conséquences au niveau des propriétés des composantes du système. On fait par exemple explorer un labyrinthe aquatique (tel que le labyrinthe de Morris mentionné dans la fig. 2) par un rat, ce qui entraîne la modification des propriétés cognitives du rat : il apprend l'architecture du labyrinthe et acquiert la capacité de s'y orienter. Lorsqu'on observe ensuite que cet apprentissage (niveau cognitif et comportemental) a eu des conséquences au niveau des neurones de l'hippocampe de l'animal, notamment la modification des synapses de la voie collatérale de Schaffer, il semble qu'on ait découvert une relation causale descendante, du niveau cognitif au niveau neuronal et moléculaire.

Selon certains auteurs⁹, le fait que l'on puisse manipuler une variable B en intervenant sur une variable A, constitue une condition nécessaire et suffisante pour l'existence d'une relation causale entre A et B. Une telle analyse de la causalité en termes de manipulabilité

⁹ L'analyse de la causalité en termes d'interventions a récemment connu un renouveau important, notamment grâce aux travaux de Woodward (2003).

apporte une justification conceptuelle à l'intuition selon laquelle les stratégies expérimentales ascendantes et descendantes que nous venons de mentionner témoignent de l'existence de relations causales qui traversent les niveaux.

Cependant, l'argument de l'exclusion causale présenté par Kim et mentionné plus haut semble montrer que l'existence de telles relations causales ne peut être qu'illusoire. En effet, Craver et Bechtel (2006) suggèrent que les processus causaux qui sont apparemment *bottom-up* ou *top-down* sont toujours le résultat d'une combinaison d'un processus causal intra-niveau et d'une relation de détermination non-causale, qu'ils interprètent comme une relation de « constitution ».

La constitution est une notion métaphysique qui permet de distinguer un objet matériel 1) de sa matière ou 2) de l'ensemble de ses parties. L'application du concept de constitution pertinente ici concerne le rapport de l'ensemble des parties d'un objet macroscopique à cet objet lui-même. Unger (1980) considère l'exemple d'un nuage, mais la situation logique est la même pour n'importe quel objet composé de parties, tels que les tables, les chaises et les êtres vivants. L'ensemble des gouttes d'eau qui forment un nuage n'est pas identique avec le nuage : si l'on considère l'évolution du nuage dans le temps, le concept de nuage est tel qu'un nuage peut persister, en d'autres termes continuer à exister en restant le même nuage, alors que des gouttes individuelles tombent en dehors du nuage ou s'y ajoutent. L'ensemble des gouttes ne peut donc pas être identique au nuage : quand une goutte tombe, le nuage continue d'exister alors que l'ensemble des gouttes cesse d'exister. Deux ensembles ne sont identiques que s'ils ont exactement les mêmes membres, et l'ensemble des gouttes dans le nuage avant la chute d'une goutte n'est pas le même que l'ensemble après la chute d'une goutte. Par ailleurs, *à un instant donné*, il aurait été *possible* que ce même nuage contienne un nombre légèrement différent de gouttes d'eau. Si le nuage était identique à l'ensemble des gouttes, il le serait nécessairement. Kripke (1972) a en effet montré que tout énoncé d'identité « A=B » où « A » et « B » sont des désignateurs rigides (des expressions qui font référence au même objet dans tous les mondes possibles) est, s'il est vrai, nécessairement vrai. La contraposée de cette thèse est qu'une relation contingente (car non nécessaire) ne peut pas être une relation identité. La possibilité qu'ils soient différents suffit donc à montrer que le nuage n'est pas identique à l'ensemble des gouttes. Là aussi, le concept de constitution permet de désigner le rapport entre l'ensemble des parties et le tout : l'ensemble de ses gouttes constitue le nuage sans qu'ils soient identiques¹⁰.

Nous ferons référence, dans ce qui suit, à trois propriétés de la relation de constitution : premièrement, il s'agit d'une relation antisymétrique au sens où, si A constitue B, il est impossible que B constitue A. L'ensemble de gouttes constitue le nuage, mais non l'inverse : on ne peut pas dire que le nuage constitue l'ensemble de ses gouttes. Deuxièmement, un même objet peut être, successivement ou alternativement, constitué de différentes matières ou de différents ensembles de parties. On peut exprimer cela en disant que la constitution peut être multiple. Troisièmement, la constitution est une relation logique et métaphysique, ce qui l'oppose à des relations épistémiques et nomologiques. L'ensemble de gouttes constitue le nuage, indépendamment de notre connaissance ou ignorance de ce fait. La constitution n'est pas une relation *épistémique* dans la mesure où elle préexiste à la connaissance que nous pouvons acquérir d'elle. La manière dont on justifie une relation de constitution la distingue d'une relation *nomologique* : nos hypothèses sur les lois de la nature doivent être justifiées de

¹⁰ L'argument qui montre que la constitution diffère de l'identité est développé dans Johnston, 1992 et Lowe, 2002, p. 73-76.

manière *a posteriori*, en s'appuyant sur des observations indépendantes (et souvent indirectes) des termes de la relation nomologique. On confirme la loi de la chute libre en mesurant la distance et le temps de la chute de différents objets en chute libre, *indépendamment* l'un de l'autre. Si je ne mesure que le temps de la chute pour déterminer la distance, en calculant la distance à partir du temps mesuré et de la loi $d=1/2gt^2$, il ne s'agit pas d'une confirmation indépendante de la loi, mais d'une pétition de principe. En revanche, si je connais la distribution précise des gouttes d'eau dans le ciel, je connais le nuage avec toutes ses propriétés : sa position, sa forme, sa densité etc. L'inférence de l'ensemble de gouttes au nuage est *conceptuel*, ce qui signifie que les objets de ces inférences – l'ensemble de gouttes et le nuage - sont liés par un lien de nécessité logique ou métaphysique et non nomologique.

