

Quelle méthode pédagogique utiliser pour rendre un cours facile à comprendre et à mémoriser ? Cette question a fait couler beaucoup d'encre : beaucoup de réponses ont été formulées, certaines mauvaises, d'autres meilleures.

A l'heure actuelle, le débat sur les méthodes pédagogiques est essentiellement une confrontation entre deux camps : les pédagogies actives, et la pédagogie traditionnelle.

La **pédagogie traditionnelle** est celle à laquelle vous avez normalement eu droit quand vous étiez à l'école. Une heure de cours est composée d'un cours magistral qui « transmet" du savoir », suivie d'une phase de pratique autonome.

Le second camp est celui des **pédagogies actives**, dans lesquelles le professeur ne doit pas faire trop de cours magistraux, mais doit aider l'enfant à redécouvrir par lui-même le savoir à inculquer. Cette pédagogie prend la motivation des élèves très au sérieux et considère qu'on apprend en faisant, par séries d'essais et d'erreurs, par tâtonnements expérimentaux.

Un troisième camp est celui des **techniques behavioristes**, basées sur le concept du conditionnement Pavlovien, et le conditionnement opérant de Skinner.

On pourrait croire que ces pédagogies actives sont motivantes pour les élèves, mais le projet d'étude *Follow through*, réalisé aux états-unis a réduit ce mythe à néant. Ce projet a évalué 9 méthodes pédagogiques, sur une durée de 10 ans, avec 352 000 élèves, dans 180 écoles. Dans cette étude, une méthode dépassait les autres sur tous les points : connaissances académiques, savoir-faire, et motivation/estime de soi. Cette méthode pédagogique était la **direct instruction**.

Dans les années 1970, B. Rosenshine, un psychologue cognitif a utilisé les résultats du projet *Follow through* et effectué des observations dans les classes pour créer une nouvelle méthode pédagogique : la pédagogie explicite. Celle-ci incorpore les techniques de la pédagogie behavioriste, avec quelques ajouts.

Depuis, les confirmations expérimentales sont arrivées. L'expérience de City springs (Baltimore), et l'expérience du Winsconsin ont été autant de confirmations expérimentales supplémentaires. Bilan : la pédagogie explicite est clairement supérieure à toutes les autres approches, aussi bien actives que traditionnelles.

Par la suite, les découvertes sur le fonctionnement de la mémoire ont permis d'expliquer une partie des observations de Rosenshine, et d'en affiner les conseils. La **théorie de la charge cognitive**, ainsi que les travaux de Sweller et de ses collègues, ont notamment révolutionné cette pédagogie.

En plus de cette pédagogie, les chercheurs en psychologie cognitive ont découvert quelques techniques assez importantes, basées sur **l'organisation et le fonctionnement de la mémoire**. On peut notamment citer les techniques utilisées dans la méthode multi-épisode d'Alain Lieury, un chercheur assez connu pour ses ouvrages de vulgarisation sur la mémoire.

D'autres travaux plus récents ont une utilité pédagogique dans certains apprentissages spécifiques : on peut citer les travaux de Stanislas Dehaene sur l'acquisition de la numération, de la lecture, et de l'écriture.

### Plan du cours

Dans cet article, nous allons parler de l'ensemble des techniques pédagogiques tirées de la psychologie cognitive et des neuro-sciences. Nous n'allons pas nous focaliser sur une pédagogie particulière, mais aborder le sujet d'une manière assez large, en mentionnant tous les conseils établis à l'heure actuelle.

Nous allons d'abord parler du fonctionnement de la mémoire et de la cognition humaine. Savoir comment on apprend passe forcément par une étude des mécanismes d'apprentissage, afin d'en tirer partie. Nous verrons que l'être humain dispose de plusieurs types de mémoire, ainsi que la manière dont son structurées les connaissances en mémoire.

Ensuite, nous verrons quelles sont les différences entre experts et novices dans un domaine. Cela nous permettra de savoir ce qui différencie experts et novices, et donc, bons et mauvais élèves. De l'étude de ces différences, nous en déduirons sur quel paramètres jouer pour rendre nos apprentissages plus efficaces.

Ensuite, nous rentrerons dans le cœur du sujet, et donnerons des techniques clé en main, et des recommandations, qui vous permettront de faire des cours clairs, faciles à comprendre, et qui feront rapidement progresser vos lecteurs et/ou élèves.

# Introduction : la mémoire

Pour commencer, nous allons faire un petit rappel sur la mémoire. Enfin, je devrais plutôt dire les mémoires, car un humain en possède plusieurs.

## Mémoires à court terme

Les **mémoires à court terme** peuvent stocker temporairement des informations et s'effacent en quelques dizaines de secondes. C'est ces mémoires qui servent à mémoriser un numéro de téléphone avant de le composer, ou qui servent à retenir temporairement ce que dit le professeur. Elles sont aussi utiles pour stocker les phrases que le professeur vient tout juste de prononcer.

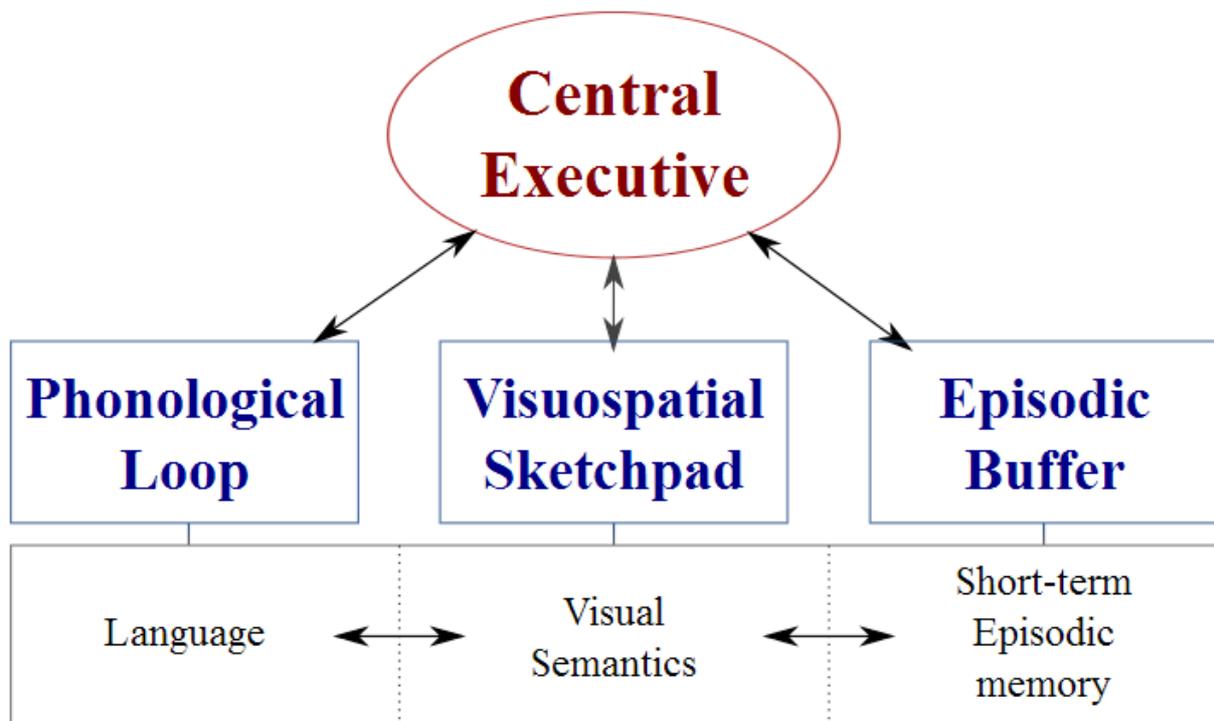
Dans ce qui va suivre, j'appellerai ces Mémoires à Court Terme : MCT.

Elles sont au nombre de 3 :

- une mémoire visuelle : le calepin visuospatial ,
- une mémoire verbale : la boucle phonologique ;
- et une mémoire fourre-tout, impliquée dans la compréhension et la création d'associations, ainsi que dans la planification : l'episodic buffer.

Ces mémoires à court terme ne sont pas que des mémoires : elles sont aussi couplée à un processeur de traitement, le superviseur attentionnel. Celui-ci est un centre de traitement chargé de la gestion des mémoires à court terme. Il est chargé de :

- décider ce qui rentre dans les diverses mémoires à court terme ;
- d'échanger des informations entre les mémoires ;
- de faire passer des informations en mémoire à long terme ;
- de combiner le contenu des diverses mémoires, lors des situations de compréhension et de résolution de problèmes.



Ce superviseur attentionnel est une sorte de petite machine, qui fonctionne avec un carburant : **l'attention**.

Lorsqu'un élève a ses mémoires totalement remplies, ou encore en situation de surcharge, ce superviseur fonctionne

n'importe comment et a du mal à gérer la situation.

De plus, il ne peut utiliser qu'une quantité finie d'attention lors d'un court laps de temps : il ne peut pas utiliser tout son stock durant 2 à 3 secondes, même si on se concentre très fort.

## Mémoire à long terme sémantique

On trouve d'abord différentes **mémoires à long terme**, qui servent à stocker les informations de manière assez durable.

Il existe différents types de mémoire à long terme, mais l'une d'entre elle va nous intéresser plus que les autres : celle spécialisée dans les connaissances, la mémoire sémantique. La psychologie de l'éducation s'intéresse à la mémoire à long terme, et notamment aux mécanismes d'oubli et d'entrée dans cette mémoire.

Vos connaissances sont stockées dans la **mémoire sémantique**, qui stocke des connaissances, reliées entre elles par des associations. Ces associations sont de différents types : associations verbales entre deux mots ; associations visuelles, ou associations d'idées créées lorsqu'on comprend quelque chose.

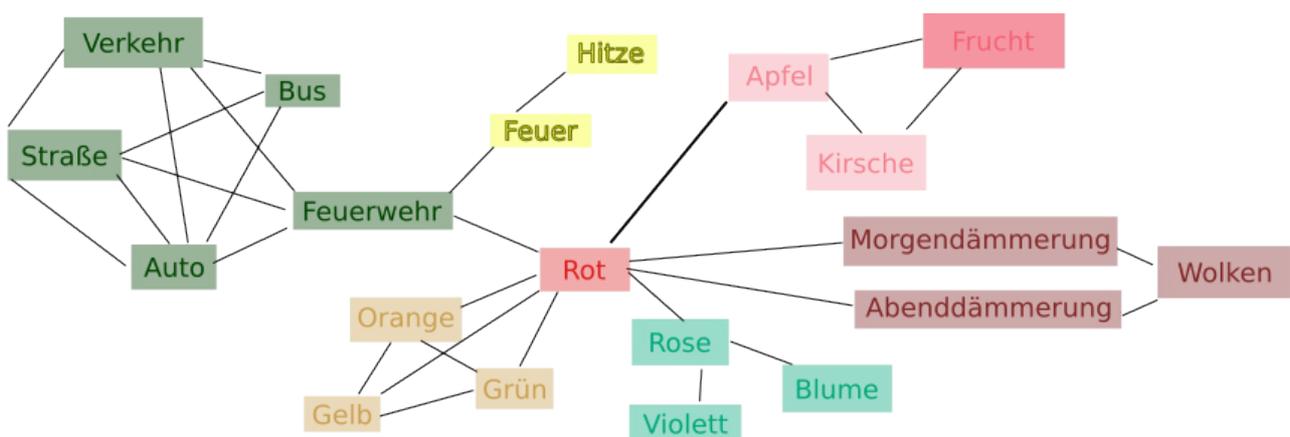
Si l'on en croit les expériences effectuées par les neuro-scientifiques, ainsi que les observations sur des patients ayant des lésions au cerveau, cette mémoire sémantique stocke des **catégories** d'objets, des concepts.

Chaque concept est stocké dans le cerveau par un ensemble d'informations élémentaires, que l'on appelle des **traits sémantiques**. Ces traits sémantiques sont des petits morceaux d'informations, qui peuvent être des informations sensorielles (images, sons, odeurs, etc), que des informations plus évoluées. Par exemple, le concept canari contiendra des traits sémantiques sensoriels : couleur jaune ; et des informations plus élaborées : peut voler.

Les traits sémantiques sensoriels sont stockés dans les aires du cerveau concernées : les odeurs dans l'aire du cerveau spécialisée dans la mémorisation des odeurs, les images dans le cortex visuel, les sons dans le cortex auditif, etc. Au final, tous ces traits sémantiques sont dispersés dans tout le cortex cérébral. Les catégories sont formées par un point de convergence de ces traits sémantiques, qui sont reliés entre eux dans une région du cortex : on soupçonne que la zone de convergence serait le cortex temporal, situé sous la tempe.

De manière générale, les petits enfants construisent les catégories dans leur cerveau en commençant par assembler entre eux des informations sensorielles, visuelles, sonores, etc. Au fur et à mesure du développement, les catégories contiennent de plus en plus de traits sémantiques abstraits.

La mémoire sémantique est intégralement constituée de concepts, reliés entre eux par des associations. L'ensemble forme ce qu'on appelle un **réseau sémantique**.



Collins & Loftus (1975): Spreading Activation Network

Ces associations peuvent être de plusieurs types. Dans notre cerveau, il existe plusieurs types d'associations :

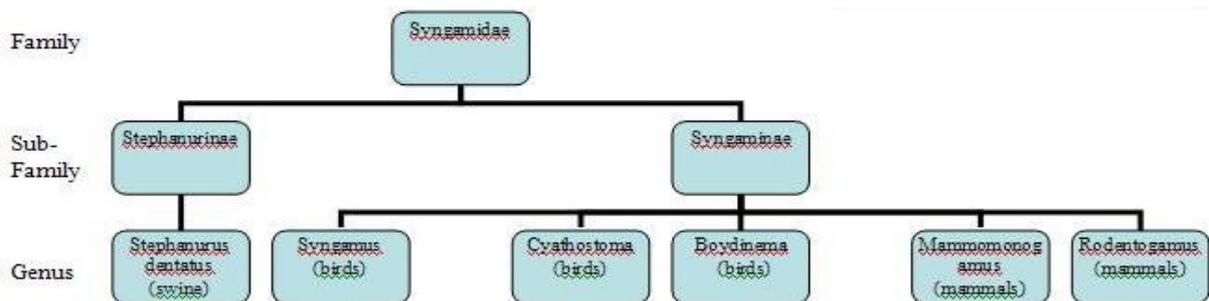
- les associations catégorielles ;
- les associations de contiguïté ;
- les associations de causalité ;
- les associations de ressemblance.

# Taxonomies

D'après les **théories hiérarchiques de classes**, les catégories sont reliées entre elles par ce qu'on appelle des **taxonomies**.

Dans celles-ci, si une catégorie est une sous-catégorie d'une autre, la sous-catégorie est appelée la catégorie fille, et l'autre la catégorie mère. Dans une taxonomie, une catégorie est reliée à une autre si elle est une catégorie-fille ou une catégorie-mère.

Quelques observations semblent supporter cette théorie : toutes les cultures utilisent des hiérarchies de catégories de ce genre. Mieux : ces taxonomies sont presque identiques à travers les cultures. Des comparaisons ont été faites sur les taxonomies d'animaux sauvages : entre les équipes de taxonomistes envoyés dans la jungle, et les taxonomies des cultures locales, il n'y a pas beaucoup de différences.



Les hiérarchies, les taxonomies, peuvent avoir un nombre indéfini de niveaux, regroupés en trois grands sous-ensembles, trois « niveaux d'abstraction » :

- le **niveau superordonné**, qui correspond aux catégories les plus générales ;
- le **niveau basique**, qui correspond aux sous-catégories du niveau superordonné ;
- et un **niveau sous-ordonné**, qui contient des sous-catégories des catégories du niveau basique.

Plus on remonte dans la hiérarchie, plus le niveau devient sur-ordonné, et moins il devient sur-ordonné. Les niveaux basiques étant alors autant sur que sous ordonnés.

## Distance sémantique

Quelques expériences semblent toutefois montrer que la hiérarchie taxonomique n'est pas toujours respectée. Ces expériences sont des expériences de **chronométrie mentale**. Le principe est simple : on donne une phrase à un cobaye, et il doit dire si celle-ci est vraie ou fausse. L'expérimentateur mesure le temps mis par le cobaye pour répondre : plus le temps mis à répondre est long, plus le cerveau a du parcourir de nœuds dans la mémoire sémantique pour tomber sur le bon.

