la théorie de l'identité

philosophie de la psychologie et des sciences cognitives
séance 4

M. Cozic

2. La théorie de l'identité

le physicalisme

- physicalisme (ou matéralisme): la réalité est ultimement et exclusivement constituée d'entités de nature physique
- Smart (1959)

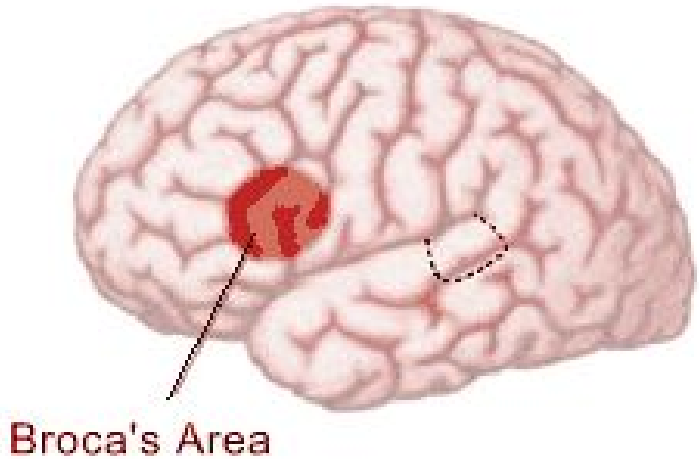
  "Il me semble que la science accrédite de plus en plus l'idée que l'on peut voir les organismes comme des mécanismes physico-mécaniques: il semble que même le comportement de l'homme sera un jour explicable en termes mécanistes. Pour autant que la science est concernée, il ne semble rien y avoir dans le monde sinon des arrangements de plus en plus complexes de constituants physiques"
- le physicalisme est la conception dominante en philosophie de l'esprit contemporaine même si tous les physicalistes ne se reconnaîtraient pas dans la formulation qu'en donne Smart donne
- nous reviendrons plus tard sur la caractérisation du physicalisme (voir Stoljar 2009 en attendant), mais d'abord nous allons nous concentrer sur une théorie physicaliste: la théorie de l'identité

la théorie de l'identité

- théorie de l'identité: les états, processus et événements mentaux sont identiques à des états, processus et événements cérébraux
- variété de physicalisme dont la forme actuelle remonte aux années 1950:
  ✓ U.T. Place "Is Consciousness a Brain Process?" (1956)
  ✓ H. Feigl "The 'Mental' and the 'Physical'" (1958)
  ✓ J.J.C. Smart "Sensations and Brian Process" (1959)
- la théorie de l'identité fonde sur les neurosciences la plausibilité d'une identification des phénomènes mentaux aux phénomènes cérébraux
- point de départ : les corrélations psychocérébrales, càd les corrélations entre phénomènes mentaux et phénomènes cérébraux