A la lumière de ces remarques sur le concept de constitution, retournons à la manière dont Craver et Bechtel proposent d'analyser la causalité descendante. Pour reprendre leur exemple, une personne décide de commencer un match de tennis. La décision cause un comportement adéquat qui entraîne une consommation augmentée de glucose dans les cellules musculaires. Il semblerait que la décision, un événement qui appartient au niveau cognitif, ait des effets à des niveaux inférieurs, et notamment aux niveaux cellulaires et moléculaires. Or, selon Craver et Bechtel, cette apparence est trompeuse : « La situation peut être décrite sans reste, en n'ayant recours qu'à des causes à l'intérieur d'un même niveau et à des relations de constitution » (Craver et Bechtel, 2006, p. 13).

L'articulation d'une relation intra-niveau avec une relation de constitution - articulation censée produire l'illusion d'un processus causal descendant - peut prendre deux formes. Dans un premier scénario possible (fig. 4), la décision *C* qui appartient au niveau cognitif se trouve dans un rapport de constitution par rapport à un événement cérébral sous-jacent *c*, de niveau neuronal ou moléculaire. L'événement moléculaire *c* cause ensuite, à ce niveau et sans aucune contribution de niveaux supérieurs, la consommation augmentée de glucose dans les cellules musculaires, *e*.

Dans un second scénario envisageable (fig. 5), la décision *C* cause, au niveau des propriétés systémiques de l'organisme, le comportement *E*. Ce dernier se trouve dans un rapport de constitution par rapport à un ensemble d'événements moléculaires *e* sous-jacents à *E* et simultanés à *E* : parmi les événements *e* qui constituent ensemble le comportement *E*, il y a par exemple des événements concernant des cellules musculaires dont la consommation de glucose augmente.

Pour parler de constitution dans ces cas, il faut généraliser le domaine d'application du concept de constitution, pour l'étendre, au-delà des substances, aux *événements* qui arrivent à ces substances. Par « événement », nous entendons ici simplement un objet à un moment donné, autrement dit, une « partie temporelle » d'une substance. En pratique, on s'intéresse normalement aux parties temporelles auxquelles surviennent des changements. Partant du fait qu'un système cognitif est constitué d'un ensemble de parties, on peut transférer la notion de constitution aux événements qui arrivent au système cognitif : les événements qui arrivent au système cognitif en tant que tout, comme la décision *C* et le comportement *E*, sont constitués d'ensembles d'événements qui arrivent à l'ensemble de ses parties.

C ne détermine pas complètement *c*, étant donné qu'une décision est compatible avec de nombreux événements neuronaux et moléculaires. C'est ce qu'on appelle la « réalisabilité multiple » des événements cognitifs. Cette réalisabilité multiple ne s'oppose pas au fait qu'un ensemble d'événements microscopiques soit constitutif d'un événement macroscopique : en effet, le même événement macroscopique pourrait être constitué par différents ensemble

d'événements microscopiques. Néanmoins, le concept de constitution ne semble pas pouvoir jouer le rôle que lui attribuent Craver et Bechtel : le scénario 1 est inadéquat parce que la décision C ne constitue pas l'événement cérébral sous-jacent c : la constitution est une relation des parties par rapport au tout ; et l'antisymétrie de la relation interdit que le tout constitue en même temps ses parties. Cela vaut tout autant de la constitution d'une substance par l'ensemble de ses parties que de la constitution d'un événement macroscopique (cognitif) par un ensemble d'événements microscopiques (neuronaux).

Le scénario 2 semble se heurter à la même difficulté : le comportement qui correspond au jeu de tennis (E) ne constitue pas l'ensemble d'événements moléculaires sous-jacents e . Le fait que les parties constituent le tout interdit la possibilité que le tout constitue l'ensemble des parties. Cependant, il est possible de réinterpréter le scénario 2 de manière à le rendre adéquat : par un processus causal au niveau de l'organisme, la décision C cause l'activité E de jouer au tennis. A chaque étape du processus allant de C à E , le fait que l'organisme soit dans un état conforme à l'exécution de l'action E décidée en C , exerce une contrainte sur ses parties. Le fait d'être dans l'état E limite l'espace des états possibles des cellules musculaires. Je suggère donc de remplacer, dans le schéma proposé par Craver et Bechtel, la relation de constitution par la relation de *contrainte*.