De ces expériences, il ressort quelques observations : plus un concept est familier, plus on peut vérifier les phrases qui y rapportent rapidement, sans que la position dans la hiérarchie taxonomique n'y change quoique ce soit. De plus, la hiérarchie est violée dans certains cas spécifiques. Par exemple, il est plus lent de vérifier la véracité de la phrase "un chien est un mammifère", que pour la phrase "un chien est un animal", alors que les théories hiérarchiques prédisent l'inverse.

De nombreux phénomènes d'**inversion de catégorie** similaires ont été documentés et seraient assez fréquents : ils montrent bien que la hiérarchie taxonomique n'est pas toujours respectée, et est au mieux approximative : d'autres types d'associations peuvent exister.

Les théories actuelles ne se contentent pas des relations taxonomiques : de nombreuses associations viendraient se surajouter aux taxonomies. Elles supposent l'existence d'**associations thématiques**, qui se forment suivant la **distance sémantique** entre deux concepts.

Pour les nouvelles théories, les concepts sont associés comme dans les modèles hiérarchiques, mais aussi suivant d'autres critères. Les concepts sont ainsi associés s'ils sont ressemblant, s'ils ont la même utilité, s'ils ont des points

communs, s'ils peuvent faire les mêmes choses, s'ils ont la même fonctionnalité, s'ils se rencontrent en même temps, s'ils s'utilisent l'un après l'autre, etc.

Selon les expériences d'association libre, ces relations sont de plusieurs types :

- les associations de contiguïté ;
- les associations de causalité ;
- les associations de ressemblance.

Deux concepts sont reliés par une association de contiguïté dans deux cas.

Avec la contiguïté spatiale, deux concepts sont associés quand ils sont proches l'un de l'autre. Par exemple, deux pays proches sur une carte seront associés. Par exemple, si je prononce le mot "France", et que je vous demande de penser à un nom de pays, vous avez de fortes chances de me sortir "Italie" ou "Allemagne", comparé à "Namibie" ou "Australie".

Avec la contiguïté temporelle, les deux sont associés car ils apparaissent souvent en même temps. Par exemple, médecin sera plus associé à "hôpital" qu'à diplôme, vu qu'on voit plus souvent de médecin quand on va à l'hôpital que quand on est face à un diplôme.

Les associations de causalité relie deux concepts ou événements, quand l'un est la cause de l'autre. Ces relations permettent d'expliquer qui cause quoi. Ces relations apparaissent naturellement quand on on explique pourquoi les choses sont ce qu'elles sont. Et cela manque dans beaucoup de cours...

Les relations de ressemblances sont de deux types. Les relations analogiques vont relier deux concepts qui ont des points communs, des points de ressemblances. Un professeur peut les utiliser en effectuant des analogies. Les relations d'oppositions vont relier deux concepts opposés : chaud/froid, etc. Ces relations sont rarement utilisées en classe, et mériteraient à l'être.

## Schémas

L'ensemble de ce réseau sémantique est structuré autour de **schémas**, des portions du réseau sémantique qui stockent une idée générale, sur laquelle viennent se coller des détails.

Un schéma regroupe tout une classe de catégories, de problèmes, d'événements, d'actions, etc ; et stocke une structure abstraite, partagée par de nombreuses catégories, de nombreux problèmes, de nombreux événements, etc. Seuls les points de communs sont conservés dans ce schéma, même si des informations contextuelles peuvent se recoller suivant les besoins.

En somme, ces schémas stockent des invariants, des généralités, le noyau dur d'un concept.

Prenons un exemple, pour montrer cet état de fait : les expériences de Bartlett. Celui-ci a travaillé sur l'apprentissage de matériel relativement complexe : cartes, schémas, textes, histoires, etc.

Une des expériences les plus marquante fut celle sur **l'effet intégrateur du titre**. Dans son expérience, il constituait trois groupes de cobayes, qui devaient tous lire un même texte assez complexe à comprendre. Le premier groupe ne connaissait pas le titre du texte. Le second connaissait le titre après avoir lu le texte, tandis que le troisième avait accès au titre avant la lecture. Le résultat : le groupe qui avait eu accès au titre avant le texte réussissait mieux que les deux autres groupes, qui avaient des performances à peu-prêt comparables. Ceci dit, mettre le titre après la fin du texte avait quand même un petit effet.

Cela vient du fait que le titre donne l'idée générale du texte, et fournit quelques informations de base, sur lesquelles les informations du texte vont venir se coller, s'associer progressivement. Cette information générale fournie par le texte permettait de structurer les informations du texte dans un schéma en mémoire sémantique, facilitant le rappel et la compréhension. Le fait que le titre doivent être placé avant le texte pour avoir un effet maximal le montre : il n'y a pas beaucoup d'associations retardées, à rebours, même s'il y en a.

Autre exemple : on a tous une représentation abstraite des différentes histoires qu'on a entendu ou lu dans notre vie : ce **schéma d'histoire** serait la base de la compréhension de textes. Expérimentalement, si on donne à des cobayes des histoires qui ne suivent pas ce schéma abstrait (une histoire d'une autre culture, une histoire conçue en vue de ne pas coller à ce schéma, etc), les cobayes se rappellent très bien des détails qui collent au schéma, et pas les autres / Bartlett (1932). Autant vous dire que la recherche sur les schémas d'histoire pourrait être utile aux professeurs de français...

Pour en donner un exemple, on peut citer une des expériences de Bartellet. Dans son expérience la plus connue, il a demandé à des cobayes de lire un texte, qui narrait une légende indienne. Dans cette légende, une bonne partie des détails de l'histoire sont relativement étranges pour un occidental, et sont assez exotiques, pas vraiment familiers. Or, Bartellet a constaté divers phénomènes dans diverses épreuves de rappel et de compréhension du texte :

- Les détails étranges du texte étaient souvent oubliés ;
- quand ils n'étaient pas oubliés, ces détails étaient rationalisés, et les cobayes étaient certains que leur version, plus « cohérente » culturellement parlant, était celle lue dans le texte ;
- les détails cohérents culturellement, familiers, étaient parfaitement retenus ;

Dans cette expérience, on voit que quelque chose, un schéma d'histoire en l'occurrence, a favorisé la mémorisation des détails familiers, et inhibé les détails bizarres, non-familiers.

On peut aussi voir certains de ces schémas comme un ensemble de cases reliées entre elles. Une partie de ces cases contient des invariants, et leur contenu est fixe, obligatoire. D'autres cases sont vides et peuvent être remplies par des informations contextuelles, qui proviennent d'une situation particulière.

Exemple, si jamais je fais face à cette équation :  $5y^2 + 27y + 3$ , je sais qu'il s'agit d'une équation du second degré : je peux rattacher cette équation à un schéma qui stocke tout ce qu'il y a à savoir sur les équations du second degré, et notamment leur « forme ». Les informations particulières, qui remplissent les cases vides, sont simplement le nom de l'inconnue, et les coefficients 5, 27, et 3.

Si les nouvelles informations ne s'encastrent pas bien dans un schéma, on aura tendance à la tordre pour la faire rentrer, en oubliant certains détails, en en modifiant d'autres, etc.

Par exemple, refiliez une liste de symptômes à apprendre à un médecin. Débrouillez-vous pour que cette liste de symptômes ne corresponde à aucune maladie connue. En même temps, faites en sorte qu'en enlevant 2 symptômes, la liste corresponde à une maladie connue du médecin. Quelques semaines plus tard, le médecin se rappellera uniquement des symptômes de la maladie courante, pas des deux autres.

Dans certains cas, cela peut aller plus loin : on peut se rappeler de choses qui n'étaient pas dans le matériel original à apprendre. Par exemple, donne une liste de symptômes à apprendre à un médecin, et fait en sorte que cette suite soit celle d'une maladie rare, mais à laquelle on a enlevé un symptôme crucial. Deux semaines plus tard, le médecin sera certain que le symptôme manquant faisait partie de la liste.

Petit détail : ces schémas ont quelques points communs avec les schémas de Piaget. Mais il s'en démarquent sur de nombreux points. Déjà, ils ne sont pas organisés selon une organisation style Moyen->but. A la place, il s'agit plus d'une organisation abstraite, spécialisée pour la reconnaissance de situation familière, qui stockent des abstractions générales.

Deuxièmement, Piaget supposait que les enfants pouvaient *accommoder* les schémas, les modifier quand ils ne sont pas adaptés à une situation. Mais avec les schémas, c'est l'inverse : ce sont les informations qui sont tordues ou omises, histoire de coller le plus possible au schéma choisit. Les schémas se modifient rarement, et ont tendance à grossir, se voir accoler des détails, mais sans possibilité de modifications.

## Utilité

Ces réseaux sémantiques et ces schémas nous permettent d'interpréter l'environnement en fonction de situations déjà connues. Grosso-modo, ils permettent de reconnaître des situations familières : si une situation colle à un schéma présent en mémoire, le schéma peut être réutilisé.

Ils permettent ainsi d'inférer : si une situation ou des informations collent partiellement à un réseau sémantiques ou un schéma, cela veut dire qu'elles sont compatibles avec : le cerveau peut réutiliser les informations dans le réseau sémantiques ou le schéma pour faire des inférences.

Mais ces réseaux sémantiques ou schémas, si utiles pour raisonner et réutiliser les connaissances apprises, peuvent aussi foutre un joyeux bordel ! Si un schéma est réutilisé alors qu'il ne le devrait pas, c'est le drame : le schéma est inadapté, les inférences faites sont totalement foireuses, et les connaissances réutilisées sont alors inadaptées à la situation. Il faudra faire attention à ce genre de situation quand on crée un cours.

De même, les nouvelles connaissances, celles qui seront à apprendre, vont parfois devoir s'intégrer dans un réseaux

sémantiques ou un schéma pré-existant. L'intégration peut mal se passer : la compréhension est mauvaise, de même que la mémorisation.

On retient mieux ce qui est familier, ce qui colle avec nos attentes, avec notre culture. La morale, c'est que plus de nouvelles informations sont cohérentes avec ce que l'on sait déjà, mieux on les retient. On retient mal ce qui est bizarre, étrange, pas naturel, non-intuitif, incohérent avec nos connaissances antérieures.

## Activation diffusante

Pour résumer, la mémoire sémantique est bien modélisée par des modèles connexionnistes à base de réseaux sémantiques et de schémas. Dans ces modèles, les concepts et catégories sont stockées par un réseau de neurone (qui peut être distribué dans tout le néo-cortex). Ces réseaux de neurones sont reliés entre eux par des synapses, qui forment des associations.

Le réseau de neurones qui stocke un concept peut émettre des influx nerveux : on dit que le concept est "activé". Plus l'activation est forte, plus le concept reçoit et produit d'influx nerveux, et plus il a de chance d'être rappelé et utilisable dans un raisonnement.

Mais cette activation ne reste pas en place : elle va se propager à travers le réseau sémantique, en passant par les associations, les synapses. Celles-ci peuvent alors propager plus ou moins bien l'activation.

Il arrive que par endroit, un concept reçoive de l'activation en provenance de plusieurs associations. Les activations convergentes s'additionnent. En somme, certains concepts servent de **points de convergence** dans le réseau sémantique.

La quantité d'activation reçue par un concept détermine sa facilité à être rappelé. Le concept est rappelé si l'activation est suffisamment importante. De plus, le concept possède une certaine force, qui définit si il est plus ou moins facile à rappeler, à activation égale.

Les recherches actuelles en psychologie et en neurosciences font penser que la compréhension, le raisonnement, la résolution de problème, l'intelligence et la créativité sont basés sur la mémoire sémantique, et plus précisément sur l'activation diffusante.

Quand on raisonne, les prémisses s'activent, et cette activation va progressivement converger vers la conclusion. Quand on résout un problème, les informations du problème activent des concepts, et l'activation diffuse jusqu'à la solution. La pensée créative découle du même phénomène, à part qu'elle crée ou parcourt des associations entre concepts distants, ayant peu de ressemblances.

Conséquence : ***l'intelligence, la créativité, et la capacité à résoudre des problèmes se basent sur la mémoire.***

Même de rien, la structure du réseau sémantique a une influence sur la propagation de l'activation diffusante : elle détermine la disponibilité d'un concept. Plus un concept est accessible par un grand nombre d'associations dans un réseau sémantique, plus on peut s'en rappeler facilement ou l'utiliser dans des raisonnements ou pour résoudre des problèmes. Et cela a des implications pédagogiques qu'on abordera plus tard.

Les points de départ de l'activation dans le cerveau sont des informations qui nous parviennent de l'extérieur. Tout texte, situation, énoncé, problème, contient des **indices de récupération**.

Dans la majorité des cas, les indices de récupération sont des mots, ou des images. Une fois que ces mots ou ces images arrivent dans la mémoire de travail ou la mémoire sensorielle, ils vont aller activer les concepts adéquats en mémoire : ils servent de porte d'entrée à la mémoire.

Conséquence : un professeur doit donner des indices de récupération les plus pertinents possibles : il utilise un langage clair, des mots connus des élèves, fournit des schémas clairs et lisibles, etc. De plus, il n'utilise pas d'expressions vagues ou ambiguës du style : "une sorte de...", "un espèce de...", etc.

Cette notion d'indice de récupération permet d'expliquer la différence de performances entre rappel et reconnaissance. Pour donner un exemple, prenons deux groupes de cobayes auxquels on fait apprendre une liste de mots. Le premier groupe devra se rappeler les mots de la liste lors d'une interrogation, tandis que l'autre devra reconnaître les mots de la première liste insérés dans une nouvelle liste.

Le rappel est plus difficile que la reconnaissance. En terme de pédagogie, cela a une petite importance. Les QCMs sont des exercices de reconnaissance, et non de rappel. Ils sont donc beaucoup plus faciles.

## Différences expert-novices

Après ce petit rappel sur la mémoire, il est temps de voir en quoi cela peut nous permettre de créer des pédagogies efficaces. Il faut d'abord savoir qu'il y a des différences entre un bon élève et un mauvais élève, qui proviennent toutes des différentes mémoires vues au-dessus.

Dans ce qui suivra, nous verrons comment exploiter ces différences pour faire progresser les élèves. En s'inspirant ce que l'on trouve dans la mémoire des bons élèves, on saura quoi faire pour adapter notre enseignement.

## Mémoire à court terme

Premièrement : un mauvais élève a souvent une mémoire de travail assez faible, qui peut contenir peu de choses. Cette MCT a une grande influence sur les performances des élèves : diverses études ont montré que la capacité de la MCT à 5 ans est un bon indicateur de réussite scolaire. De plus, on observe une corrélation entre faible capacité de la MCT et échec scolaire. Cette MCT a surtout une grande influence dans l'apprentissage de la lecture, de l'écriture, et de la numération : à cet âge-là, les enfants ont une MCT contenant peu d'informations.

La quasi-totalité des problèmes de mémorisation ou de compréhension sont dus à la gestion des MCTs par le professeur, qui a tendance à surcharger la MCT. Or, pour mémoriser une connaissance, on est obligé de passer par la MCT. Plus elle est remplie, moins la compréhension et la mémorisation sont bonnes. Si elle sature, c'est encore pire : les performances s'effondrent, l'élève fait des erreurs, etc. Le rôle du professeur est d'éviter que la mémoire à court terme se remplisse : il doit diminuer la **charge cognitive**, c'est à dire le nombre d'informations présentes simultanément dans la MCT.

La pédagogie traditionnelle ne prend pas en compte cette charge cognitive. Mais de ce point de vue, les pires pédagogies sont les pédagogies actives qui utilisent des situations à forte charge cognitive : gérer un projet n'est pas simple, résoudre des problèmes non plus. La MCT des élèves est alors totalement remplie et le bilan est catastrophique : une grande partie de l'attention est utilisée pour résoudre le problème ou gérer le projet, et n'est pas utilisée pour la mémorisation ou la compréhension.

Pour diminuer l'occupation de la MCT, le professeur peut utiliser la **théorie de la charge cognitive**, une théorie qui recense plusieurs effets ayant chacun des implications pédagogiques. Pour être franc, cette théorie est une vraie arme d'apprentissage massive, sur laquelle toutes les pédagogies tirées de la psychologie cognitive se basent, sans exception.