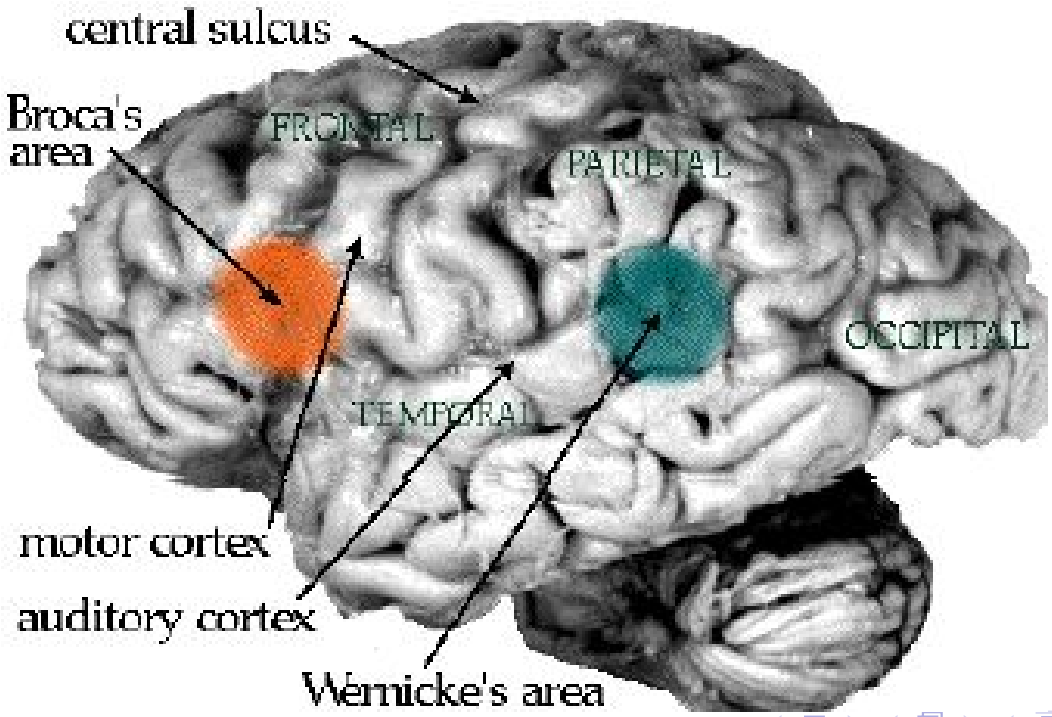
l'aire de Broca

- 1861: Paul Broca (1824-1880) présente le cas d'un patient récemment décédé, Leborgne, victime d'un A.V.C.
- Leborgne comprenait ce qu'on lui disait, mais ne parlait pas - il émettait slt le son "Tan". On parle d'aphasie (de Broca)
- l'autopsie révèle une lésion au niveau du lobe frontal gauche = aire de Broca