Une contrainte limite les possibilités d'évolution ou de changement accessibles à un système donné. Dans un système d'équations avec n variables, chaque équation impose une contrainte sur ces variables, au sens de limiter le nombre de valeurs que peuvent prendre les variables pour satisfaire l'équation. Si les variables représentent des degrés de liberté d'un système physique, c'est-à-dire des dimensions à l'intérieur desquelles l'état du système peut varier, la notion de contrainte acquiert une signification physique : chaque équation exprimant un lien entre les variables exprime une limite imposée aux possibilités d'évolution du système. Chaque équation réduit le nombre de degrés de liberté du système¹¹. Une boule qui roule sur un plan incliné est soumise à la contrainte de rester sur ce plan : ses possibilités de mouvement sont limitées dans la mesure où elle ne peut ni traverser le plan ni s'élever au dessus du plan. De la même manière, chaque propriété d'un système complexe est une source de contraintes qui s'exercent sur ses composantes. L'imperméabilité du réceptacle dans lequel il est contenu impose une contrainte sur le nombre de molécules. Chaque contrainte diminue le nombre d'états possibles dans lesquels peuvent se trouver les constituants. Les contraintes correspondent à des déterminations partielles de l'état du système. Tant qu'il y a moins de contraintes que de degrés de liberté, les contraintes auxquelles un système est soumis ne déterminent pas *totalemment* son état. La seule contrainte de la boule est de rester sur le plan incliné : si on fait abstraction des trois dimensions correspondant aux vitesses que la boule peut avoir, l'espace des états qui lui sont accessibles diminue de trois à deux : au lieu de pouvoir occuper toutes les positions selon trois dimensions spatiales, sa liberté ne s'étend que sur un plan à deux dimensions.

La notion de contrainte exprime l'idée d'une détermination partielle ; par ailleurs, la contrainte n'est pas associée à une « direction privilégiée » comme la constitution. On peut dire que l'état des parties exerce une contrainte sur l'état du tout ; mais il est tout aussi adéquat de dire que l'état du tout exerce une contrainte sur l'état des parties. En ce sens, la position d'un corps solide limite les degrés de liberté de la position des atomes qui le

¹¹ En mécanique classique, le formalisme de Lagrange permet de trouver la trajectoire d'un système dont le mouvement est soumis à un certain nombre de contraintes, en exprimant chaque contrainte par une équation.

composent, même s'il ne s'agit que d'une détermination partielle dans la mesure où les atomes gardent des degrés de liberté de position correspondant aux vibrations.

Les degrés de liberté relatifs à la position ne sont qu'un exemple. La notion de degré de liberté, et donc celle des contraintes qui restreignent ces degrés de liberté, peut être généralisée à toutes les propriétés déterminables d'un système qui peuvent prendre différentes valeurs déterminées. La température corporelle d'un animal correspond en ce sens à un degré de liberté qui est soumis à la contrainte, exercée par un mécanisme de régulation au niveau de l'organisme, de rester à l'intérieur de certaines limites. Mais cette température exerce à son tour une contrainte « descendante » sur les états de mouvement possibles des molécules qui constituent l'organisme. Le mécanisme de régulation de la température limite le nombre d'états accessibles aux molécules composant l'organisme en termes de leur énergie cinétique, en imposant une valeur moyenne à l'ensemble des énergies cinétiques des molécules. Au même sens, le fait qu'un système cognitif possède à un moment donné une propriété cognitive donnée, par exemple de percevoir consciemment un objet, ou le fait qu'il lui arrive un événement cognitif, par exemple la prise d'une décision d'exécuter l'action E, impose une contrainte sur les états possibles des parties du système cognitif. La perception consciente d'une balle de tennis en mouvement exerce une contrainte sur les états possible des parties du corps du sujet, et notamment sur l'état de l'activité de ses neurones. Elle est incompatible avec les yeux fermés, mais aussi avec beaucoup d'états perceptifs, par exemple avec la perception d'une scène immobile. Cependant, cette contrainte est partielle car l'état de perception consciente de la balle de tennis est compatible avec un grand nombre d'états neuronaux.

Il semble donc possible de modifier le scénario 2 suggéré par Craver et Bechtel en y substituant la notion de contrainte à celle de constitution : l'état du système cognitif qui correspond à l'exécution du comportement E, exerce une contrainte partielle sur l'état des composantes microscopiques, et notamment neuronales du système, sans le déterminer totalement. Cette suggestion semble s'accorder avec l'affirmation de Craver et Bechtel selon laquelle l'état global E « enrégimente (*enlists*) » ses parties. Cependant, l'impression que « l'enrégimentation » des parties du système par l'état du système soit équivalente à notre « contrainte » imposée par l'état systémique sur l'état des parties, est trompeuse. On peut même douter de la cohérence de la notion l'enrégimentation des parties par le tout : D'une part, Craver et Bechtel la présentent comme une forme d'*identité*. « Un changement dans l'activité du mécanisme en tant que tout n'est rien d'autre qu'un changement à l'égard d'une ou plusieurs composantes » (Craver et Bechtel, 2006, p. 13). Or, dans la mesure où une propriété systémique est compatible avec de nombreuses propriétés au niveau des parties du système, et un changement systémique compatible avec de nombreux changements au niveau des parties, il ne s'agit pas d'une identité. Mais d'autre part, Craver et Bechtel disent que l'enrégimentation est une forme de *constitution* : l'activité de jouer au tennis, disent-ils, « est en partie constituée par des activités aux articulations neuromusculaires » (Craver et Bechtel, 2006, p. 13). Cela rend leur caractérisation de l'enrégimentation incohérente : dans la mesure où nous avons montré plus haut que la constitution et l'identité sont incompatibles, il ne peut pas y avoir de relation qui est à la fois identité et constitution.