## Réseaux sémantiques

Ensuite, un autre facteur apparaît : la faible quantité de connaissance en mémoire sémantique.

Cela a été plus ou moins établi par les études d'Alain Lieury, datées de 1996). Pour en donner une illustration, on peut dire que la corrélation entre moyenne scolaire et connaissances encyclopédiques en classe de 5ème sont de 0.72. Et la corrélation est la même avec le taux de redoublement 4 ans plus tard. Ce qui est plus élevé que les corrélations avec les tests de QI ou de raisonnement (0.50).

Les recherches sur l'expertise sont claires : toutes les différences entre un expert du domaine et un novice sont dues à la grande quantité de schémas et de connaissances des experts, ainsi qu'à la structure de leurs réseaux sémantiques.

Un expert a vu un grand nombre situations et de problèmes bien précis et peut automatiquement reconnaître des problèmes familiers, et réutiliser les solutions qu'il a déjà apporté à ces problèmes. Là où un novice devra gérer une quantité d'informations phénoménales, l'expert reconnaîtra des situations déjà vues et les liera avec un schéma bien particulier en mémoire. Il pourra alors réutiliser ses connaissances antérieures et regrouper les informations du problèmes en chunks assez gros.

Un expert n'aura donc pas besoin de maintenir grand chose en MCT, ni même à réfléchir autant que le novice : il aura juste à charger deux schémas en MCT et à faire quelques associations simplistes pour trouver le résultat. Et ce phénomène suffit à expliquer toutes les différences entre experts et novices.

Et ce n'est pas spécifique à l'éducation, ce principe s'adapte aussi à des choses assez diverses : lecture et compréhension de texte, conception de circuits électroniques, programmation, algèbre, géométrie, diagnostic médical, etc.

Cela a notamment été montré sur des joueurs d'échec. Si on les met face à une configuration de jeu courante, ils sont capables de retenir un plus grand nombre de pièces que les débutants, et leur performances en cours de jeu sont

supérieures. Mais face à des configurations de jeu qui n'apparaissent presque jamais dans une vraie partie, leurs performances sont strictement identiques à celles des débutants.

## Schémas

Dans les études sur l'expertise, mentionnées plus haut, un dernier facteur est apparu : l'organisation des connaissances dans le réseau sémantique. Celle-ci est différente chez les experts : leurs réseaux sémantiques contiennent un plus grand nombre d'associations, et celles-ci sont plus pertinentes. Les experts ont plus tendance à associer des concepts sur la base de principes ou de connaissances abstraites, sémantiques, que sur leurs caractéristiques de surface, perceptives.

On peut résumer cela en disant que les experts ont structuré leurs connaissances et ont très bien comprises celles-ci. Un expert n'est pas qu'un simple répertoire de connaissances. Les vrais experts sont ceux qui ont une tête bien faite, et pas seulement une tête bien pleine.

On en déduit donc qu'un bon cours devra fournir une grande quantité de connaissances, et devra chercher à maximiser le nombre et la qualité des associations entre concepts. Le résultat sera un cours qui met l'accent sur la compréhension. Cette compréhension aura plusieurs définitions, et de nombreuses astuces permettent de faciliter celle-ci : les modèles hiérarchiques de la mémoire sémantique mettent en avant l'impact de la catégorisation sur le plan et les explications, et les divers types d'associations peuvent s'appliquer facilement dans un cours.

De plus, chez l'expert, les connaissances sont souvent organisées sous forme de schémas : elles sont structurées autour d'une grande idée générale, sur laquelle des idées un peu plus particulières viennent s'ajouter.

Et cela marche aussi bien pour les compétences des joueurs d'échec (travaux de de Groot), le calcul mental, que pour des domaines aussi divers que la conception de circuits électroniques, les mathématiques, l'économie, ou la programmation - voir page Wikipédia anglais sur l'expertise.

On en déduit qu'un bon cours devra être structuré autour d'une idée générale, donnée au début du cours, sur laquelle viendront s'agglomérer des détails : cela permettra de former des schémas en mémoire sémantique.

## Mémoire sémantique et apprentissage : l'importance de la compréhension et des connaissances antérieures

L'activation diffusante a son importance en pédagogie. Tout le challenge du professeur est de faire en sorte que l'activation arrive à bon port, qu'elle arrive à activer les bons concepts quand il faut. Pour cela, il dispose de deux solutions.

Les associations dans le réseau sémantique possèdent un poids, qui définit leur capacité à transmettre l'activation facilement. Ce poids peut évoluer au cours du temps. Typiquement, plus on répète quelque chose, plus les connexions qui mènent au matériel sont consolidées. On en déduit donc que la **répétition** semble être un bon moyen pour mémoriser. Mais c'est oublier que cette répétition a un faible effet, et demande du temps.

Mais ce n'est pas le seul : si on veut augmenter l'activation qui converge vers un concept, on peut aussi multiplier les points de convergence. Dans ce réseau, les associations sont très importantes : elles vont déterminer l'accessibilité d'une information en mémoire.

Plus un concept est relié à d'autres par des associations pertinentes, plus le travail de l'activation diffusante est facilité par la présence de nombreux points de convergences, qui servent d'indices de récupération. Plus on associe un concept à d'autres, mieux on le mémorise, et plus on structure la mémoire sémantique.

Les confirmations expérimentales proviennent des travaux sur la **profondeur de traitement**. Les expériences qui ont montré cet effet de la structuration du réseau sémantique et l'importance des indices de récupération sont nombreuses. Les premières expériences furent celles de Craik et Tulving (1973 et 1975). Elles ont été suivies d'une quantité assez impressionnante d'études sur le sujet.

Par exemple, on peut citer : l'expérience de Moreau (1973). Celle-ci a donné des listes de villes à apprendre à deux groupes de cobayes. Le premier groupe voyait les villes dans le désordre, tandis que le second groupe voyait ces villes en suivant le trajet d'un tour de France qui avait eu lieu dans le passé. Le second groupe avait un taux de rappel bien supérieur.

D'autres expériences faites sur des listes de dates ont montré qu'explicitement les relations de causalités entre événements, permet de mieux mémoriser la liste des dates associées.

Même chose pour les listes de mots : donner des informations sur les relations (souvent catégorielles) entre mots permet de mieux mémoriser les listes.

Bref, pour augmenter au maximum ce nombre d'associations, le professeur doit jouer sur la profondeur de traitement, la **qualité de compréhension**. Plus on comprend, plus on crée des associations avec le matériel à apprendre : comprendre, c'est faire des liens, relier.

Plus on peut relier de nouvelles informations à des informations déjà présentes en mémoire, mieux on comprend, et mieux on mémorise. Cela permet de structurer la mémoire sémantique et l'optimise fortement, ce que le par cœur ne peut pas faire. Un bon professeur doit considérer que la compréhension est le mécanisme de mémorisation par excellence, supérieur au par cœur.

Conséquence, le professeur doit mettre fortement l'accent sur la **compréhension** : c'est la technique de mémorisation par excellence, sur laquelle tout le cours est basé. On peut donc résumer avec le slogan suivant : ***il faut comprendre pour apprendre.***

## **Généralités sur le contenu d'un cours magistral**

D'après ce que l'on sait des réseaux sémantiques et des schémas, la compréhension et la mémorisation est fortement améliorée en associant et en intégrant de nouvelles connaissances dans un schéma ou un réseau sémantique pré-existant.

Le professeur peut utiliser ce principe en jouant sur l'ordre des notions dans un cours pour faciliter la compréhension des élèves. Un enseignant doit bâtir son cours et sa progression sur des connaissances antérieures, sur des schémas pré-établis. Ces connaissances antérieures peuvent avoir été abordées auparavant dans le cursus scolaire, ou dans le cours même.

Cette vision de la compréhension basée sur une association d'une nouvelle association à des connaissances antérieures prévoit l'apparition de phénomènes de **transferts positifs** : l'apprentissage préalable de certaines notions permet de faciliter des apprentissages ultérieurs. Cela a des implications sur la construction du plan, mais aussi sur l'ordre des cours dans l'année.

## **Quantité de connaissances**

Quand on associe une nouvelle information à un réseau sémantique, elle ne s'associe pas qu'avec les pré-requis : celle-ci s'associe à un maximum de connaissances antérieures avec lesquelles elle a un lien. Conséquence : plus on a de connaissances antérieures, plus on crée d'associations, et mieux on se souvient, mieux on comprend.

On en déduit une chose : plus quelqu'un possède de schémas ou de réseaux sémantiques fournis, plus les phénomènes de transfert positif vont apparaître. Plus le réseau sémantique est riche, plus il sera facile à un élève d'associer une nouvelle information au réseau préexistant :

- l'élève peut réutiliser une bonne partie de ce qu'il sait pour comprendre les nouvelles notions ;
- l'intégration dans le réseau sémantique ou le schéma est facilitée : les nouvelles informations s'encastrent bien dans son réseau sémantique.

Une grande partie des mauvais résultats des élèves serait donc due à réseau sémantique trop pauvre, qui empêcherait les indices de récupération et l'activation de faire leur travail. Le problème n'est donc pas l'effacement de l'information, mais son accessibilité, sa récupération.

Pour en donner une illustration, on peut dire que les corrélations entre moyenne scolaire et connaissances encyclopédiques en classe de 5ème sont de 0.72. Et la corrélation est la même avec le taux de redoublement 4 ans plus tard. Ce qui est plus élevé que les corrélations avec les tests de QI ou de raisonnement (0.50). cf Études d'Alain Lieury (1996).

Morale : plus un cours contient de connaissances, plus on constate de transfert positif. Le professeur doit donc multiplier les exemples et contre-exemples différents, donner un cours magistral bien rempli, fournir des explications complémentaires, multiplier les anecdotes, etc. Il faut éviter de survoler certaines notions, de zapper rapidement d'une

notion à l'autre. Au contraire, il faut voir chaque notion à fond, avec beaucoup de détails.

C'est paradoxal, mais un cours facile à apprendre est un cours qui contient beaucoup de détails pertinents et d'informations hautement structurées. Mais trop peu de professeurs cherchent à remplir leur cours à raz-bord : au contraire, nombreux sont les professeurs qui cherchent à ne pas surcharger les élèves, en diminuant la quantité de détails.

C'est notamment ce qui arrive dans les classes de niveau. Dans la classe de niveau la plus faible, la durée du cours magistral est souvent plus courte que dans les autres classes. Cela pourrait expliquer l'effet assez délétère sur la réussite scolaire des classes de niveau sur les classes de niveau le plus faible : d'après les sociologues, ces classes ont tendance à aggraver les difficultés des élèves faibles comparé à des classes homogènes.

## **Qualité et structuration des connaissances**

Reste qu'empiler les informations les unes sur les autres n'est pas une bonne idée. Comme on l'a vu plus haut, faire des liens est une nécessité, trop souvent négligée dans les cours traditionnels. Ceux-ci ont tendance à ne pas trop faire de liens entre les notions : ces cours ont un aspect assez descriptif, académique. Les connaissances semblent empilées les unes après les autres, avec certes une structure organisée par un plan, et une certaine cohérence, mais peu d'effort est fait que que les explications elle-mêmes relient les informations entre elles, l'effort mis sur la compréhension est relativement faible.

Un bon professeur doit au contraire chercher à maximiser le plus possible le nombre et la qualité des associations entre les concepts vus dans son cours.

Pour maximiser le nombre d'associations, rien ne doit être admis : tout doit être associé à des connaissances déjà acquises. Pour cela, un bon professeur doit expliquer pourquoi les choses sont ce qu'elles sont, d'où sortent les faits appris dans le cours, quelles sont les raisons, les principes physiques derrière les choses, comment les déduire à partir de principes de base, etc. C'est tout le contraire de la pédagogie traditionnelle, qui donne des cours descriptifs, dans lesquels on admet beaucoup.

Pour donner un exemple, un professeur de mathématique doit faire beaucoup de démonstrations, afin que les élèves associent les formules à apprendre avec d'autres : le professeur doit démontrer même les formules les plus simples.

## **Activer les connaissances antérieures**

Mais comment créer des associations de manière efficace ? Quelles solutions peut avoir un professeur. On peut déjà regarder ce qui est spécifiquement basé sur l'activation diffusante.

Comme on l'a vu, comprendre, c'est associer à des informations déjà apprises. Ainsi, l'apprentissage doit être comme un processus cumulatif, dans lequel on part de bases, sur lesquelles on élabore de plus en plus. Faire en sorte que tous les pré-requis soient su avant d'aborder une notion est donc une chose importante.

Mais même si les pré-requis ont été vu, il se peut qu'ils soient légèrement oublié. Cela arrive si ils n'ont pas été vu depuis longtemps. Dans ce cas, ils ne se sont pas effacés de la mémoire : on ne s'en rappelle pas parce que l'activation n'arrive pas à atteindre les pré-requis. En clair : l'élève ne comprend pas parce que les pré-requis ne sont pas assez activés.

Pour éviter tout problème, les pré-requis utilisés dans le cours sont pré-activés par **une phase de rappels**. Cette phase n'est pas une révolution, mais est malheureusement trop souvent négligée par les professeurs traditionalistes. Dans l'idéal, la phases de rappel est systématique : on y a droit à chaque cours.

Cette technique donne de bons résultats d'après les méta-analyses faites sur le sujet. Voir à ce sujet : *Classroom Instruction that Works*, Marzano, R et al (2001).

## **Advance organisers : donner la primauté aux idées générales**

Les connaissances sur les schémas peuvent aussi nous éclairer et nous donner quelques conseils pour améliorer l'ordre des notions, ainsi que la manière de les associer.

Logiquement, ces schémas sont organisés autour d'une idée générale. Cela implique qu'un cours doit commencer par

donner quelques idées générales avant d'élaborer et de rajouter des détails tout autour.

Prenons un exemple, pour montrer cet état de fait : les expériences de Bartellet. Celui-ci a travaillé sur l'apprentissage de matériel relativement complexe : cartes, schémas, textes, histoires, etc.

Une des expériences les plus marquante fut celle sur **l'effet intégrateur du titre**. Dans son expérience, il constituait trois groupes de cobayes, qui devaient tous lire un même texte assez complexe à comprendre. Le premier groupe ne connaissait pas le titre du texte. Le second connaissait le titre après avoir lu le texte, tandis que le troisième avait accès au titre avant la lecture. Le résultat : le groupe qui avait eu accès au titre avant le texte réussissait mieux que les deux autres groupes, qui avaient des performances à peu-prêt comparables. Ceci dit, mettre le titre après la fin du texte avait quand même un petit effet.

Cela vient du fait que le titre donne l'idée générale du texte, et fournit quelques informations de base, sur lesquelles les informations du texte vont venir se coller, s'associer progressivement. Cette information générale fournie par le texte permettait de structurer les informations du texte dans un schéma en mémoire sémantique, facilitant le rappel et la compréhension. Le fait que le titre doivent être placé avant le texte pour avoir un effet maximal le montre : il n'y a pas beaucoup d'associations retardées, à rebours, même s'il y en a.

Cette expérience donne quelques conseils aux professeurs de français, c'est certain. Mais on peut aussi s'inspirer de cette expérience, et chercher quelque chose qui serait à un cours ce que le titre est au texte. Ce quelque chose prend la forme d'une **phase de mise en place**, qui donne l'idée générale du cours.

Elle permet d'assembler les nouvelles informations dans un schéma, et donc de mieux organiser ses pensées.

Dans cette phase de mise en place, le professeur va expliquer aux élèves :

- ce qu'il vont apprendre durant le cours ;
- quels sont les sujets qu'ils vont aborder ;
- quel est le but de la leçon ;
- ce que les élèves seront capables de faire après la leçon ;
- etc.

On peut aussi utiliser des **advances organiser**. Ces avances organisers sont des informations, données aux élèves avant un « apprentissage » (une partie de cours), qui leur permet de mieux organiser les informations abordées dans le cours. Ces *advances organisers* servent ainsi de structure de base, sur laquelle vont se relier de nouvelles informations.

En gros, le cours magistral doit commencer par des idées relativement générales sur le sujet abordé dans le cours, et devenir de plus en plus précis au fil du temps. Ces idées générales doivent servir à donner une vue d'ensemble de ce qui va être abordé.