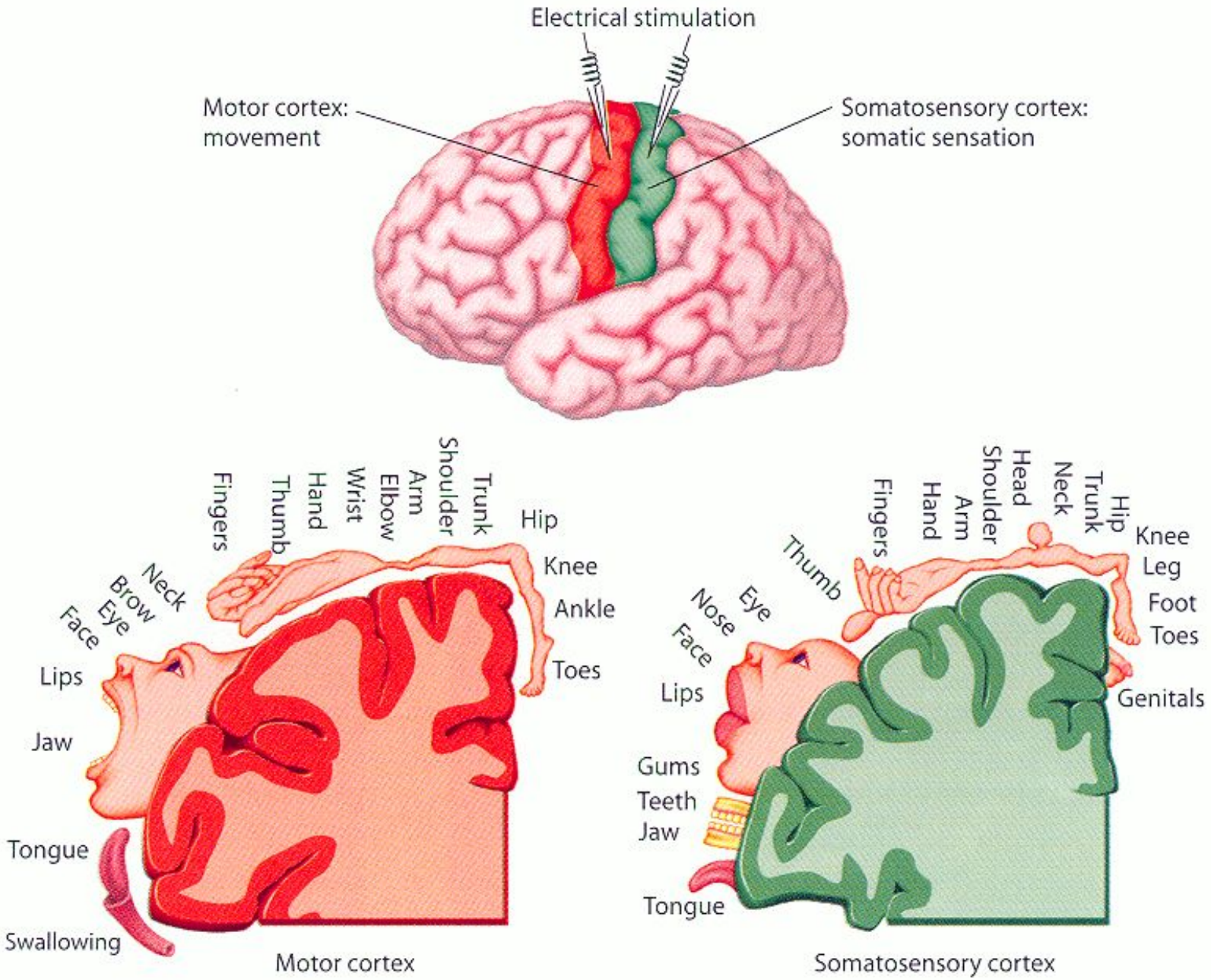
l'aire de Wernicke

- 1876: C. Wernicke (1848-1902) publie une étude sur un autre individu victime d'un A.V.C., qui peut parler sans difficulté mais dont le discours est largement dépourvu de sens
- ne comprend pas non plus le langage (écrit ou oral)
- lésion à un autre endroit du cerveau = aire de Wernicke



cartes fonctionnelles du cortex

- cartes du cortex moteur et somesthésique obtenues par stimulation électrique (Penfield et Jaspet, 1940-50)
- on parle d'homonculus pour désigner ce genre de correspondance


Electrical stimulation
Motor cortex: movement
Somatosensory cortex: somatic sensation
Fingers
Thumb
Hand
Wrist
Elbow
Arm
Shoulder
Trunk
Hip
Knee
Ankle
Toes
Face
Eye
Brow
Neck
Lips
Jaw
Tongue
Swallowing
Motor cortex
Fingers
Thumb
Hand
Arm
Shoulder
Head
Neck
Trunk
Hip
Knee
Leg
Foot
Toes
Genitals
Face
Nose
Eye
Lips
Gums
Teeth
Jaw
Tongue
Somatosensory cortex

les corrélations psychocérébrales

- extrapolation (optimiste) des corrélations psychocérébrales (Kim, 1998):
  Pour tout type d'état mental M et tout organisme o, il existe un type d'état cérébral C tel que

o est dans l'état M ssi o est dans l'état C

comment expliquer des corrélations?

- on peut expliquer l'existence de corrélations entre A et B de plusieurs façons:

(i) causalité : on peut explique le fait que A (la température chute sous 0) est corrélé à B (le lac gèle) par le fait que A *cause* B

(ii) cause commune: on peut expliquer le fait que A (l'heure affichée à une horloge h_a chez un horloger) est corrélé à B (la même heure affichée à une horloge h_b chez le même horloger) par le fait que A et B ont une cause commune (la synchronisation par l'horloger de ses horloges)

(iii) on peut expliquer le fait que A (la température d'un gaz) est corrélé à B (la pression du gaz) par le fait que A et B sont deux aspects d'un même processus (l'agitation moléculaire)

comment expliquer des corrélations?

(iv) identité: on peut expliquer le fait que A (l'éclair) soit corrélé à B (la décharge électrique entre nuages ou entre nuage et terre) par le fait que A et B sont identiques

- la théorie de l'identité choisit cette dernière voie : expliquer une corrélation par une identité.

quelle identité?

a priori et *a posteriori*

- cas paradigmatiques d'identité qui guident la TI :
- L'eau, c'est H20
- La température d'un gaz est l'énergie cinétique moyenne de ses molécules
- il y a des identités qui sont accessibles a priori, sans que l'on ait besoin de recourir à l'expérience:
- 2 est le plus petit nombre premier
- à l'inverse, il y a des identités que l'on découvre a posteriori, empiriquement, comme eau=H20
- pour les théoriciens de l'identité (la plupart en tous cas), les identités psychocérébrales sont de ce genre

quelle identité?

signification et référence

- la TI ne soutient pas que, si l'état mental *M* est l'état cérébral *C*, "*M*" et "*C*" aient la même signification.
- "*M*" et "*C*" n'ont pas la même signification, mais ils ont la même référence: ils désignent la même chose (cf. D. Lewis (1966))

types vs. exemplaires

- combien y a-t-il de mots sur la ligne ci-dessous?