De toute manière, la relation entre l'état du système, comme celui d'exécuter l'action de frapper une balle de tennis, et l'ensemble des états des parties du système, ne peut pas être une forme de constitution. Il ne s'agit pas de constitution 1) parce qu'on ne peut pas dire que le tout soit constitutif de ses parties (on peut seulement dire que les parties sont constitutives du

tout), alors que l'on peut dire, comme nous l'avons vu, que le tout contraint l'état des parties et 2) parce que la détermination (partielle) de l'état des parties par l'état du tout n'est pas conceptuelle, mais repose sur des lois de la nature. La position d'un corps solide impose une contrainte sur la position des molécules qui le composent. Cependant, on ne peut pas savoir *a priori* quels sont les mouvements des molécules qui sont compatibles avec la contrainte. Cela requiert la connaissance *a posteriori* des lois de la nature qui imposent des contraintes supplémentaires sur les molécules en fonction de leur propriétés et en fonction de l'état systémique du solide. Il est encore plus clair que des recherches empiriques sont nécessaires pour déterminer quels états d'activité neuronale d'un système cognitif sont compatibles avec le fait que le système perçoit, par exemple, consciemment une balle de tennis.

Nous avons montré qu'il est cohérent de supposer qu'il existe des relations causales qui traversent les niveaux de composition. Nous avons vu que les notions de constitution et de contrainte permettent d'envisager la possibilité qu'un événement qui arrive à une partie ait un effet sur les propriétés systémiques et la possibilité qu'un événement qui se produit au niveau du système ait une influence causale sur les parties. L'application des concepts de constitution et de contrainte aux systèmes macroscopique présuppose que l'on puisse identifier des niveaux auxquels correspondent différentes sciences, et des propriétés et lois identifiées par ces sciences.

Cependant, la possibilité conceptuelle d'une influence causale qu'un système exerce sur ses propres parties en vertu de ses propriétés systémiques a été remise en cause par l'argument de Kim que nous avons présenté plus haut. Résumons brièvement cet argument : étant donné que l'état des parties du système e , au temps t , est un événement physique, le principe de la complétude du domaine physique garantit que l'événement e a une cause complète de même niveau physique, c , au temps t^* . Kim fait alors appel à un second principe méta-scientifique, le principe d'exclusion explicative, pour montrer qu' e ne peut pas avoir d'explication causale indépendante de son explication microscopique en termes de c , ce qui exclut en particulier l'explication d' e en termes de la propriété mentale C .

Je dois me contenter ici de quelques remarques qui suggèrent qu'il est après tout conceptuellement cohérent d'envisager la possibilité que l'état global d'un système complexe détermine partiellement et de manière non-causale l'état de ses parties. Dans un système complexe, il est possible que certains événements microscopiques – par exemple les modifications cellulaires ou moléculaires dans l'état d'un organisme vivant – ne soient pas complètement déterminés, à long terme, par des événements de même niveau. On peut en effet faire l'hypothèse selon laquelle le cerveau obéit aux lois du « chaos classique »¹². La possibilité de faire des prédictions sur l'évolution d'un système chaotique est limitée à un bref laps de temps. On ne peut expliquer, de manière déductive-nomologique (DN), un événement moléculaire (tel que la transformation d'une molécule d'ATP en ADP, ce qui dégage l'énergie nécessaire à la contraction musculaire), en termes d'événements moléculaires qui ont lieu une heure plus tôt.

Pour pouvoir tirer une conclusion ontologique – que l'état e de l'ensemble des parties du système à t n'est pas causalement déterminé par l'état c de l'ensemble des parties du système à t^* - de l'impossibilité d'une prédiction à long terme dans un système chaotique, il faut faire deux présuppositions. La première concerne l'interprétation de la notion de détermination causale. On peut analyser les relations causales sur deux niveaux¹³.

¹² Cf. Skarda et Freeman, 1990, Lehnertz et Elger, 2000, Newman, 2001.

¹³ Cette analyse est développée dans Kistler, 1998, 1999, et 2006c.

1. Les relations causales peuvent être conçues comme des relations entre événements particuliers, où on entend par « particulier » une chose ou un événement concret ayant de nombreuses propriétés. Sur ce plan, on peut faire l'hypothèse que la causalité repose sur la transmission d'une quantité d'énergie (ou d'une autre grandeur conservée) d'un événement sur l'autre.