Pour donner un exemple, présenter le plan et le déroulement du cours avant celui-ci peut aider les élèves. Il s'agit d'une forme minimale d'*advance organiser*. Mais il en a d'autres, et j'espère que vous aurez des idées lorsque vous préparerez vos cours.

Dans le même genre, **faire un résumé du cours à la fin de celui-ci** favorise fortement la mémorisation. Ce résumé permet d'explicitier les idées importantes, celles qui sont générales, et favorise donc la formation de schémas.

Ces techniques donnent de bons résultats d'après les méta-analyses faites sur le sujet. Voir à ce sujet : *Classroom Instruction that Works*, Marzano, R et al (2001).

## **Réseau sémantiques : sur quelle base relier des informations ?**

Comment créer des associations de manière efficace ? La partie précédente a en partie répondu à la question. Mais les connaissances sur les réseaux sémantiques peuvent aussi nous éclairer et nous donner quelques conseils pour améliorer l'ordre des notions, ainsi que la manière de les associer.

La majorité des professeurs a tendance à créer des plans en se basant uniquement sur les relations de pré-requis. Si vous regardez les plans écrits par les professeurs, ceux-ci sont souvent assez linéaires : on passe d'une notion à une autre, chaque notion ayant souvent les précédentes comme pré-requis. En quelque sorte, placer les pré-requis d'une notion avant celle-ci oblige le professeur à suivre un ordre bien précis.

Si vous regardez bien certains cours, vous verrez qu'il est possible de réorganiser l'ordre des notions en voyant tout de

même les pré-requis avant la notion associée. Parmi ces progressions, certaines faciliteront fortement la mémorisation. Et malheureusement, peu de professeurs les utilisent.

La raison est simple : les professeurs ont tendance à voir un concept immédiatement après avoir abordé ses pré-requis : dès que les bases sont là, on voit le concept immédiatement. Or, dans certains cas, on peut utiliser un autre ordre des notions, de manière à faciliter l'apprentissage. Cela demande de regrouper des concepts similaires ensemble, en suivant quelques astuces.

Pour comprendre comment organiser une progression, et faire ces regroupements, nous pouvons utiliser la distance sémantique entre les concepts, ainsi que les taxonomies.

De même, ces réseaux sémantiques peuvent directement nous dire comment créer des associations assez facilement. Je rappelle que ces associations sont classées dans différents types, que l'on a déjà abordés :

- Les associations taxonomiques ;
- les associations sur la base de la distance sémantique, qui dépendent des relations de ressemblance, d'opposition, de contiguïté temporelle et spatiale, et de causalité.

## Distance sémantique

Commençons par la distance sémantique.

On peut certainement utiliser les associations de contiguïté dans certains cas, mais je n'ai malheureusement aucun exemple qui ne me viennent à l'esprit.

Par contre, les associations de causalité, qui permettent d'expliquer qui cause quoi, ne sont pas à négliger. Ces relations apparaissent naturellement quand on explique pourquoi les choses sont ce qu'elles sont, quelles est l'origine des phénomènes, etc. Ces associations sont très importantes, et manquent dans beaucoup de cours...

Pour donner un exemple, les démonstrations mathématiques permettent de faire émerger naturellement ce genre d'associations. Mais elles sont très peu utilisées en cours de mathématique, dans l'enseignement français. Et c'est bien dommage !

## Analogies, similitudes, et différences

On peut continuer en rappelant que deux concepts ont une faible distance sémantique s'ils sont ressemblants, s'ils sont proches, s'ils ont des points communs. Ainsi, un bon cours doit regrouper les notions qui sont proches, qui se ressemblent.

Pour donner un exemple, regardons les programmes de mathématique français. Dans ceux-ci, on voit comment résoudre les équations du second degré à deux endroits différents : à l'entrée en lycée, en classe de seconde, on voit comment résoudre certaines équations (celles dont le delta est positif ou nul). Puis, deux ans plus tard, on voit comment les résoudre lorsque le résultat est un nombre complexe. En utilisant ces histoires de distance sémantique, on devrait regrouper le tout dans un seul chapitre portant sur la résolution des équation du second degré, placé après celui sur les nombres complexes.

De même, le professeur ne doit pas hésiter à faire des liens entre une nouvelle situation/concept à apprendre, et des concepts/situations déjà connus. L'usage des relations de ressemblance peut aussi se baser sur des analogies bien choisies.

Le professeur peut aussi utiliser les relations d'opposition : chaud/froid, doux/rugueux, etc. Cela peut se faire en comparant deux concepts en opposition. Le seul exemple que j'ai ne sera pas très parlant, mais je tiens à citer un de mes cours sur les architectures dataflow, des processeurs spéciaux, très différents dans processeurs usuels. Dans ce cours, j'expliquais certaines fonction des processeurs dataflow en comparant celle-sic avec ce qui se faisait sur les processeurs normaux (Von neumann avec Program Counter). Cela marchait relativement bien.

Mais, chose assez originale, la recherche pédagogique s'est pas mal portée sur l'importance des relations taxonomiques. Voyons cela le plus en détail...

# Taxonomies

Un cours magistral doit respecter le plus possible une hiérarchie taxonomique. Alain Lieury a proposé d'organiser le plan en suivant cette organisation en catégories. Cette technique n'est toutefois pas présente dans les autres méthodes pédagogiques, comme la pédagogie explicite.

Avec un cours parfait, cette hiérarchie servirait à créer le plan directement. Dans les faits, peu de professeurs construisent un plan autour de hiérarchies catégorielles : la majorité me semble plutôt utiliser des progressions linéaires, assez différentes. Alors certes, il n'est pas toujours possible d'utiliser un plan hiérarchique. Mais c'est un truc utile à savoir.

Si jamais cela ne peut pas servir à créer le plan, on peut toutefois organiser certains morceaux du cours autour d'une telle taxonomie. Penser à classer les informations, les regrouper de manière à voir les informations d'une même catégorie ensemble peut fortement aider. Mais cela demande de correctement calibrer les explications, ainsi que l'ordre des notions dans un paragraphe, dans une explication, etc. Dans le pire des cas, le professeur peut donner directement cette taxonomie aux élèves, sous la forme de schémas.

Plus le matériel à mémoriser est structuré autour d'une telle hiérarchie, plus on peut l'apprendre facilement. Et les expériences réalisées sur le sujet vont dans ce sens.

Prenons un exemple : l'expérience de Bousfield (1953). Dans cet expérience, il était demandé à des élèves de mémoriser une liste de 60 mots. Ces 60 mots appartenaient à 4 catégories différentes : animaux, profession, légumes, et vêtement. Les mots sont présentés aléatoirement aux cobayes. Seulement, au fil des essais, on s'aperçoit que les cobayes ont tendance à regrouper les mots dans leurs réponses. Au final, après un grand nombre d'essais, ils rappellent les mots catégorie par catégorie. Et ce regroupement est fortement corrélé à un taux de rappel de plus en plus élevé.

Autre expérience : Clark, Lesgold et Winzenz (1969) ont fait apprendre des listes de mots à deux groupes de cobayes. Le premier avait une liste de mots organisée hiérarchiquement, et l'autre une liste complètement désorganisée. Le bilan est très clair : les participants du groupe hiérarchique ont rappelés 3 fois plus de mots que les autres. Et dans 90 % des cas, le rappel d'un mot superordonné ou basique est suivi du rappel des mots du niveau d'en dessous. Comme si la catégorie super-ordonnée servait d'indice de récupération aux autres catégories.

Autre expérience, effectuée par Alain Lieury : on fait apprendre à des élèves des listes de mots, classés en catégories. Ces mots forment des groupes de trois mots, appartenant à une catégorie intermédiaire. De plus, les trois catégories intermédiaires font partie d'une sur-catégorie qui regroupe tous les mots de la liste. Le rappel est meilleur que pour une liste désorganisée, avec les mots mis dans le désordre.

D'autres expériences sur des listes de mots sont arrivées aux mêmes résultats : classer les informations dans une hiérarchie taxonomique permet de faciliter fortement le rappel et l'apprentissage, d'un facteur pouvant aller de 3 à 5.

De plus, si on donne les noms des catégories, et qu'on montre un petit dessin montrant la hiérarchie des catégories, le rappel de la liste de mots est clairement meilleur comparé à une liste de mot convenablement triée. Mine de rien, cela peut servir pour les informations à apprendre par cœur. Les départements sont un bon exemple : les catégories sont alors les régions. Les listes de dates sont aussi concernées ; les catégories sont alors des périodes de l'histoire : moyen-âge, antiquité, etc ; avec des sous-catégories du style : règne de Louis 16, etc. Pareil pour l'apprentissage de cartes.

Pour aller plus loin, il faut savoir que certains niveaux sont plus faciles à utiliser, le niveau le plus simple étant le niveau intermédiaire. Les objets de ce niveau de catégories sont ceux que les enfants savent classer en premier, tandis que les adultes raisonnent le plus souvent avec des objets placés dans ce niveau.

Généralement, les catégories découvertes sont celles du niveau intermédiaires, suivies par les niveaux inférieurs ou supérieurs. C'est ce niveau qui est utilisé le plus facilement dans les raisonnements, dont les adultes se rappellent plus facilement. Visiblement, le niveau basique serait le plus important pour la cognition humaine.

Mieux : les études inter-culturelles montrent que les catégories basiques sont celles qui sont communes à toutes les cultures : les variances apparaissant surtout pour les niveaux sur et sous ordonnés. De plus, dans toutes les cultures, la grosse majorité des catégories sont des catégories de niveau basique. Et enfin, les mots désignant les catégories basiques sont les plus courts, ce qui signifie qu'ils sont utilisés les plus souvent (loi de Zipf). Malgré tout, quelques observations montrent qu'il existe des exceptions.

Une bonne tactique d'apprentissage est donc de commencer par apprendre les concepts basiques, avant de descendre puis de remonter dans la hiérarchie. Le plan d'un cours se structurerait en tranches, chacune abordant d'abord un concept

de niveau basique, puis les concepts sous-ordonnés affiliés. Une fois toutes les tranches d'un concept sur-ordonnée vues, on pourrait l'aborder en généralisant les points communs des concepts basiques.

Le cours utiliserait donc la méthode déductive dans chaque tranche : on utilise les connaissances de niveau basique pour aborder les niveaux sous-ordonnés. Ensuite, les concepts sur-ordonnés sont abordés par pédagogie inductive, à partir des concepts basiques.

Pour identifier les niveaux, il faut savoir que les objets du niveau intermédiaire d'une même catégorie ont tendance à partager pas mal d'attributs en commun, et la quantité de point communs est encore plus élevée pour le niveau inférieur.

Ce n'est pas le cas des objets placés dans le niveau le plus élevé : le niveau superordonné. Les catégories de ce niveau sont des catégories très générales, dans lesquelles les membres n'ont pas vraiment d'attributs en commun, mais sont plus définies par des corrélations d'attributs. Ce niveau sert surtout pour les catégories les plus abstraites, les plus floues, qui ne correspondent pas à des objets physiques clairement identifiés. Par exemple, une catégorie "véhicule", "comportement", "objet", etc en sont de bons exemples.

## ***Outils pédagogiques pour créer des associations***

Les professeurs peuvent aussi créer des associations entre idées en utilisant divers outils pédagogiques plus ou moins bien connus. Cela peut passer par l'utilisation de cartes mentales, de dessins, de diagrammes, etc.

### **Cartes mentales : ça fonctionne !**

Créer des associations, des taxonomies, cela a certainement du vous faire penser aux cartes mentales. Ces cartes mentales sont en effet un outil qui permet de visualiser les liens entre concepts. Si on se base sur ce simple fait, les cartes mentales sont donc censées être un outil pédagogique efficace.

Qu'en est-il ? La recherche est très claire là-dessus : ça fonctionne, et les professeurs ne devraient pas s'en priver !

Quelques références, pour ceux qui ne me croient pas :

- Cunningham (2005) : *Mindmapping: Its Effects on Student Achievement in High School Biology*
- Farrand, Hussain, and Hennessy (2002) : "[The efficacy of the mind map study technique](#)". *Medical Education*.
- Brian Holland, Lynda Holland, Jenny Davies (2004). *An investigation into the concept of mind mapping and the use of mind mapping software to support and improve student academic performance*.
- Marzano, R et al (2001) : *Classroom Instruction that Works*.

### **Dessins, diagrammes, et autres informations visuelles**

Autre outil assez efficace : les informations visuelles. En bref, ***un bon schéma vaut mieux qu'un bon discours***.

Il est depuis longtemps constaté que les informations sont mieux retenues quand elles se basent sur des indices visuels : c'est le **Picture superiority effect**. Cela viendrait de la profondeur de traitement : les images génèrent plus d'indices de récupération.

Dans tous les cas, l'ajout d'informations visuelles dans un cours aide les élèves. Cela peut passer par l'utilisation de cartes mentales, de schémas au tableau, de projections Power Point, ou d'autres techniques du genre. Faire des schémas, des diagrammes, des tableaux, est d'une grande aide pour les élèves.

Cela demande souvent d'utiliser le numérique et les technologies associées.

Dans certaines expériences, les chercheurs ont pu constater que l'effet sur la mémoire et la compréhension serait maximal si l'élève crée lui-même le dessin, diagramme, ou la carte mentale à apprendre. Cela peut faire l'objet de devoirs, ou se faire en classe.

Par exemple, demander aux élèves de faire une carte mentale à la fin d'une présentation (d'un cours), aurait une certaine efficacité.

## ***Transférer : comment réutiliser ce que l'on a appris ?***

On peut se demander si les phénomènes de transferts positifs apparaissent entre des domaines différents. Cela arrive quand un des domaines est un pré-requis pour l'autre. Par exemple, plus on est bon en mathématiques, mieux on se débrouillera en physique.

Mais on peut se demander si cela vaut aussi pour des domaines, quand l'un n'est pas un pré-requis pour l'autre. Par exemple, est-ce que faire des mathématiques forme réellement à la logique, d'un point de vue général ? Est-ce que faire du Latin permet de structurer l'esprit ?

On peut aussi se demander si un élève est capable d'appliquer ce qu'il a appris dans des situations différentes de celles qu'il a vu en cours.

Par exemple, un élève qui a appris à résoudre des calculs simples (fractions, divisions, multiplication, etc) est-il capable d'appliquer ces compétences pour résoudre des problèmes d'épicerie ? Même chose pour l'apprentissage des règles de grammaire : un élève qui sait appliquer ses règles de grammaire dans des exercices scolaires sera-t-il capable de les utiliser lors de la rédaction d'un texte assez court ?

Les expériences montrent que ce n'est pas forcément le cas seuls 60% des élèves seront capable de résoudre un calcul déjà vu auparavant en condition réelle. Pareil pour la grammaire : la recherche nous dit que c'est loin d'être trivial.

De manière générale, ces phénomènes de transferts entre domaines sont assez difficiles à obtenir. Les recherches faites sur le sujet ont montrées que ces phénomènes de transferts sont particulièrement rares. A vrai dire, on observe le plus souvent le phénomène inverse : des transferts négatifs, dans lequel des apprentissages antérieurs nuisent aux apprentissages futurs. Les transferts positifs sont beaucoup plus rares, et ont souvent lieu entre situations très semblables.

Les études faites sur l'apprentissage de la programmation (Pea and Kurland 1984, Salomon and Perkins 1987) montrent que celui qui apprend à programmer apprend juste à programmer, et n'acquiert pas de méthodes de résolution de problèmes, de capacité à raisonner ou à abstraire, ni quoique ce soit d'autre. Même chose pour le latin : cf. Thorndike and Woodworth (1901), Thorndike (1923), qui, contrairement aux idées reçues, ne permet pas de favoriser la mémoire, ou de structurer l'esprit. Et des études dans le genre, on en a des tonnes.

Et c'est tout à fait normal ! La psychologie cognitive de la résolution de problèmes et du raisonnement nous dit qu'il n'existe pas de compétence générale à résoudre des problèmes, raisonner, etc ; et que tout se base sur des connaissances spécialisées, des schémas et des réseaux sémantiques. Donc ceux qui veulent améliorer des compétences générales comme la rigueur, la logique, la concentration, l'abstraction, la capacité à résoudre des problèmes, etc ; font fausse route.

Mais il existe tout de même des solutions pour obtenir des phénomènes de transferts. Ces transferts apparaissent souvent entre problèmes ou connaissances fortement semblables, par transfert de connaissances spécialisées à un domaine dans des situations comportant des points communs avec la situation d'apprentissage.