  CHIEN CHIEN CHIEN

- 3 exemplaires du même type
- un type est une entité abstraite, générale ; un exemplaire est un particulier, concret

quelle identité?

identité exemplaire vs. typique

- il s'agit d'une identité typique, et non d'une identité exemplaire, entre phénomènes mentaux et phénomènes cérébraux

(i) identité exemplaire (*token identity*) : tout événement mental particulier *x* est identique à un événement physique (cérébral) particulier *y* (cf. Davidson, 1970)

- ▷ la douleur que j'ai ressentie aujourd'hui à 12h est un certain événement physique

(ii) identité typique (ou des occurences, *type identity* ou *type-type identity*) : tout type d'événement mental *M* est un type d'événement physique (cérébral) *C*

- ▷ La douleur est l'activation des fibres de type *C*

- L'identité typique est plus forte que l'identité exemplaire: on ne peut avoir identité typique sans identité exemplaire, mais il est possible d'avoir identité exemplaire sans identité typique

- Fodor, J. (1974)

  “Le physicalisme exemplaire dit simplement que tous les événements dont parlent les sciences sont des événéments physiques...
  ...le physicalisme exemplaire est plus faible que ce qu’on pourrait appeler le "physicalisme typique", doctrine selon laquelle toute *propriété* mentionnée dans les lois de n’importe quelle science est une propriété physique. Le physicalisme exemplaire ne permet pas d’inférer le physicalisme des types parce que l’identité contingente entre deux événements ne garantit pas en principe l’identité des propriétés que les événements exemplifient.”

faiblesse de la théorie de l'identité des exemplaires

- Kim (2006/2008)

  "...le physicalisme des occurrences est une doctrine faible qui ne nous raconte pas grand-chose: ...[elle] ne dit rien des rapports entre les propriétés mentales et les propriétés physiques, des rapports entre, d'un côté, les douleurs, les démangeaisons, les pensées ou les états de conscience et, de l'autre côté, les types d'événements mentaux. Le physicalisme des occurrences resterait vrai même s'il n'y avait pas l'ombre d'un lien systématique entre le mental et le physique...Soulignons que, pas plus que "tout objet qui a une couleur a une forme" n'implique que "les objets de la même couleur ont la même forme", le physicalisme des occurrences n'implique que (II) les événements (et les objets) qui sont semblables du point de vue physique doivent être semblables du point de vue mental."

TI et réductionnisme

- la théorie de l'identité (typique) est liée au réductionnisme si bien qu'on l'appelle parfois physicalisme réductionniste
- on conçoit traditionnellement la réduction comme une relation entre théories: une théorie T_1 (la théorie réduite) est réduite à une autre théorie T_2 (la théorie réductrice) si les lois de T_1 peuvent être dérivées de celles de T_2
- exemple: réduction de la thermodynamique (relations entre variables macroscopiques comme volume, température, pression) à la théorie cinétique des gaz (théorie du mouvement des particules dont sont faites les gaz)

TI et réductionnisme

- en général, les deux théories n'ont pas le même vocabulaire - n'emploient pas les mêmes concepts
- pour qu'il y ait réduction, il faut des lois qui établissent des liens entre les concepts des deux théories - lois que l'on appelle des lois-ponts
- exemple: température = énergie cinétique moyenne des molécules
- on peut concevoir la TI comme affirmant l'existence de lois-points (les identités) entre la psychologie et les neurosciences

2. Arguments en faveur de la TI

arguments en faveur de la TI

(i) explication des corrélations psychocérébrales

- question: mais pourquoi invoquer l'identité, plutôt que la causalité ou la causalité commune ?