2. Mais dans la plupart des contextes où l'on s'intéresse à une relation causale, à savoir parce qu'elle fait l'objet d'une explication scientifique, on envisage la causalité sur le plan de certaines *propriétés* bien déterminées des événements : on ne veut, par exemple, pas seulement établir qu'il *existe* une relation causale entre deux épisodes successifs dans la vie d'un organisme qui est en train d'apprendre une association par conditionnement classique. On cherche plutôt à comprendre un *fait* F_2 portant sur l'organisme à l'instant t_3 après un épisode d'apprentissage, en fonction d'un *fait* F_1 portant sur l'épisode d'apprentissage, de telle sorte que F_1 est causalement responsable de F_2 . F_2 peut être par exemple le fait que la force associative V_A entre un stimulus conditionné A et un stimulus inconditionné US a augmenté de ΔV_A , au cours d'un épisode de conditionnement où le sujet a été exposé, à t_1 , à un stimulus composé de A et de X, avant d'être exposé, à t_2 qui suit immédiatement t_1 , au stimulus US. Selon la théorie de Rescorla et Wagner¹⁴, ce qui est causalement responsable de F_2 est un fait qui porte sur le système cognitif à l'instant t_1 et sur l'épisode d'apprentissage qui a lieu entre t_1 et t_2 : F_1 porte sur les forces associatives V_A et V_X des stimuli A et X avant l'épisode de conditionnement¹⁵, ainsi que sur différents paramètres¹⁶. F_1 détermine causalement le fait F_2 , par l'intermédiaire d'une loi découverte par Rescorla et Wagner, selon laquelle la modification de la force associative est égale à $\Delta V_A = \alpha_A \beta_1 (\lambda_1 - V_{AX})$. Nous pouvons interpréter la détermination causale de F_2 par F_1 de manière ontologique : la force associative après l'épisode d'apprentissage est indépendante de notre connaissance et de nos descriptions de ces faits. Pour interpréter ce second aspect de la relation causale, on peut partir de l'analyse déductive-nomologique de l'explication causale : expliquer causalement F_2 , c'est produire un argument déductif-nomologique dont F_2 est la conclusion et dont les prémisses contiennent la condition initiale F_1 , ainsi qu'un certain nombre d'énoncés de lois de la nature. Dans une perspective réaliste, on peut inférer que, pour que cette explication causale puisse être vraie, il faut qu'il existe des faits F_1 , F_2 et des lois N exprimées par les énoncés nomologiques, tels que F_1 est causalement responsable de F_2 , étant donné les lois N¹⁷.

La seconde présupposition concerne l'interprétation de l'indétermination dans un système chaotique¹⁸ : il s'agit là aussi d'une indétermination ontologique et non seulement épistémique. Lorsqu'il s'agit de déterminer la valeur d'une grandeur mesurable d'un système physique dont les valeurs parcourent un continu, on ne peut attribuer aucun sens empirique à l'hypothèse que cette grandeur ait une valeur déterminée avec une précision *absolue*. Cela prive de fondement l'hypothèse selon laquelle ces grandeurs mesurables possèdent malgré tout objectivement des valeurs d'une précision infinie, ce qui aurait pour conséquence que

¹⁴ Rescorla et Wagner, 1972, p. 76.

¹⁵ Rescorla et Wagner supposent que la force V_{AX} du stimulus composé de A et X est égale à la somme $V_A + V_X$ des forces des stimuli A et X.

¹⁶ Il s'agit des paramètres suivants : α_A représente la saillance du stimulus A, β_1 est propre au stimulus inconditionné US utilisé et λ_1 représente la force maximale d'association qu'il est possible d'atteindre avec le stimulus US.

¹⁷ Pour une défense de cette analyse de la causalité, voir Kistler, 1999 et 2006c.

¹⁸ Nous nous limitons ici à la considération du chaos classique. La prise en considération de la mécanique quantique qui prédit l'existence de limites absolues à la précision des mesures, selon les relations dites « d'incertitude », pose des problèmes que dépassent le cadre de cet article.

l'évolution du système serait elle aussi objectivement déterminée avec une précision infinie pendant en temps infini. Le contenu d'une telle hypothèse outrepasserait les limites de ce qui est, même en principe, accessible à la connaissance scientifique. En ce sens, même si l'on suppose avoir déterminé l'état d'un système chaotique au temps t^* avec la précision objectivement et absolument maximale, cet état à t^* ne détermine pas complètement l'état du système à des temps ultérieurs t suffisamment éloignées de t^* .

L'incomplétude objective, dans un tel système chaotique, de la détermination « horizontale » des événements physiques sur le plan physique, remet en cause le « principe de la complétude du domaine physique ». Dans un tel système il n'existe, pour un événement e donné qui a lieu à l'instant t , et pour des instants t^* qui précèdent t par un laps de temps suffisamment long, aucun événement c de même niveau et ayant lieu à t^* , qui détermine e complètement. Cela ne signifie pas qu'un tel événement e soit totalement indéterminé. Le succès de l'éthologie et de la psychologie dans l'explication de nombreux comportements animaux et humains montre que ces derniers obéissent à des « lois de système »¹⁹, et notamment aux lois cognitives de la détermination des actions par le raisonnement et la décision. Le fait qu'un organisme obéisse à de telles lois systémiques signifie qu'il existe des contraintes globales sur son évolution. Il est possible qu'un événement moléculaire qui a lieu dans un organisme vivant soit déterminé en partie par des contraintes d'ordre moléculaire et en partie par des contraintes « descendantes » : l'organisme doit évoluer en accord avec certaines lois de système.

Il est important d'être conscient de la nature de la démarche qui a été ici la nôtre : aucun argument philosophique ne peut démontrer qu'il existe des événements physiques qui ne sont pas, à des instants t^* qui les précèdent d'un intervalle suffisamment long, complètement déterminés par des événements physiques ayant lieu à t^* . Aucun argument philosophique ne peut montrer non plus qu'il existe des lois systémiques qui déterminent l'évolution des systèmes complexes indépendamment de leur état physique précis, en fonction de leurs propriétés systémiques. Notre but ne peut donc être que celui de montrer 1) que cela n'est pas conceptuellement incohérent et 2) que cela ne contredit aucun principe métaphysique fondé sur l'interprétation de la science dans son ensemble, comme les principes kimien de complétude et d'exclusion.