Les mécanismes principaux de transferts sont surtout basés sur les schémas, qui permettent de reconnaître les situations familières. A vrai dire, on peut considérer que les experts, étudiés dans les études sur l'expertise et la résolution de problèmes, sont des transféreurs particulièrement efficaces. Leur grande quantité de schémas, et l'élaboration de leurs réseaux sémantiques permet aux indices de récupération de fonctionner à plein : toute situation un tant soit peu familière va immédiatement évoquer un schéma bien précis en mémoire, récupéré grâce au jeu de l'activation diffusante.

La conséquence, c'est que les schémas sont la base : ils doivent être accessibles à partir d'un grand nombre de situations, et doivent être accessibles dans le réseau sémantique. Conséquence : il faut jouer sur l'organisation des schémas et du réseau sémantique : la compréhension est la base.

La plus grosse difficulté, qui empêche souvent le transfert, est de reconnaître que deux problèmes sont semblables. Une solution pour cela, c'est créer des relations de ressemblance entre concepts. On a vu précédemment que les analogies sont un bon moyen pour transférer des connaissances d'une situation source, vers une situation-cible.

Un bon exemple, assez classique dans la littérature scientifique, est celui fourni par Gick et Holyak (1980). Dans son expérience, il demande à ses cobayes de trouver une solution pour guérir un patient d'une maladie à l'estomac avec des rayons X. Seul problème : les rayons ne doivent toucher que la tumeur, et pas les tissus environnants. Or, les rayons forment un faisceau un peu trop large. Comment faire ? La solution est de concentrer les rayons.

Les cobayes sont séparés en deux groupes. Un premier groupe n'a aucune autre indication. La réussite dans ce groupe n'est pas fameuse : 8% des personnes trouvent la solution.

L'autre, étudie un autre problème avant de résoudre celui sur la tumeur. Le problème est le suivant : une forteresse est accessible par un ensemble de routes, qui converge vers elle à la manière des rayons d'une roue. Un général veut prendre la forteresse d'assaut, et est certain qu'avec tout ses hommes, il peut faire tomber la forteresse. Mais un espion lui indique que les routes sont minées, et que si trop d'hommes passent dessus, les mines exploseront, mais de petits groupes d'hommes peuvent passer sans danger. La solution de ce problème est la même : il faut disperser les forces, avant de les concentrer au point d'attaque.

Ce second groupe a un taux de réussite nettement supérieur : 30%. Et si les expérimentateurs disent que la solution du premier problème peut aider à résoudre le second, on passe à 90%.

Un autre effet vient jouer dans les phénomènes de transfert : **l'effet de contexte**. Cet effet nous dit que plus les circonstances du rappel sont proches des circonstances d'apprentissage, plus il est facile de se rappeler. En quelque sorte, le contexte peut servir d'indice de récupération.

Pour donner un exemple, il faut savoir que l'on se rappelle mieux de quelque chose si l'on se trouve dans l'endroit dans lequel on l'a appris. Quelques expériences ont ainsi comparé deux groupes de cobayes, qui devaient apprendre une liste de mots. Après avoir appris la liste, on leur donnait un temps de repos d'un jour, et on faisait passer un test de rappel aux deux groupes. Les cobayes du premier groupe passaient le test dans la salle dans laquelle ils avaient appris, alors que le second passait le test dans une autre salle. Le premier groupe avait des résultats 15 à 20 % supérieurs.

L'effet a aussi été observé en faisant varier non pas la salle de classe, mais l'expérimentateur. Si l'examineur est la même personne que celle qui a surveillé l'apprentissage, le rappel est meilleur. Certains psychologues ont même été jusqu'à garder la même salle, mais ont changé l'ordre des tables. L'effet restait !

Et on peut citer d'autres expériences du même genre, qui évaluaient la présence d'une odeur. Lors de l'apprentissage de la liste par les cobayes, les expérimentateurs ont diffusé une odeur particulière dans la pièce. Lors du test, un groupe sentait cette odeur juste avant le rappel, pas l'autre groupe. Le groupe qui a senti d'odeur a toujours des résultats supérieurs. Comme quoi, le coup de la madeleine de Proust, c'était pas du chiqué !

Et ce ne sont pas les seuls exemples. Le contexte cognitif joue aussi. Par exemple, se rappeler quelque chose lors d'un test dépend de la langue d'apprentissage. Si on vous fait apprendre quelque chose en allemand, vous aurez plus de facilité à donner la réponse si la question est posée en allemand qu'en français. Cet effet a été mis en évidence chez des étudiant anglais-espagnols bilingues (Marian and Fausey, 2006), mais peut certainement se généraliser à tout le monde.

Même de rien, cet effet a une influence sur la fiabilité du contrôle continu. On en déduit facilement que le contrôle continu permet de faire facilement augmenter les notes de quelques points. On comprend mieux les revendications de certaines associations étudiantes universitaires.

Mais cet effet nous dit aussi que se contenter de répéter des exercices identiques ne servira qu'à résoudre de mieux en mieux ces exercices. Pour l'application dans des situations autres, comme dans la vie courante ou professionnelles, c'est autre chose : le contexte ne sera pas du tout le même.

Par exemple, apprendre des règles de grammaires avec des exercices types, académiques, sera difficilement transférable à l'écriture ou la lecture d'un vrai texte. Surtout que les mécanismes d'apprentissage de ce genre de compétence est loin d'être simple, cf <http://www.cahiers-pedagogiques.com/L-erreur-orthographique-l-apprentissage-implicite-et-la-question-des-methodes-de-lecture-ecriture>.

La morale est claire : on apprend peut-être avec le bled et en répétant des exercices, mais cela ses limites. De nombreux élèves peuvent avoir de très bonnes notes dans des exercices académiques, mais ne pas savoir appliquer leurs connaissances lors de rédaction de textes simples. Des expériences sur la gestion du pluriel l'ont montré : les enfants sont parfaitement capables de détecter des erreurs (un S là où il ne devrait pas y en avoir, ou inversement) avec un taux e réussite de 95%, alors qu'environ 30% des mots orthographiés dans un texte qu'ils ont écrit peuvent comporter une erreur.

Même chose pour l'apprentissage mathématique : apprendre à résoudre des fractions ne signifie pas que l'on sera capable de résoudre des calculs d'épiciers en transférant les méthodes de résolution de problèmes acquises en cours. Dans une expérience, on a demandé à des élèves de faire des calculs de fraction dans deux contextes.

Dans le premier contexte, les élèves devaient résoudre des exercices scolaires, comme  $4/6 * 8/9$ . Dans l'autre, les élèves

devaient faire des calculs d'épiciers. Si la réussite était de 87% dans le premier groupe, elle tombait à 60% dans le second groupe. Dans les deux cas, le contexte cognitif ne sera absolument pas le même.

Il y a une solution : faire en sorte que le contexte d'utilisation soit le même que celui d'apprentissage. Un professeur doit donc multiplier les indices de récupération potentiels en variant les contextes d'apprentissages.

Autre problème : les connaissances apprises dans un contexte auront du mal à être abstraites et utilisées dans d'autres contextes. En clair : on doit décontextualiser les connaissances, avant de les appliquer dans des situations diverses. On doit d'abord apprendre d'abord sans contexte, puis ajouter des exemples d'application divers et variés, avec beaucoup de contextes...

Par exemple, on peut donner des exemples d'application tirés de la vie réelle ou proches de conditions d'emploi. Pour les exercices de fraction, cela correspond à donner des calculs d'épiciers en guise d'exercice. Au lieu d'un banal exercice du type : *calculer 43,25 - 23,18*, on pourrait donner des exercices ancrés dans un contexte réel.

Le professeur peut aussi expliciter les situations de transfert potentiels : « attention, là vous allez pouvoir réutiliser quelque chose qu'on a vu il y a deux semaines » !

## ***Transferts négatifs: quand ce que l'on sait empêche l'apprentissage !***

Le transfert, c'est bien. Mais il arrive parfois qu'associer avec des connaissances antérieures aie des effets négatifs : un apprentissage antérieur peut perturber l'apprentissage en cours. On fait alors face à des phénomènes d'**interférence**.

## **Mis-conceptions**

On distingue deux types d'interférence, dont le premier est l'interférence pro-active : d'anciennes connaissances empêchent la mémorisation de nouvelles idées. C'est ce qui explique que quelqu'un qui a appris l'espagnol aura plus de mal à apprendre l'italien, par exemple.

L'autre type d'interférence est l'interférence rétro-active, dans laquelle des connaissances plus récentes ont tendance à favoriser l'oubli d'anciennes informations. Par exemple, quelqu'un qui apprend l'italien aura tendance à oublier ses leçons d'espagnol qu'il aura vu il y a quelques années.

Ces interférences n'ont lieu qu'entre informations semblables. Des informations très semblables ou très dissemblables ne donnent pas lieu à des phénomènes d'interférence. L'interférence proviendrait de l'activation diffusante. Lors d'un processus de rappel quelconque, celle-ci pourrait activer plusieurs informations semblables et les élire pour le rappel. Seul problème, ces informations étant ressemblante, le cerveau ne saurait pas laquelle est la bonne : un phénomène de confusion empêcherait un rappel correct.

Prenons un exemple d'un élève qui a une erreur dans sa mémoire. Si on corrige celle-ci, la correction s'ajoutera dans le réseau, mais ne modifiera rien : l'erreur et la correction sont accessibles à partir des mêmes indices de récupération. Au final, l'activation activera les deux en même temps : l'erreur et la correction interféreront et entreront en compétition pour le rappel, perturbant celui-ci. Généralement, il faut une grande quantité de répétitions pour que la correction domine l'erreur.

Je me permet de faire une parenthèse sur l'interférence de manière générale. Celle-ci est loin d'être un mal dans certains cas. En effet, c'est elle qui est à l'origine du **processus d'abstraction**. Ce sont elles qui font que multiplier les contextes permet un meilleur transfert positif.

En effet, quand on étudie plusieurs choses similaires ou d'une même catégorie, ces deux choses auront tendance à être reliées ensemble dans le réseau sémantique. Évidemment, lors de la récupération, les détails des différentes situations auront tendance à se confondre, et à interférer s'ils sont relativement ressemblants. Par contre, les généralités, les informations valides pour une grande partie de ces situations, n'interféreront pas : elles se ressemblent trop (et sont parfois totalement identiques) pour éviter toute confusion.

On voit donc que donner une grande quantité d'exemples peut permettre à un élève d'extraire des généralités, et de former un schéma cohérent d'un concept. En conséquence, donner beaucoup d'exemple est une bonne pratique. De plus, certains concepts très abstraits peuvent s'apprendre à partir d'exemple. Il arrive que pour certains concepts bien ciblés, la pédagogie inductive puisse donner de bons résultats, quoi qu'on en dise ! Les concepts en question doivent être toutefois assez abstraits pour l'étudiant.

Mais ces interférences ont aussi des défauts : ce sont elles qui sont à l'origine des phénomènes de transferts négatifs. En dehors des situations citées juste avant, les éviter est donc une nécessité.

En cours, ces phénomènes de transferts négatifs proviennent de deux sources principales. La première source est l'utilisation de schémas dans des situations où ils ne sont pas adaptés. Si jamais un élève utilise un schéma pour comprendre une nouvelle information, et que le schéma choisit n'est pas le bon, le transfert ne se produit pas.

On peut éviter cela en faisant des cours clairs, et en utilisant convenablement les phases de rappel, de mise en place, et les *advances organisées*.

Ensuite, la seconde source est la présence de défauts dans les schémas et dans le réseau sémantique, qu'on appelle des **mis-conceptions**. Pour simplifier, ces mis-conceptions sont des connaissances incorrectes, qui interfèrent avec de nouveaux apprentissages. Ces misconceptions empêchent de nouvelles informations de s'encaster dans un schéma ou un réseau sémantique, en empêchant certaines associations.

Ces mis-conceptions sont parfois des idées reçues, des connaissances naïves, qui sont tirées de l'expérience.

Par exemple, demandez à n'importe quel enfant de dire ce qui est le plus lourd entre un kilo de plume et un kilo de plomb : 90% vont se tromper et dire qu'un kilo de plomb est plus lourd. Ils vont naturellement faire une confusion entre densité et poids. Il faut dire que dans notre environnement, nous avons rarement des exemples nous permettant de faire la distinction.

Autre exemple : demandez à un enfant ce qui se passe si une fusée spatiale, qui voyage dans un vide parfait, coupe ses moteurs. Tous vont vous dire que la fusée va freiner. C'est normal : ils se basent sur ce qu'ils ont l'habitude de voir, à savoir qu'un objet qui roule finit toujours par s'arrêter tout seul.

Ces connaissances intuitives, ces conceptions du monde erronées, sont des schémas erronés. La plupart naissent spontanément, ou par observation de l'environnement. Les exemples mentionnés plus haut sont d'ailleurs des mis-conceptions qui proviennent d'une étude attentive de l'environnement. Comme quoi, laisser un élève construire son savoir est une bonne manière de lui faire générer une chieée de mis-conceptions.

Ces mis-conceptions peuvent aussi venir de la généralisation abusive d'apprentissages antérieurs. Dans ce cas, on peut considérer qu'ils s'agit de phénomènes de transferts négatifs. A savoir, le fait de vouloir appliquer quelque chose alors qu'on ne devrait pas.

Par exemple, prenons un débutant en programmation, qui commence à programmer en C. Ce langage de programmation comprend la notion de variable, ainsi que celle de fonction. Mais ces deux concepts sont très différents des concepts de variables et de fonction abordés en mathématique. Et bien les étudiants auront tendance à faire le transfert, donnant lieu à des erreurs.

Et des erreurs dans le genre, on en trouve quand même pas mal dans l'apprentissage de la programmation. Ceci dit, il existe diverses manières pour éviter l'apparition de ces mis-conceptions. Les travaux sur les machines sont relativement mal connus, mais ils existent et donnent des résultats.

Ces mis-conceptions peuvent aussi être de simples erreurs de compréhension, qui ont exactement les mêmes conséquences. Ces erreurs sont des défauts dans le réseau sémantique, et naissent quand un élève n'associe pas correctement une information au bon endroit dans le réseau sémantique : cela peut être une association en trop, une qui manque, une autre mal placée, etc. Elles ont tendance à déformer ceux-ci, à les fausser.

## Conséquences

La présence de mis-conceptions peut empêcher l'acquisition de nouvelles connaissances, qui auront du mal à s'encaster dans le réseau sémantique aux endroits concernés par le défaut. De plus, elles empêchent l'activation diffusante de diffuser convenablement. Et c'est facile à comprendre : si le schéma est inadéquat, il est normal que les informations aient du mal à s'y insérer.

Pire : elles peuvent interférer avec les tentatives de correction et se cristalliser en mémoire. A force d'utilisation, celles-ci se consolident et deviennent de plus en plus difficiles à corriger.

Les pédagogies actives sont toutes une véritable catastrophe de ce point de vue : les élèves procèdent par séries d'essais et d'erreurs, chaque erreur étant mémorisée. La progression est donc très lente, les élèves assimilant des schémas

boiteux, des erreurs de compréhension et des erreurs procédurales par paquets.

Un professeur doit absolument éviter l'apparition de mis-conceptions et d'erreurs de compréhension. Les schémas des élèves doivent être les plus fiables possible : **toute connaissance apprise ne doit pas interférer avec des apprentissages ultérieurs.**

Conséquence : **le cours doit être rigoureux, précis, et n'admet pas la moindre simplification ou généralisation abusive.** Ce qui est appris ne doit pas être remis en cause dans la suite de la scolarité.

Par exemple, un professeur explicite de mathématique ne dira pas : « la racine carrée d'un négatif n'existe pas », mais « la racine carrée d'un négatif n'est pas un réel ». Oui, je sais : l'exemple est stupide, et le conseil semble être trivial. Mais il existe des situations assez subtiles dans les cours de physique ou de science qui font que ce genre de choses est plus courante qu'on ne le croit.

Autre solution : **bien préciser les différences entre concepts déjà vus et ceux à apprendre.** Un professeur doit bien mettre l'accent sur ce qui différencie des concepts semblables. De plus, il doit bien préciser quand certaines connaissances peuvent être généralisées, et quand elles ne le peuvent pas.