(ii) simplicité et économie (Smart, 1959)
"Si l'on accepte qu'il n'y a pas d'arguments philosophiques contraignants qui nous forcent à accepter le dualisme, et si la théorie des processus cérébraux et le dualisme sont également compatibles avec les faits, alors les principes de parsimonie et de simplicité me semblent trancher radicalement en faveur de la théorie des processus cérébraux." (Smart, 1959)

arguments en faveur de la TI

- pourquoi la théorie de l'identité serait-elle "plus simple" ?

  ✓ simplicité ontologique: on ne suppose l'existence que d'une seule sorte de chose

  ✓ simplicité nomologique: on n'a pas à supposer l'existence d'un très vaste ensemble de lois de corrélations psycho-cérébrales (ou psycho-physiques en général).
  objection: la TI nous conduit à accepter un très vaste ensemble de lois d'identifications psycho-cérébrales.

  réponse: ces lois d'identification sont précisément bien moins coûteuses que des lois de corrélation *sui generis*.

arguments en faveur de la TI

(iii) les propositions d'identité ont un pouvoir explicatif que n'ont pas les simples corrélations (Block & Stalnaker, 1999)

analogie: pourquoi chauffer l'eau cause son ébullition ?
explication par les corrélations:
(c_1) la chaleur est corrélée à l'énergie cinétique moléculaire
(c_2) l'ébullition est corrélée à un certain mouvement moléculaire
augmenter l'énergie cinétique moléculaire cause un certain mouvement moléculaire

---

chauffer est corrélé à quelque chose qui cause quelque chose qui est corrélé à l'ébullition

$\Rightarrow$ pas une vraie explication de l'explanandum

arguments en faveur de la TI

explication par l'identité:

(i_1) la chaleur est l'énergie cinétique moléculaire

(i_2) l'ébullition est un certain mouvement moléculaire

augmenter l'énergie cinétique moléculaire cause un certain mouvement moléculaire

---

chauffer l'eau cause son ébullition

- de la même façon, les identités psychophysiques permettent d'expliquer pourquoi un événement mental (ex: douleur) peut causer un autre événement mental (ex: affliction), ce que ne parviennent pas à faire de simples corrélations. Il faut les accepter par inférence à la meilleure explication (IME).
- objection (Kim, 2006/2008): ce sont les lois neurophysiologiques qui font le travail explicatif, pas les énoncés d'identité. On ne peut donc invoquer (IME).

arguments en faveur de la TI

(iv) si l'on accepte la TI, le problème de la causalité mentale, si difficile à résoudre pour le dualiste, semble réglé

3. objections à la théorie de l'identité

objections à la TI

- 2 (familles d') objections principales à la TI:

(1) les objections basées sur l'indiscernabilité des identiques (??)

(2) l'objection de la réalisabilité multiple (???)

l'indiscernabilité des identiques

- Principe d'Indiscernabilité des Identiques (PII): si deux entités sont identiques, alors elles ont les mêmes propriétés
- à distinguer de la réciproque, le principe (beaucoup plus contestable et beaucoup plus contesté) d'identité des indiscernables)
- on peut construire plusieurs objections à la TI à partir du (PII): il suffit de mettre en avant des propriétés qu'auraient les états mentaux mais pas les états cérébraux (ou inversement, qu'auraient les états cérébraux, mais pas les états mentaux)

l'indiscernabilité des identiques, suite

v1 états mentaux et états cérébraux ne sont pas connaissables de la même manière:

- Pierre sait (immédiatement) qu'il ressent de la douleur
- Pierre ne sait pas que ses fibres de type *C* sont activées
- ▶ réponse: si quelque chose est nommé ou décrit de deux modes, et si l'on ne sait pas que ces deux modes renvoient à la même chose, notre savoir peut être sensible à cette différence de mode
- Pierre sait que Molière est l'auteur du *Misanthrope*
- Pierre ne sait pas que Jean-Baptiste Poquelin est l'auteur du *Misanthrope*

l'indiscernabilité des identiques, suite

v2 la croyance que *P* est sémantiquement évaluable: elle est vraie ou fausse selon que *P* est vraie ou fausse. Mais quel sens cela a-t-il de parler de la vérité ou de la fausseté de l'activation de la zone *Z* du cortex ?