Nous sommes parvenus au résultat que l'existence de relations causales descendantes est conceptuellement cohérente. Il est *possible* que le domaine des événements physiques ne soit pas toujours clos, contrairement à ce que soutient Kim et de nombreux autres physicalistes. En l'absence d'une explication complète et exclusivement physique d'un événement physique e donné, le principe d'exclusion explicative ne s'applique pas : il est clair que l'explication de e qui mentionne l'état cognitif C , ainsi qu'une loi psychologique en vertu de laquelle C conduit à l'action E ne peut pas être une explication *complète* de l'événement physique e . E ne fait qu'exercer une contrainte sur les états microscopiques possibles du système. A l'intérieur de cette contrainte, c'est l'état microscopique d qui précède e de peu qui détermine quel état microscopique e est effectivement réalisé. Pour expliquer l'état physiologique e de manière complète, il est donc nécessaire de faire appel à la fois à des facteurs cognitifs et à des facteurs physiologiques. L'état e est déterminé par un mécanisme que l'on ne peut comprendre et expliquer qu'en prenant en considération plusieurs de ses niveaux constitutifs.

¹⁹ L'expression est de Schurz, 2002. Dans la mesure où l'évolution ou le comportement d'un organisme obéit à des régularités, l'organisme est ce que Cartwright, 1999, appelle une « machine nomologique ».

Conclusion

L'analyse expérimentale des mécanismes, notamment en neurosciences cognitives, semble attester l'existence de relations causales « descendantes » : certaines interventions sur des propriétés systémiques (par exemple, cognitives) d'un organisme ont un impact au niveau des propriétés de ses composantes (par exemple des neurones et molécules). L'analyse de la notion de relation causale descendante conduit à l'hypothèse que certains événements microscopiques sont conjointement déterminés par des facteurs causaux de même niveau et par des contraintes descendantes dues à l'existence de lois de système. De tels événements n'obéiraient pas au « principe de la clôture causale du domaine des événements physiques », tel qu'il a été introduit par Kim²⁰.

Références

- Bickle J., 2003, *Philosophy and Neuroscience: A Ruthlessly Reductive Account*, Dordrecht, Kluwer.
- Broad C.D., 1925, *The Mind and its Place in Nature*, Londres, Harcourt, Brace and Co.; repr. Londres, Routledge, 2000.
- Cartwright N., 1999, *The Dappled World. A Study of the Boundaries of Science*, Cambridge, Cambridge University Press.
- Causey R.L., 1977, *Unity of Science*, Dordrecht, Reidel.
- Churchland P.S., 1986, *Neurophilosophy: Toward a Unified Science of the Mind-Brain*, Cambridge (Mass.), MIT Press ; trad. par M. Siksou, *Neurophilosophie: l'esprit-cerveau*, Paris, Presses Universitaires de France, 1999.
- Craver C.F. 2002, « Interlevel Experiments and Multilevel Mechanisms in the Neuroscience of Memory », *Philosophy of Science*, 69, S 83- S 97.
- Craver C.F. et Bechtel W., 2006, « Top-down causation without top-down causes », *Biology and Philosophy*, 1-17.
- Darden L. et Maull N., 1977, « Interfield Theories », *Philosophy of Science*, 44, 43-64.
- Davidson D., 1963, « Actions, Reasons, and Causes », in Davidson D., *Essays on Actions and Events*, Oxford, Clarendon Press, 1980, p. 3-19 ; trad. par P. Engel, *Actions et événements*, Paris, P.U.F., 1993.
- Glennan S., 1996, « Mechanisms and the Nature of Causation », *Erkenntnis*, 44, 49-71.
- Johnston M., 1992, « Constitution Is Not Identity » *Mind*, 101, 89-105.
- Kandel E.R., 2000, « Cellular Mechanisms of Learning and the Biological Basis of Individuality », in Kandel E.R., Schwartz J.H. et Jessell T.M., *Principles of Neural Science*, chap. 63, 1247-1279, New York, McGraw-Hill.
- Kim J., 1998, *Mind in a Physical World*. Cambridge (Mass.), MIT Press ; trad., *L'esprit dans un monde physique*, Paris, Syllepse, 2006.
- Kistler M., 1998, « Reducing Causality to Transmission », *Erkenntnis*, 48, 1-24.
- Kistler M., 1999, *Causalité et lois de la nature*, Paris, Vrin, Collection Mathesis, 1999. (Traduction anglaise, *Causation and Laws of Nature*, Londres, Routledge, 2006).
- Kistler M., 2005, « Is Functional Reduction Logical Reduction ? », *Croatian Journal of Philosophy*, 5, 219-234.

²⁰ Je remercie mes auditeurs à Genève et Madrid, ainsi que Reinaldo Bernal pour leur commentaires critiques sur des versions antérieures de cet article.