Autre solution : déminer explicitement toute mis-conception. De nombreuses recherches en didactiques cherchent à identifier les erreurs de compréhensions courantes. Les connaître permet de prendre les devants et de **déminer préventivement** ces erreurs avant leur apparition.

De même, le professeur peut connaître les mis-conceptions courantes, que les élèves ont avant de commencer un cours. Ce qui permet encore une fois de **prendre les devants.**

Par exemple, prenons une mis-conception courante : les corps tombent dans le vide à une vitesse qui dépend de leur poids. Demandez à n'importe quel élève assez jeune, il croira que c'est vrai (alors qu'en fait, tous les corps tombent à la même vitesse dans le vide). Un professeur de physique pourra faire l'expérience devant ses élèves, afin de leur prouver que si, les objets tombent à la même vitesse.

Autre solution : corriger au plus vite des erreurs et mis-conceptions. Pour cela, le professeur doit se réserver une période de **feedback**, qui permet aux élèves d'identifier leurs erreurs de compréhension et de les corriger. Dans cette phase, le professeur va tester les élèves avec des questions de compréhension, voire par une série d'interrogations écrites.

La solution privilégiée est l'utilisation massive d'interrogation écrite systématique à la fin de chaque cours, qui peuvent porter sur le contenu du cours qui vient juste d'être fait, ou sur le contenu des cours précédents.

## **Mémoire sémantique et apprentissage : l'importance de la répétition**

Maintenant, après le cours, il faut aussi regarder ce qui en reste après les cours, ou quelques années après. Or, des expériences sur l'apprentissage de listes de mots semblent montrer que l'on oublie très rapidement ce que l'on a appris, en à peine quelques jours, voire semaines.

Pour montrer cet effet de dégradation rapide de la mémoire suivi d'une stabilisation définitive, il suffit de mesurer à intervalles réguliers ce que l'élève a retenu du matériel appris. Le bilan est ce qu'on appelle une **courbe de l'oubli.**

Contre-intuitivement, cet oubli s'arrête rapidement. Le taux d'oubli diminue rapidement avec le temps, et au final, les élèves retiennent une grande partie de ce qu'ils ont appris en classe. Entre 30 à 20% de ce qu'on a vu en classe est conservé, même plusieurs dizaines d'années après avoir quitté l'école.

Le taux exact dépend fortement du niveau de maîtrise initial : ceux qui retiennent le mieux sont ceux qui ont appris facilement le matériel, et qui n'ont pas eu de difficultés particulières à le comprendre.

Le temps d'oubli peut être assez long, et finit souvent par se stabiliser. Pour donner un exemple, je vais citer une étude faite sur les connaissances en espagnol apprises durant la scolarité d'élèves américains. Cette étude a montré que la phase de déclin peut prendre 3 ans, avant que l'oubli ne se stabilise. Une fois stabilisé, les mots rescapés ne s'oublient plus par la suite.

Le chercheur qui a découvert cette stabilisation lui a donné le mot d'**effet permastore**. Ce mot fait référence au

permafrost, à savoir l'état totalement gelé de certains sols dans les régions froides (en Sibérie, notamment). Tout se passe comme si les souvenirs finissait par geler, par se consolider si fortement en mémoire qu'ils ne peuvent plus être oubliés.

D'autres études sur la capacité à reconnaître ses anciens camarades de classe sur des photographies ont montrées que l'oubli peut être encore plus lent, avec des courbes de l'oubli quasiment horizontales.

Le bilan est que l'oubli s'effectue en deux phases : une phase d'oubli dans laquelle un grand nombre de connaissances sont oubliées, suivie d'une phase de stabilisation, dans laquelle l'oubli est particulièrement lent.

## Raisons de l'oubli

Bref, ces courbes d'oubli semblent malheureusement stables : on oublie vite, mais cet oubli s'arrête rapidement et les connaissances qui ont survécus sont apprises définitivement. Mais quel est la cause de ces oublis ? Est-ce un effacement des informations en mémoire, ou un problème de disponibilité ?

Les expériences montrent clairement que la courbe d'oubli en reconnaissance est différente de celle obtenue par des taches de rappel. De plus, les taches de rappel indicées n'ont pas la même courbe d'oubli que les taches de rappel pur. Si les souvenirs en rappel, rappel indicé et reconnaissance se dégradent à la même vitesse, on peut reconnaître deux fois plus d'informations en reconnaissance/rappel indicé qu'en rappel.

De plus, la pente de la courbe dépend fortement de la nature du matériel à apprendre. Du matériel non-familier, peu intéressant, ou difficile à apprendre/comprendre est oublié rapidement. Par contre, tout ce qui est facilement compris est rarement oublié : c'est l'effet de la profondeur de traitement.

En clair : une grande partie des oublis ne vient pas d'un effacement de l'information, mais d'un problème de disponibilité, d'accessibilité dans le réseau sémantique.

Le professeur doit donc jouer sur la pente de la courbe d'oubli. C'est ce que font les techniques basées sur la catégorisation, et la compréhension : le fait de pouvoir accéder facilement à l'information permet d'éviter certains oublis, et donc de diminuer la pente.

En plus, la pente de la courbe est dépendante de la qualité de la mémorisation : les techniques basées sur la catégorisation et les associations forment des souvenirs beaucoup plus solides, plus résistants à l'effacement.

## Sur-apprentissage

Embinghaus a supposé que pour mieux diminuer la perte avec le temps, il fallait utiliser le **sur-apprentissage** : à savoir continuer à réviser, même une fois que le matériel à apprendre était totalement su. Cette technique a été reprise par la pédagogie explicite, et c'est une erreur grave !

Expérimentalement, cette technique ne marche pas ! Différentes études ont montrées que le sur-apprentissage permettait d'améliorer la rétention du matériel répété, mais seulement sur une période de temps très faible : les effets semblent s'estomper après quelques semaines, voire quelques mois.

Une de ces études ("The Effect of Overlearning on Long-Term Retention") a demandé à des étudiants de retenir des faits géographiques et des définitions de mots. Les étudiant qui avaient sur-appris retenaient mieux, mais les effets s'estompaient assez rapidement, quelques semaines plus tard. Une autre de ces études (The Effects of Overlearning and Distributed Practise on the Retention of Mathematics Knowledge) a testé la même chose, mais pour des exercices de mathématique. Le bilan est le même : pas d'effets à long terme.

## Nature

Une solution pour limiter l'oubli, est la **répétition**. Plus on répète quelque chose, plus on s'en souvient. Le principe de la répétition est que la mémoire s'efface rapidement, et que son contenu se détériore si on ne le répète pas.

Là encore, la répétition est importante. Reste que cette répétition peut varier suivant différents paramètres, qu'il vaut mieux optimiser le plus possible. On peut jouer sur deux paramètres :

- le type de répétition ;
- et sa distribution dans le temps.

On distingue trois types de répétition, classées de la moins efficace à la plus efficace :

- répétition de maintien ;
- répétition d'élaboration ;
- répétition de rappel.

## Répétition de maintien

La **répétition de maintien** consiste à revoir le matériel à apprendre à l'identique. Il s'agit d'une forme de répétition utilisée dans le par cœur, la répétition d'exercices, etc. Malheureusement, cette répétition de maintien ne permet pas de créer de nouvelles associations, ni d'enrichir les réseaux sémantiques : elle joue sur la disponibilité de l'information, mais pas sur son accessibilité.

Tous les psychologues considèrent qu'il s'agit d'un bon moyen pour maintenir des informations en mémoire à court terme, mais pour le passage en mémoire à long terme, il faut trouver autre chose.

Pour donner un exemple flagrant de la mauvaise efficacité de la répétition de maintien, des chercheurs ont demandé à des volontaires de décrire du mieux qu'ils pouvaient une pièce de un penny, un centime américain. De plus, ils leur ont posé quelques questions simples sur cette pièce.



Bilan : les volontaires ne se souvenaient pas des détails de la pièce. Par exemple, ils ne savaient pas si le visage était tourné vers la droite ou la gauche. Pourtant, ces personnes voient cette pièce tous les jours.

Dans le même genre, on peut demander à des cobayes quelle est la disposition des chiffres (les touches) sur leur téléphone. Le taux de fausses réponses est particulièrement élevé.

La répétition de maintien est donc à éviter : le " par cœur " ne fonctionne pas. De même, relire ses cours n'est pas efficace. Il faut trouver d'autres méthodes de travail. Pourtant, beaucoup d'étudiants utilisent en priorité la répétition de maintien lors de leurs révisions.

## Répétition de rappel

Mais il existe une méthode de répétition qui surpasse tous les autres : la **répétition de rappel**. Expérimentalement, augmenter le nombre de rappels est un meilleur moyen de mémorisation que la répétition de maintien ou d'élaboration. C'est ce qu'on appelle le **Testing effect**.

Pour avoir un effet optimal, il vaut mieux que les questions posées aient un niveau de difficulté moyen à faible. Avec des questions trop dures, le nombre de rappels serait trop faible pour avoir un effet significatif.

Pour donner un exemple, je vais prendre une expérience effectuée par des chercheurs de l'université de St Louis, aux USA. Ils ont conçus une épreuve de mémorisation d'une liste de 50 mots, répartis dans 10 catégories différentes. Ils ont ensuite pris trois groupes d'élèves. Le premier devait réviser et bachoter durant 8 séances. Le second devait bachoter durant 6 séances, et se faire tester dans deux autres. Enfin, le troisième groupe devait Bachoter durant 4 séances de répétition de maintien, et se faire tester durant 4 séances. Immédiatement après, les étudiants étaient testés : pas de différence entre les groupes. Mais une semaine plus tard, les étudiants repassaient devant l'examinateur. Bilan : plus un groupe se faisait tester, plus les résultats étaient meilleurs.

Autre expérience, réalisée par Roediger & Karpicke. On demande à des étudiants de lire des textes scientifiques

relativement courts. Un premier groupe devait lire le texte 4 fois. Un second groupe devait le lire trois fois, et se faisait tester une fois. Un troisième groupe devait lire le texte une seule fois, mais devait subir trois interrogations. Une semaine plus tard, les trois groupes étaient testés : plus un groupe avait été testé, mieux c'était. A vrai dire, le troisième groupe réussissait à rappeler 50 de mots en plus ! Imaginez si vos notes augmentaient dans cette proportion : passer de 8 à 12 n'est pas négligeable !

Quelques études récentes semblent montrer que relire ses cours est moins efficace que les stratégies d'auto-test ou de test par un pair. Cela viendrait du fait que relire ses cours donne une illusion de compétence : un texte est rempli d'indices de récupération qui font qu'on comprend facilement. En situation de rappel, on n'a pas autant d'indices de récupération

Et des expériences dans le genre, on en trouve beaucoup d'autres. Et elles convergent toutes vers la même conclusion : la répétition de rappel est beaucoup plus efficace, et cela peut fortement augmenter les résultats. En clair : c'est une arme de mémorisation massive. Dommage qu'elle ne soit pas utilisée dans la pédagogie explicite, ni dans les autres pédagogies du même genre.

Il vaut donc mieux utiliser des exercices remplis de question de rappel et de compréhension lors de ses interrogations. Il faut considérer les interrogations et évaluations comme une méthode d'apprentissage, et non un moyen de vérifier les connaissances de l'élève. De plus, la dernière phase du cours est une phase de test, dans laquelle le professeur pose beaucoup de questions de compréhension, voire organise une interrogation écrite systématique à chaque fin de cours.

Bilan : au lieu de relire leur cours, on gagne plus à se faire tester par un ami, ou à s'auto-tester. Se faire poser des questions, ou s'auto-poser des questions est donc un bon moyen de mémorisation.

## Et si on ajoutait de la compréhension ?

Une autre technique est la **répétition d'élaboration**, qui consiste à revoir le matériel à apprendre en ajoutant de nouvelles associations à chaque répétition. Cette répétition est une sorte de mélange entre répétition pure, et compréhension.

Le principe caché derrière cette forme de répétition, c'est les réseaux sémantiques : cette forme de répétition permet de créer de nouvelles associations qui serviront d'indices de récupérations ou de points de convergence vers le concept à rappeler.

Pour élaborer, on peut mélanger la répétition de maintien avec de la compréhension. Ce la peut se faire en :

- complétant son cours avec des informations tirées de Wikipédia, ou d'autres sources extérieures au cours ;
- en faisant faire des exposés aux élèves, etc.

Un étudiant ou un professeur peut aussi multiplier les « épisodes d'apprentissage », afin de rajouter des informations supplémentaires, et d'affiner de plus en plus la compréhension du sujet. Cette méthode d'apprentissage, appelée multi-épisode par son concepteur, Alain Lieury, est des plus efficace.

Le professeur peut aussi mixer répétition d'élaboration et répétition de rappel dans les mêmes exercices. Il suffit de jouer sur le contenu des exercices. Par exemple, la mémorisation est meilleure quand on demande à l'élève de ré-expliquer avec ses propres mots. C'est le **Self-explanation effect**.

Les exercices qui mélangent élaboration et rappel vont donc demander à l'élève de :

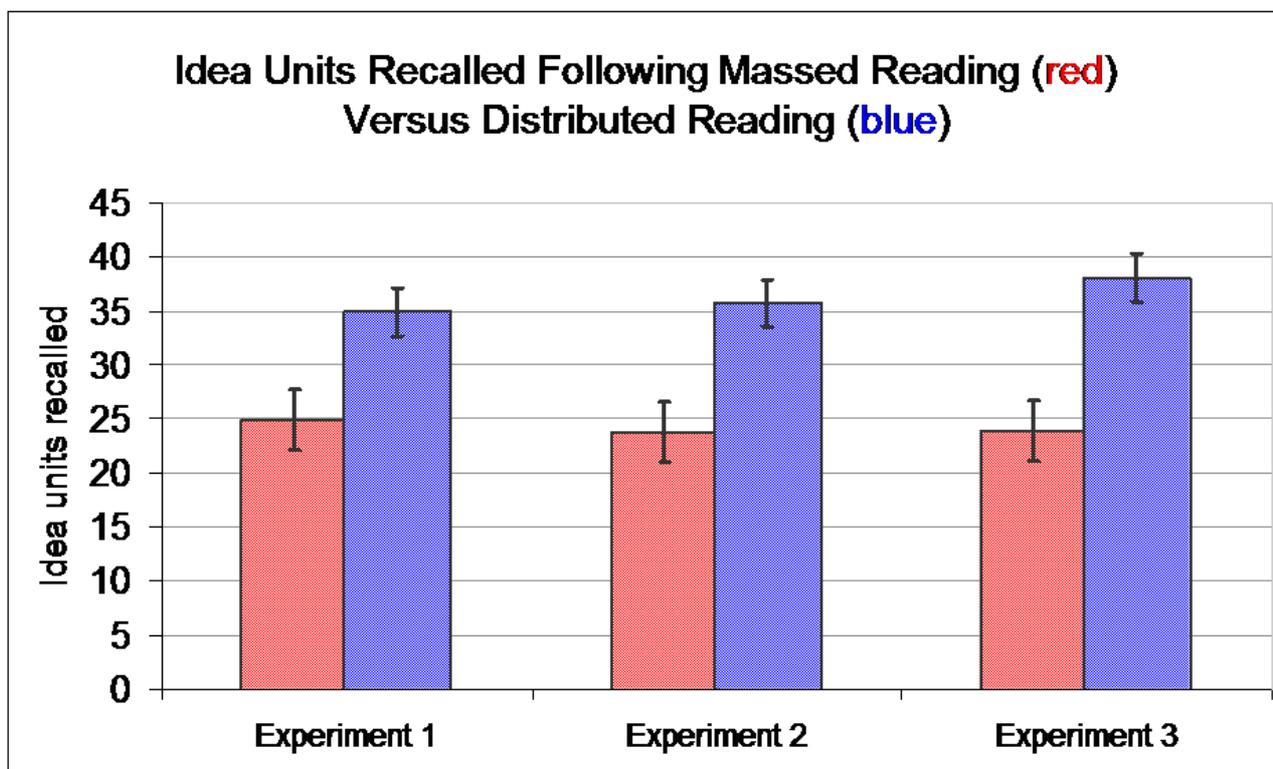
- répondre à des questions de compréhension ;
- réaliser un exposé ;
- réaliser des résumés du cours ;
- faire des synthèses ;
- créer une carte mentale du cours ;
- etc.