- réponse: ce n'est pas l'état mental qui est vrai ou faux, mais son contenu. Pourquoi un état cérébral n'aurait-il pas un contenu, lui aussi ?

l'indiscernabilité des identiques, suite

v3 on peut qualifier une douleur d'aigüe; mais cela semble très étrange de qualifier un état cérébral d'aigü ?

- réponse: il ne faut pas concevoir "aigü" comme exprimant une propriété d'un état mental.Une douleur aigüe est un type d'état mental, lequel est identique à un type d'état cérébral.

la réalisabilité multiple

- l'argument de la réalisabilité multiple est l'argument qui, historiquement, a eu l'effet le plus dévastateur sur la TI
- la TI aboutit à une forme de chauvinisme du mental: pour avoir l'état mental M, il faut (et il suffit) d'avoir l'état cérébral C_M. Pour ressentir une douleur, il faut avoir une activation des fibres C. Ce corollaire a des conséquences contre-intuitives.
- Il y a des états mentaux (ex : la douleur) qu'on attribue volontiers à des animaux dont la structure neurologique est passablement différente de la nôtre et où il y a peu de chances qu'il y ait une activations de neurones C dans les situations où on leur attribue une expérience de douleur.

la réalisabilité multiple, suite

- on peut concevoir que les Martiens aient une vie mentale analogue à la nôtre mais une constitue chimique radicalement différente : à notre sytème nerveux central correspond chez eux une machinerie complexe en silicone.
- N'y a-t-il pas certains états mentaux que l'on a envie d'attribuer à des artefacts comme les robots ? ou que l'on aura envie d'attribuer dans un futur plus ou moins proche à des artefacts plus perfectionnés que ceux que nous fabriquons aujourd'hui ?

la réalisabilité multiple, suite

- Putnam, H. (1975)

  “Si (...) nous pouvons trouver un seul prédicat psychologique qui peut être clairement appliqué à la fois à un mammifère et à une pieuvre (par exemple "avoir faim"), mais dont le corrélat physico-chimique est différent dans les deux cas, la théorie de l’état cérébral s’effondre. Il me paraît plus que probable que nous en soyons capables.”

la réalisabilité multiple, suite

- certains partisans de la TI font valoir des "identités restreintes", (Braddon-Mitchell & Jackson). Supposons par exemple que la douleur corresponde à l'activation des fibres *C* chez l'homme, de neurones *D* chez le cobra, etc. On aurait alors un ensemble d'identifications :
- douleur chez l'homme = activation des fibres *C*
- douleur chez le cobra = activation des fibres *D*
- douleur chez le dauphin = activation des fibres *E*
- ...

- aux identités restreintes correspondraient des réductions locales des propriétés mentales aux propriétés physiques.

références

- Esfeld, M. (2005) *La philosophie de l'esprit*, Paris: A. Colin, chap. 6
- Kim, J. (2006/2008), *Philosophie de l'esprit*, Paris: Ithaque, chap. 4
- Block & Stalnaker (1999) "Conceptual Analysis, Dualism and the Explanatory Gap", *Philosopical Review*, 108(1), pp. 1-46
- Fodor, J. (1974/1980) "Les sciences particulières (l'absence d'unité de la science: une hypothèse de travail)", trad.fr. in Jacob (ed.) *De Vienne à Cambridge*, Paris: Gallimard
- Putnam, H. (1975), "The Nature of Mental States", dans *Mind, Language, and Reality : Philosophical Papers*, vol. 2, Cambridge UP, 1975 ; trad.fr. "La nature des états mentaux" dans D. Fisette et P. Poirier (ed.), *Philosophie de l'esprit*, vol. 1, Vrin, 2002
- Smart, J.J. (1959) "Sensations and Brain Processes", *Philosophical Review*, 58, pp. 141-156
- Stoljar, D. (2009)"Physicalism", The Stanford Encyclopedia of

pour repartir du bon pied...

- qu'est-ce que le physicalisme (ou matérialisme) ?
- qu'est-ce que la théorie de l'identité ?
- quel genre d'identité sert d'analogie à la théorie de l'identité ?
- qu'est-ce que la différence type/exemplaire ?
- comment fonctionnent les objections basées sur le principe des indiscernables ?