- Kistler M., 2006a, « The Mental, the Macroscopic, and their Effects », *Epistemologia*, 29, 79-102.
- Kistler M., 2006b, « Les causes des actions », *Le temps philosophique*, 12, 141-175.
- Kistler M., 2006c, « La causalité comme transfert et dépendance nomique », *Philosophie*, 89, 53-77.
- Lehnertz K. et Elger C.E. eds., 2000, *Chaos in Brain?*, World Scientific, Singapore.
- Lowe, E.J., 2002, *A Survey of Metaphysics*, Oxford, Oxford University Press.
- Machamer P., Darden L. et Craver C.F., 2000, « Thinking about Mechanisms », *Philosophy of Science*, 67, 1-25.
- Nagel E. 1961, *The Structure of Science*, Londres, Routledge and Kegan Paul.
- Newman D.V., 2001, « Chaos, Emergence, and the Mind-Body Problem », *Australasian Journal of Philosophy*, 79, 180-196.
- Rescorla, R. A. et Wagner A.R. (1972), « A Theory of Pavlovian Conditioning : Variations in the Effectiveness of Reinforcement and Nonreinforcement », in : A.H. Black et W.F. Prokasy (éds.), *Classical Conditioning, II: Current Research and Theory*, p. 64-99, New York, Appleton-Croft-Century.
- Salmon W., 1984, *Scientific Explanation and the Causal Structure of the World*, Princeton, Princeton University Press.
- Schaffner K., 1993, *Discovery and Explanation in Biology and Medicine*, Chicago, Chicago University Press.
- Schurz G., 2002, « *Ceteris paribus* Laws: Classification and Deconstruction », *Erkenntnis*, 57, 351-372.
- Skarda C. et Freeman W.J., 1990, « Chaos and the New Science of the Brain », *Concepts in Neuroscience*, 1, 275-285.
- Unger P., 1980, « The Problem of the Many », *Midwest Studies in Philosophy*, 5, French, P. A., Uehling, Jr. T. E., and Wettstein, H. K., eds., University of Minnesota Press, Minneapolis, p. 411-67.
- Woodward J., 2003, *Making Things Happen*, Oxford, Oxford University Press.

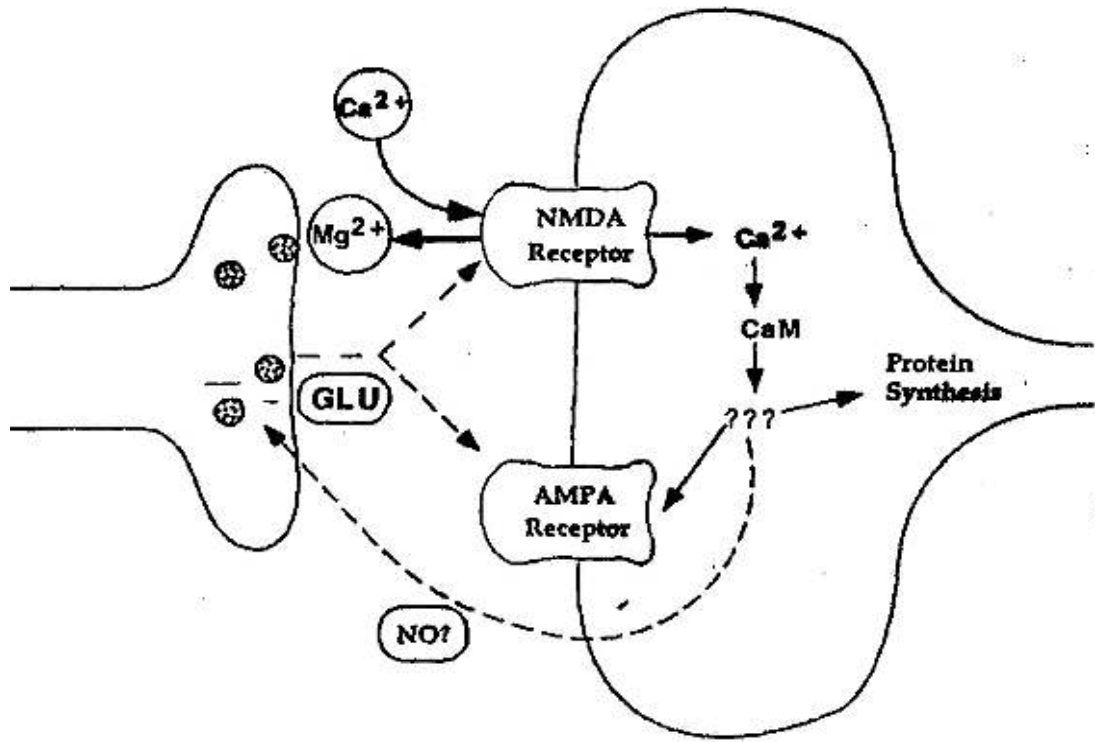


Fig. 1. Esquisse du mécanisme de la potentialisation à long terme (LTP) (de Craver 2002, p. S 87).