## **Distribuer**

Nous avons vu que répéter était la seule technique pouvant limiter l'oubli, et que toutes les formes de répétitions ne se

valent pas. Ceci dit, nous n'avons pas encore parlé du dernier paramètre : la répartition temporelle. On peut répartir nos répétitions dans le temps de deux façons : soit on effectue toutes les répétitions en une fois, soit on les étale dans le temps. La première solution s'appelle l'apprentissage massé, et la seconde l'apprentissage distribué.

Concrètement, la seconde solution consiste à faire de nombreuses pauses assez longues entre les répétitions : on effectue quelques unes, puis on fait une pause, avant de reprendre.



Expérimentalement, on constate que l'apprentissage distribué démonte littéralement l'apprentissage massé : le seul fait de distribuer l'apprentissage peut permettre d'améliorer la mémorisation de plusieurs ordres de grandeurs. En clair : on sort l'artillerie lourde. Et cet effet apparaît dans presque toutes les situations d'apprentissage. Cette supériorité de la distribution sur l'apprentissage massé s'appelle le **spacing effect**.

Ainsi, un élève aura tout intérêt à répartir ses devoirs sur la semaine de façon à en faire un petit peu à chaque fois. De même, faire de longues pauses tout les deux-trois exercices aidera fortement l'élève. Répartir et gérer son temps est une obligation pour les élèves.

De même, les professeurs gagent à étaler les devoirs et de distribuer les révisions : un bon professeur donne fréquemment des devoirs, mais en donne peu à chaque fois. De ce point de vue, les statistiques montrent clairement qu'il y a du chemin à faire.

Les origines de cet effet ne sont pas claires : il semblerait que cet effet aurait plusieurs explications. Une des raisons est que distribuer l'apprentissage empêcherait certains phénomènes d'interférence d'intervenir, et d'autres théories du genre postulent des mécanismes ayant un rapport avec les réseaux sémantiques. Ceci dit, ce phénomène a aussi été observé pour des tâches de mémoire implicite : cela marche aussi pour l'apprentissage sportif !

## Mémoire à court terme et théorie de la charge cognitive

Pour mémoriser une connaissance, on est obligé de passer par la MCT. Plus elle est remplie, moins la compréhension et la mémorisation sont bonnes. Si elle sature, c'est encore pire : les performances s'effondrent, l'élève fait des erreurs, etc.

Cette MCT a une grande influence sur les performances des élèves : diverses études ont montré que la capacité de la MCT à 5 ans est un bon indicateur de réussite scolaire. De plus, on observe une corrélation entre faible capacité de la MCT et échec scolaire. Cette MCT a surtout une grande influence dans l'apprentissage de la lecture, de l'écriture, et de la numération : à cet âge-là, les enfants ont une MCT contenant peu d'informations.

La quasi-totalité des problèmes de mémorisation ou de compréhension sont dus à la gestion des MCTs par le professeur, qui a tendance à surcharger la MCT. Le rôle du professeur est d'éviter que la mémoire à court terme se remplisse : il doit diminuer le **charge cognitive**, c'est à dire le nombre d'informations présentes simultanément dans la MCT.

La pédagogie traditionnelle ne prend pas en compte cette charge cognitive. Mais de ce point de vue, les pires pédagogies sont les pédagogies actives qui utilisent des situations à forte charge cognitive : gérer un projet n'est pas simple, résoudre des problèmes non plus. La MCT des élèves est alors totalement remplie et le bilan est catastrophique : une grande partie de l'attention est utilisée pour résoudre le problème ou gérer le projet, et n'est pas utilisée pour la mémorisation ou la compréhension.

Pour diminuer l'occupation de la MCT, le professeur peut utiliser la **théorie de la charge cognitive**, une théorie qui recense plusieurs effets ayant chacun des implications pédagogiques. Pour être franc, cette théorie est une vraie arme d'apprentissage massive, sur laquelle toutes les pédagogies tirées de la psychologie cognitive se basent, sans exception.

Certains de ces effets permettent au professeur de créer des supports de cours, d'interrogation ou d'exercice plus efficaces :

- *l'element interactivity effect*,
- le *worked exemple effect*,
- le *problem completion effect* ;
- *l'expertise reversal effect* ;
- le *guidance fading effect* ;
- le *goal-free effect* ;
- le *modality effect* ;
- le *redundancy effect* ;
- le *split attention effect*.

## ***Element interactivity effect : ne pas surcharger la mémoire de travail***

Nous allons commencer par aborder le premier effet : *l'element interactivity effect*. Les autres effets seront abordés progressivement, au long de cet article.

La MCT stocke des morceaux d'informations plus ou moins complexes qu'on appelle des **chunks**. *L'element interactivity effect* nous dit plus on stocke de chunks en mémoire à court terme, et plus ces chunks interagissent entre eux, plus les performances s'effondrent. Diminuer à la fois le nombre de chunks, et leurs relations, est donc primordial.

Une bonne solution pour réduire la charge cognitive est de regrouper plusieurs informations dans un seul *chunk*. Par exemple, essayez de mémoriser le numéro de téléphone suivant : 0 3 2 4 5 3 4 6 7 8. Et maintenant essayez avec celui-ci : 56 23 98 45 22. Cette fois-ci, c'était plus facile : le fait de regrouper les chiffres fait que le nombre de *chunks* à mémoriser était plus faible : deux fois moins.

Ce chunking s'applique dans la construction d'un cours : cela donne un cours hyper-structuré, découpé en chunks. Cela passe par une factorisation des concepts à aborder en sous-concepts, eux-même potentiellement découpés. Chacun de ces chunks doit contenir peu d'informations, entre 3 à 4 maximum.

### ***Structurer son cours en chunks est la clé.***

Évidemment, on ne va pas chunker un cours n'importe comment. Le meilleur découpage est un découpage sémantique : chaque *chunk* doit avoir une signification.

Par exemple, regardez les suites de mots suivantes :

*shampoing, voiture, champ, bal, lance-roquette.*

*crata, scronno, munu, carrac, shullp.*

La première est plus facile à mémoriser que la seconde. Cela vient du fait que les mots de la seconde sont inconnus : pour les mémoriser, on doit mémoriser chacun des sons, et l'ordre de chaque son : la MCT déborde. Pour la première, il suffit de mémoriser quelques idées.

Regrouper des informations suivant leur sens permet de créer des associations entre eux, facilitant leur rappel. Cela permet aussi d'utiliser des schémas préexistants et d'intégrer ces informations dans un schéma en mémoire sémantique, facilitant fortement le travail de l'activation diffusante.

Dans tous les cas, **le découpage d'un cours en petites unités notionnelles associées progressivement les unes aux autres** est un bon moyen pour rendre le cours très facile à comprendre. Le résultat du découpage est très fin : songez que ce découpage se base sur un faible nombre d'informations, maintenues durant une dizaine de secondes. Ce niveau de granularité est rarement atteint dans les cours traditionnels.

Bref, *enseigner, c'est chunker*.

## Semantic/categorical chunking

Maintenant, je vais me servir de vous comme cobaye, cher lecteur. Rien de bien dangereux, rassurez-vous.

Essayez de vous souvenir de cette phrase, mot pour mot :

*le chasseur a tué un lapin, et a fini par le manger*

et celle-ci :

*lapin le a chasseur tué un, fini a et par manger le*

Plus dur pour la seconde, non ? Pourtant, les deux comportent les mêmes mots : le fait d'associer des chunks a permis de diminuer votre charge cognitive.

Autre exemple : lisez une liste de mots à des élèves, et demandez leur de rappeler les mots qu'ils viennent d'entendre immédiatement après avoir lu la liste. Dans ces conditions, les mots de la liste sont en mémoire à court terme. On remarque que classer les mots en groupes de mots appartenant à une même catégorie permet de faciliter le rappel d'environ 10 à 20%, voire plus.

Associer, catégoriser, comprendre, tout cela diminue fortement la charge cognitive. Cela permet à notre mémoire à court terme de stocker non seulement beaucoup d'informations, mais aussi les liens et associations d'idées entre ces informations.

C'est parce que la mémoire à court terme peut contenir des catégories, des schémas ou des morceaux de réseaux sémantiques. D'ailleurs, les théories récentes (modèle de Cowan, long term working memory, et retrieval structures model) pensent que la mémoire à court terme serait simplement une sous-partie des concepts activés en mémoire à long terme. Le superviseur attentionnel serait capable de sélectionner les concepts activés pertinents en focalisant l'attention dessus, ou en les inhibant s'ils ne sont pas pertinents.

La conséquence, c'est que le fait de regrouper des chunks suivant leur distance sémantique ou leur catégories permet de diminuer fortement la charge cognitive.

On peut aussi en déduire que plus un élève a acquis de connaissances, plus il pourra chunker son cours facilement : remplacer 25 informations par un seul chunk est clairement positif. Cela donne un argument supplémentaire pour les cours remplis, très complets, qui voient les choses en profondeur.

## Isolated element effect

Conjuguer chunking et découpage sémantique donne la technique du *scaffolding* : on aborde chaque notion en présentant les éléments essentiels, et on rajoute progressivement des détails, étape par étape. Cela demande de découper chaque concept en un chunk principal, auquel on rajoute progressivement des chunks plus petits. Cela donne une progression qui va du simple au complexe.

Prenons un exemple, pas forcément convaincant, mais qui montre le principe : l'apprentissage des polynômes du second degré. Une explication non-chunkée dirait qu'un polynôme du second degré est une équation de la forme :

$$a * (x^2) + b * x + c = 0.$$

Une explication chunkée découperait cette explication en deux morceaux : un terme égal à  $a * (x^2)$ , et un autre terme égal à un polynôme du premier degré de forme  $b * x + c$ , connu des élèves. L'ensemble est ensuite assemblé progressivement.

Cette technique a une conséquence pour l'apprentissage de concepts composés d'un grand nombre de sous-concepts fortement reliés entre eux : il vaut mieux aborder indépendamment les sous-concepts de base, et les assembler progressivement. En quelque sorte, le professeur doit jouer aux légos pour réduire la charge cognitive. En faisant cela, la compréhension est facilitée : c'est ce qu'on appelle l'*Isolated element effect*, découvert par Sweller et Chandler, dans leurs travaux sur la charge cognitive.

Donnons un exemple : les démonstrations qui utilisent des lemmes. Dans certains de mes cours d'université, les professeurs commençaient la démonstration du théorème, et l'interrompait pour démontrer les lemmes dont ils avaient besoin. Au lieu de faire cela, ils auraient du démontrer les lemmes avant de commencer la démonstration.

Autre exemple : les schémas ou cartes de géographie. Expérimentalement, on a plus de mal à retenir un schéma si celui-ci est trop chargé en détails, et limiter le nombre d'éléments est une bonne chose. Il vaut mieux limiter la quantité de concepts dans un schéma (pays pour une carte, organes pour un schéma de corps humain, etc) à environ 4 : la limite de la mémoire à court terme pour les concepts, ainsi que pour la mémoire visuo-spatiale.

Ainsi, au lieu de mettre beaucoup de choses sur un seul schéma, il vaut mieux décomposer le schéma en plusieurs sous-schémas ayant chacun peu d'éléments/d'informations. Par exemple, une carte des départements de la France a tout intérêt à être découpée en sous-cartes, une par région.

## Chunking des procédures

Ce qui est fait pour le cours magistral peut aussi être fait pour l'acquisition des procédures : on peut les chunker en petites unités procédurales pour mieux gérer la charge cognitive.

Par exemple, regardons l'apprentissage de la multiplication. Les professeurs de mathématique abordent celle-ci d'un seul bloc : d'abord on apprend les tables, et ensuite on apprend à multiplier deux nombres entre eux. Il n'y a pas d'étapes intermédiaires.

Un professeur doit chercher à découper l'algorithme de multiplication en sous-étapes qu'il combinera ensemble :

- Il commencera avec les tables de multiplication.
- Puis il expliquera comment multiplier un nombre par une puissance de dix.
- Puis il utilisera les deux étapes précédentes pour obtenir l'algorithme qui permet de multiplier un nombre à un chiffre par un autre nombre à plusieurs chiffres.
- Puis il réutilisera les deux étapes précédentes pour indiquer comment calculer un produit partiel.
- Et enfin, il terminera avec l'étape de multiplication proprement dite : calculer plusieurs produits partiels, et les additionner.

Chacune de ces étapes est acquise par de la pratique guidée et consolidée à grand coup d'exercices.

Autre solution pour réduire la charge cognitive : faire en sorte que les exercices soient relativement simples au début de la pratique autonome, et augmenter progressivement la difficulté.

## Séparation procédures et connaissances déclaratives

Une autre solution pour diminuer la charge cognitive consiste à séparer les connaissances théoriques des connaissances procédurales (comment faire).

### **Modality effect : visuel et auditif main dans la main**

Ensuite, le professeur peut utiliser le *modality effect*. Cet effet nous dit que le professeur doit utiliser adéquatement le bon support de transmission : visuel, oral, écrit, etc. Ces supports n'ont pas la même efficacité, et le choix du bon support peut grandement aider les élèves.

## Un bon schéma vaut mieux qu'un long discours

Les informations sont mieux retenues quand elles se basent sur des indices visuels : c'est le **Picture superiority effect**. Cela vient de la profondeur de traitement : les images contiennent plus d'indices de récupération.

Dans tous les cas, l'ajout d'informations visuelles dans un cours aide les élèves. Cela peut passer par l'utilisation de cartes mentales, de schémas au tableau, de projections Power Point, ou d'autres techniques du genre. Cela demande souvent d'utiliser le numérique et les technologies associées.

En bref, *un bon schéma vaut mieux qu'un bon discours*.

## Dire au lieu de lire

On peut aussi profiter des connaissances sur le fonctionnement de la mémoire à court terme verbale pour améliorer ses cours. Premier conseil : on mémorise mieux des mots quand on les lit à haute voix. Ce qui peut être utile pour apprendre ses listes de verbes irrégulier ou toute autre liste de mots. Utile aussi pour l'apprentissage des langues étrangère ou du français dans les petites classes.

Autre conseil : la parole est nettement mieux comprise que le texte. Cela vient du fait que le texte écrit passe dans la mémoire à court terme visuelle, avant d'être traduit en sons. La charge cognitive du texte est donc plus élevée, et le texte écrit utilise un peu d'attention pour être traduit.

Autant vous dire que ceux qui lisent des livres ou apprennent sur internet au lieu d'aller en cours sont un peu mal barrés. En conséquence, au lieu de faire lire un texte à vos élèves, lisez-le à haute voix. Par exemple, dans un cours de français, pré-lire le texte à haute voix avant de faire travailler les élèves dessus peut les aider à mieux comprendre le sens du texte. Ce conseil peut s'appliquer en cours de français, ou dans tout cours dans lequel la lecture de documents écrits a de l'importance (histoire, éducation civique, anglais, etc).

## On peut faire les deux

Une technique apparemment contre-productive donne de bons résultats : mixer les informations verbales et visuelles.

Première raison : notre cerveau dispose de plusieurs mémoires à court terme, dont l'une est spécialisée dans les informations orales/écrites, et l'autre dans la vision. En répartissant les informations entre ces deux MCT, on cumule les capacités des deux mémoires : on peut alors gérer plus d'informations, et faire des liens entre elles plus facilement. Comme quoi, ajouter du visuel dans un cours est une bonne tactique.

Si vous faites une présentation orale devant une classe, ou que vous créez un support de cours ou une feuille d'exercices : ajoutez des informations visuelles. Concrètement, cela passe par un usage de schémas au tableau, éventuellement par des photocopies, mais surtout : des présentations Powerpoint.

Si mixer des informations visuelles et verbale est une bonne idée, il ne faut pas le faire n'importe comment : représenter une information à la fois sous forme visuelle et sous forme verbale pourrait avoir des effets délétères sur l'apprentissage.

Certaines expérimentations semblent montrer qu'ajouter de la redondance peut augmenter la charge cognitive. Par exemple, présenter un texte à la fois sous forme écrite et orale sur un écran d'ordinateur pourrait réduire les performances d'apprentissage. Insi, il faudrait éviter de doubler les informations : c'est le **redundancy effect**.

Mais les expériences ne sont pas claires sur le sujet, et nombreux sont ceux qui se demandent si cet effet existe pour de vrai. Certaines expériences semblent aller dans ce sens, d'autres non. Bref, le débat n'est pas clos.