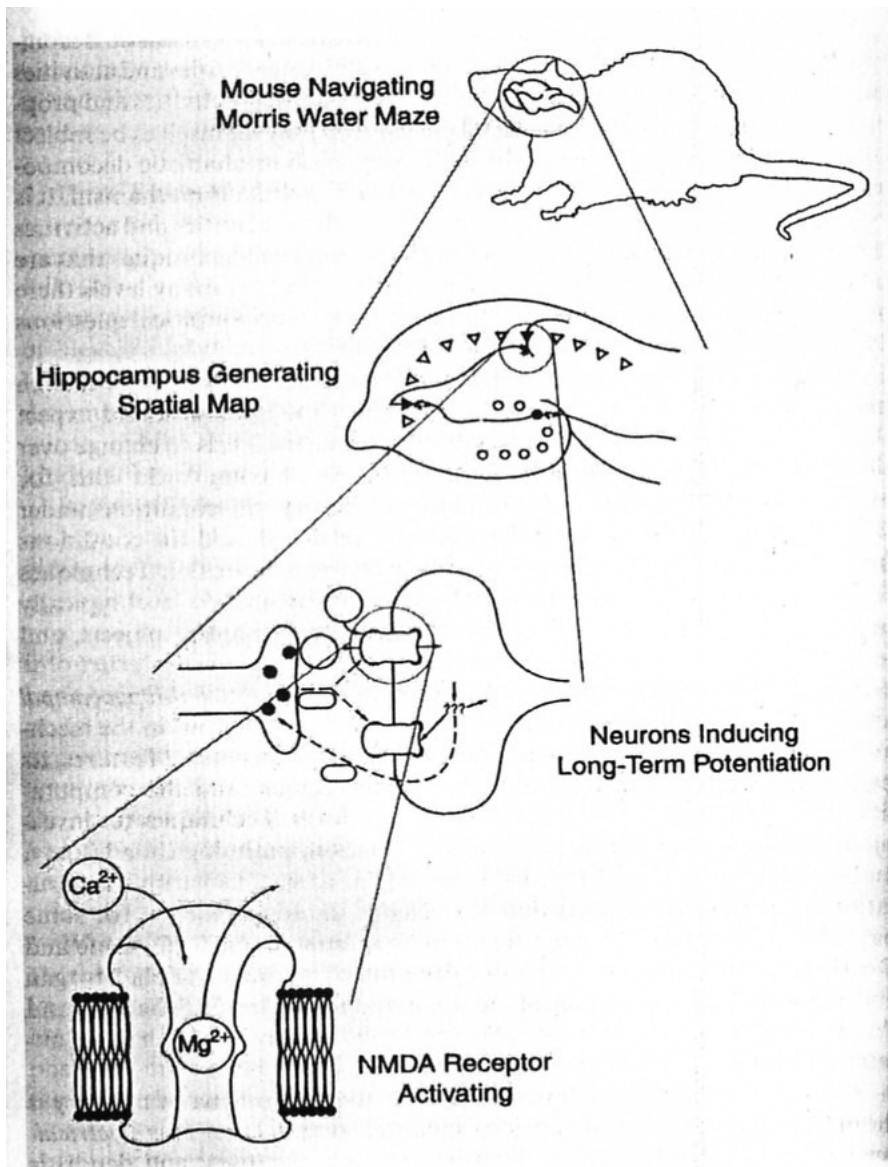


Fig. 2. Différents niveaux de l'analyse du mécanisme de la LTP (d'après Craver 2002, p. S90).

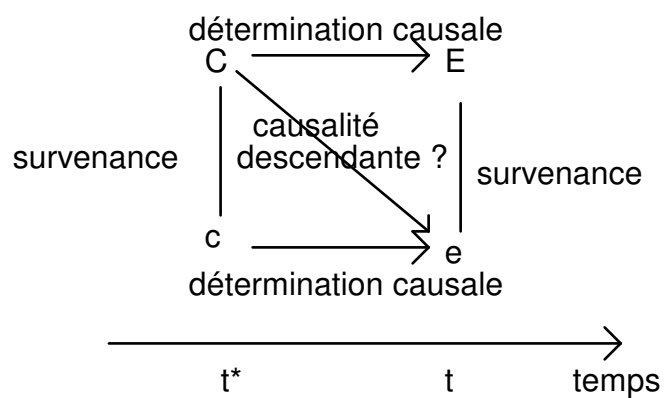


Fig. 3. La causalité descendante

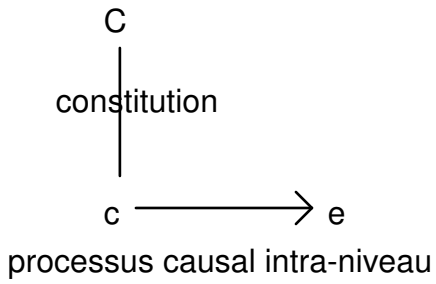


Fig. 4. Scénario 1 de la décomposition d'un processus causal descendant apparent

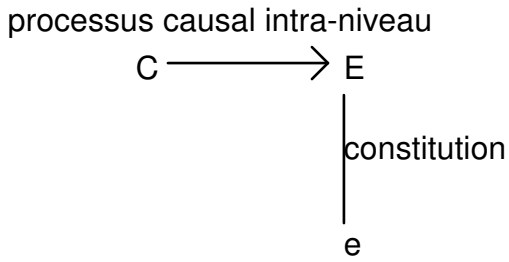


Fig. 5. Scénario 2 de la décomposition d'un processus causal descendant apparent

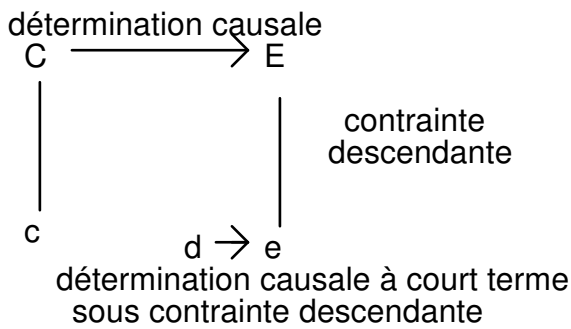


Fig. 6. Modèle de la causalité descendante: causalité intra-niveau et contrainte descendante