## **Split attention effect : ne pas disperser les informations**

Lorsque l'on mixe des informations visuelles et textuelles, ou que l'on veut organiser un document pédagogique à donner aux élèves, il faut faire attention à ne pas disperser les informations.

Si jamais plusieurs informations sont compréhensibles indépendamment, les séparer ne posera pas de problème.

Dans le cas contraire, l'élève devra alors visiter plusieurs endroits sur la feuille ou le manuel, avant de rassembler celles-

ci et de les traiter. Durant ce temps, l'élève devra maintenir toutes ces informations dans sa mémoire de travail, augmentant fortement sa charge cognitive.

Pour donner un exemple, il vaut mieux éviter de faire des légendes en dessous des schémas.

De même, le texte ne doit pas être assez éloigné de l'image qui correspond.

Expérimentalement, les élèves comprennent mieux quand le texte est incorporé aux schémas, et placé immédiatement à côté du dessin correspondant.

## **Worked-exemples effect : exemples travaillés**

Un humain peut résoudre des problèmes en utilisant un grand nombre de stratégies différentes. La psychologie cognitive a longtemps cherché à découvrir ces stratégies. Elle s'est aussi intéressée aux différences entre experts et novices dans un domaine, pour voir quelles étaient les différences entre ces deux catégories de personnes.

Les résultats de la recherche sont clairs : les stratégies de résolution de problèmes sont connues de tous. Les experts n'utilisent pas de stratégies de résolution de problèmes meilleures que celles des débutants. Toutes les stratégies de résolution de problèmes utilisées semblent carrément être innées, selon certains, et seraient des formes de connaissances « biologiquement primaire ». Selon certains, ces stratégies seraient innées, et l'on aurait pas besoin de les apprendre.

En réalité, les différences entre un expert du domaine et un novice sont dues à la grande quantité de schémas des experts, ainsi qu'à l'élaboration de leurs réseaux sémantiques. Là où un novice devra gérer une quantité d'informations phénoménales, l'expert reconnaîtra des situations déjà vues et les liera avec un schéma bien particulier en mémoire. Il pourra alors réutiliser ses connaissances antérieures et regrouper les informations du problèmes en chunks assez gros.

De plus, la qualité de ses réseaux sémantique permet à l'activation diffusante de faire son travail efficacement, et de faire des inférences correctes.

Un expert n'aura donc pas besoin de maintenir grand chose en MCT, ni même à réfléchir autant que le novice : il aura juste à charger deux gros schémas en MCT et à faire quelques associations simplistes pour trouver le résultat. Et ce phénomène suffit à expliquer toutes les différences entre experts et novices.

Cela a notamment été montré sur des joueurs d'échec. Si on les met face à une configuration de jeu courante, ils sont capables de retenir un plus grand nombre de pièces que les débutants, et leur performances en cours de jeu sont supérieures. Mais face à des configurations de jeu qui n'apparaissent presque jamais dans une vraie partie, leurs performances sont strictement identiques à celles des débutants.

Et ce principe s'adapte aussi à des choses assez diverses : lecture et compréhension de texte, conception de circuits électroniques, programmation, algèbre, géométrie, etc.

La conséquence, c'est que tenter d'apprendre les élèves à raisonner, à résoudre de problèmes, ou à former des élèves créatifs est impossible. Les phénomènes de transferts sont assez restreints, limités à des domaines très semblables. Mais apprendre des stratégies de résolution de problèmes ou de raisonnement générique est impossible. On peut former des experts dans un domaine en leur transférant des connaissances particulières.

On peut exploiter cette technique dans un cours, si on veut que nos élèves apprennent facilement à résoudre des exercices. Un élève sans connaissance préalable devra réfléchir intensément pour résoudre le problème : le contenu de sa mémoire à long terme ne peut pas être utilisé pour diminuer sa charge cognitive. Il faut éviter cela.

## **Exemples travaillés**

La solution, c'est de transmettre aux débutants les connaissances que les experts utilisent pour diminuer leur charge cognitive. La solution consiste à ajouter une phase de **pratique guidée** dans le cours. Dans cette pratique guidée, le professeur résout des exercices devant les élèves : il pense à haute voix, montre explicitement comment il résout le problème, montre bien quelles sont les étapes de résolution et comment il les enchaîne, il explicite ses raisonnements, etc.

Cela permet de transférer des schémas de résolution de problèmes ou des exemples de problèmes résolus, que l'élève pourra réutiliser lors de la pratique autonome, réduisant ainsi sa charge cognitive.

Cette pratique guidée doit être systématique : on y a droit à chaque cours. Et enfin, autre innovation : elle précède la pratique autonome.

Cette pratique guidée est une des plus grosse différence avec la pédagogie traditionnelle. Dans la pédagogie traditionnelle, il n'y a pas de transferts positifs, et toute repose sur la répétition d'exercices. Seul problème : la charge cognitive de ces techniques est fortement élevée. Et ne parlons pas des pédagogies actives, qui cherchent à faire pratiquer les élèves le plus tôt possible et dont le principe même est contraire à l'esprit de la pratique guidée : il n'y a pas de transferts.

Dans les années 1985 et 1987, Sweller and Cooper ont testé l'efficacité des exemples travaillés. La première catégorie de groupes d'élèves passait directement aux exercices après le cours. L'autre, subissait une phase de pratique guidée. L'effet était sans appel : la pratique guidée a augmenté de manière significative les résultats des élèves.

De nombreuses expériences faites sur le sujet vont toutes dans le même sens. Cela fonctionne, que ce soit en mathématique, en physique, en biologie, géologie, en programmation, en électronique, pour l'apprentissage d'un second langage, pour l'apprentissage de la musique, pour apprendre à reconnaître le style d'un peintre en particulier, pour l'apprentissage de la littérature anglaise, et j'en passe.

De plus, les recherches montrent clairement que les meilleurs professeurs sont ceux qui utilisent un temps assez long pour la pratique guidée. Une étude réalisée dans le secondaire, en France, a montré que les meilleurs enseignants étaient ceux qui passaient le plus de temps à faire de la pratique guidée : les "bons" enseignants font 24 minutes, tandis que les plus mauvais font dans les 11 minutes.

## **Expertise reversal, problem completion et guidance fading effects : la pratique, ça sert !**

Ces expériences ont été répétées de nombreuses fois, pour vérifier quelque chose : quand cette pratique guidée devient inutile ?

La théorie de la charge cognitive et les expériences, nous disent que celle-ci devient inutile lorsque l'étudiant a acquis suffisamment de schémas de résolution. Répéter ces schémas ne sert alors pas à grand chose : il vaut mieux laisser l'élève automatiser ces schémas à grands coups d'exercice et de pratique autonome. Plus un étudiant a de connaissances préalables, moins les exemples travaillés seront utiles. C'est ce qu'on appelle l'*expertise reversal effect*.

Le mieux est que la transition entre exemples travaillés et pratique autonome soit la plus douce possible. Pour cela, on peut utiliser des exemples partiellement travaillés dans lesquels le professeur résout partiellement les exercices présentés et laisse les élèves finir. Expérimentalement, on peut constater que cela donne de bons résultats sur pas mal d'élèves. C'est ce qu'on appelle le *problem completion effect*.

On peut alors combiner les deux effets précédents. Après une phase de pratique guidée pure, le professeur poursuit par une phase de pratique guidée partielle, avant de passer progressivement à une phase de pratique autonome. En utilisant cette technique les résultats sont meilleurs : c'est le *Guidance fading effect*.

## **Goal-free effect : privilégier les problèmes ouverts**

Le goal-free effect est un effet qui porte sur des problèmes particuliers. Ceux-ci sont définis par deux choses : un état final, un état initial. Le but du problème est de trouver une série d'étapes permettant de passer de l'état initial à l'état final. Chacune de ces étapes correspond à l'application d'une petite procédure, qu'on appelle un opérateur.

Pour donner un exemple, prenons un problème de physique simple.

*Un train part d'une gare A, pour aller à une gare B.  
Sur le trajet entre les gares, ce train va à une vitesse constante de 50 km/h.  
Il parcourt 200 kilomètres.  
Trouvez la durée du trajet.*

Dans ce cas, l'état initial est constitué des données du problème.

L'état final est celui qui donne la durée du trajet.

L'étape à effectuer pour trouver le résultat est simple : il s'agit d'une simple division entre la longueur du trajet, et la vitesse du train. L'opérateur utilisé est donc une simple division.

La difficulté de ce genre de problème est de se le représenter mentalement, et d'en déduire une modélisation mathématique permettant d'utiliser les outils vus en classe.

Autre exemple plus complexe :

*Un train part d'une gare A, pour aller à une gare B.  
Ce train part de la gare A à une vitesse de 50 km/h.  
Il arrive à la gare B à une vitesse de 100 km/h.  
Sur le trajet, son accélération est constante.  
La durée du trajet est de 2 heures.  
Trouvez la distance parcourue lors du trajet.*

Dans ce cas, le problème est plus complexe: une seule étape ne suffit pas pour le résoudre. Il faut calculer accélération, en déduire la vitesse en fonction du temps, et intégrer en fonction du temps pour obtenir la distance parcourue. Cela fait trois étapes.

Dans ce cas, l'élève doit découvrir quelle série d'étape permet de déduire l'état final, le résultat. Chaque opérateur sera donc un petit calcul mathématique, permettant de calculer la vitesse, l'accélération, etc.

La difficulté est alors non seulement de bien modéliser mentalement le problème, mais aussi de découvrir la suite d'étapes, de la planifier convenablement.

## Means-end analysis

Pour résoudre ce genre de problème, l'élève va utiliser une stratégie qu'on appelle la *means-end analysis*. Il en existe deux types.

Dans la première stratégie, il va partir de l'état initial, va choisir un opérateur et va l'appliquer. Il obtiendra un nouvel état, qu'il comparera à l'état final. Si les deux sont identiques, alors il a trouvé le résultat. Sinon, il va poursuivre récursivement : il va de nouveau choisir un opérateur à appliquer sur ce nouvel état, et ainsi de suite.

Dans la seconde stratégie, l'élève va faire la même chose, mais va partir de l'état final, pour poursuivre à rebours.

Évidemment, cette stratégie marche aussi très bien en découpant le problème en sous-problèmes.

Cette stratégie est utilisée très couramment dans la vie courante. Par exemple, c'est cette stratégie que vous utilisez quand vous voulez savoir à quelle heure partir de chez vous pour être à l'heure à un rendez-vous. Dès que l'on ne dispose pas de méthode de résolution pour un problème, nous allons naturellement utiliser cette stratégie pour tenter de le résoudre. Certains pensent que cette stratégie sera une technique innée, un « savoir biologiquement primaire ».

On remarque une chose : cette means-ends analysis ne s'applique pas quand on dispose de la méthode de résolution pour un problème. Par exemple, si jamais un élève connaît la méthode de résolution d'une équation du second degré, utiliser le goal-free effect ne servira à rien pour résoudre des équations du second degré. Faites bien attention à correctement identifier les exercices qui pourront profiter de cet effet.

Comprendre comment rendre celle-ci plus efficace du point de vue de la charge cognitive est le but du goal-free effect. Comme je l'ai dit, il y a deux types de means-end analysis: une qui part du but, et une autre qui part des données du problème.

La littérature scientifique sur le sujet est claire : les débutants dans un domaine ont tendance à partir du but, tandis que les experts dans un domaine partent des données du problème. Cela a été montré dans un environnement scolaire par les études de Larkin, McDermott, Simon, et Simon (1980a, b), sur des exercices de physique.

L'élève novice part du but : il commence par résoudre le problème, puis s'aperçoit qu'il faut résoudre un sous-problème, et continue ainsi de suite. Il résout les sous-problèmes quand ils surviennent. Seul problème, cette méthode demande de maintenir beaucoup de choses en MCT : le but, et tous les sous-problèmes.

Les experts commencent partent des sous-problèmes les plus basiques, et remontent progressivement vers les sous-problèmes les plus haut.

Cela vient du fait qu'un expert a vu beaucoup de problèmes similaires lors de son apprentissage. Il peut ainsi réutiliser ce qu'il a appris : il sait comment combiner les données du problème pour obtenir le résultat, et n'utilise donc pas vraiment de means-end analysis. Par contre, un débutant n'a rien dans sa mémoire pour savoir comment planifier les étapes pour résoudre le problème.

Or, ces deux stratégies ne sont pas neutres du point de vue de la charge cognitive. La stratégie utilisée par les débutants, qui consiste à utiliser la means-end analysis partant du but a tendance à surcharger rapidement la mémoire de travail. Dans ces conditions, les élèves gaspilleront toute leur attention pour résoudre le problème, et rien ne passera en mémoire à long terme. Les exercices seront difficiles à résoudre et auront peu d'efficacité lors de l'apprentissage.

Prenons un exemple, on donne cet exercice à un élève novice :

$y = x + 6$ ,  $x = z + 3$ , et  $z = 6$ .  
Trouvez la valeur de  $z$ .

Un novice va se focaliser sur le but : trouver la valeur de  $y$ .

Il va d'abord lire  $y = x + 6$ . Il va alors chercher à résoudre le sous-problème : trouver  $x$ .

Il va ensuite lire  $x = z + 3$ . Il va alors chercher à résoudre le sous-problème : trouver  $z$ .

Enfin, il va lire  $z = 6$ , et il va rembobiner le tout.

Le fait de rembobiner le tout signifie qu'il faut maintenir beaucoup de choses en mémoire de travail.

En comparaison, l'expert va aller dans le sens inverse, et va substituer immédiatement les valeurs dans les équations, réduisant fortement le nombre de choses à maintenir en mémoire.

## Exercices goal-free

Comment faire pour que le novice utilise la bonne méthode ? La solution est radicale : supprimer le but ! Le fait de ne pas donner de but explicite aide les novices à résoudre les problèmes correctement : c'est ce qu'on appelle le **Goal-free effect**.

Par exemple, dans l'exemple avec l'équation, on aurait un énoncé du style :

$y = x + 6$ ,  $x = z + 3$ , et  $z = 6$ .  
Trouvez le maximum de valeurs.

Ou encore :

*Une particule a une vitesse de 2 mètres par seconde à un instant  $t$ .  
Elle accélère de 5 mètres par secondes carrés durant 25 secondes.  
Trouvez un maximum de valeurs.*

Des expériences (Owen and Sweller, 1985) ont comparé deux groupes d'élèves de même niveau qui recevaient des exercices de géométrie : un groupe avait un énoncé goal-free, et l'autre un exercice conventionnel. D'autres expériences complémentaires ont été faites, encore une fois sur des exercices de géométrie, montrant encore l'efficacité de la stratégie goal-free.

Bobis, Sweller, and Cooper (1994). Sweller, Mawer, et Ward (1983), ont testé l'efficacité de cette technique sur des exercices de physique. La différence était en faveur du groupe goal-free dans tous les cas.

## Métacognition

# Béhaviorisme

## **Feedback**

### **Évaluation formative**

Dans une salle de classe, les renforçateurs sont assez variés, qui portent presque tous sur le comportement de l'enseignant. Mais on trouve aussi les notes, bien que celles-ci soient des renforçateurs assez mauvais : elles ne sont pas assez précises. Elles ne permettent pas renforcer un comportement bien précis, vu qu'elles jugent des copies complètes. Il faut donc trouver des renforçateurs plus précis.

Pour cela, on peut utiliser l'**évaluation formative**. Cette évaluation formative consiste à dire explicitement aux élèves quelles sont leurs forces et leurs faiblesses. Cette forme d'évaluation consiste à donner un *feedback* sur les compétences et connaissances de l'élève.

Par exemple, au lieu de mettre une note sur une copie, on peut donner des commentaires normés sur la copie, au lieu d'une note. Ce commentaire devrait expliciter clairement, pour chaque copie :

- ce qu'il a su faire,
- ce qu'il a compris,
- et ce qui est su,
- et ce qu'il n'a pas su faire,
- ce qu'il ne sait pas,
- comment il aurait du traiter l'exercice,
- et ce qu'il doit revoir, afin d'indiquer aux élèves les difficultés qu'il a, afin qu'il sache quoi faire pour y remédier (ce qui est plus efficace que de simplement dire de bosser à un élève).

L'autre avantage de cette forme d'évaluation est qu'elle permet aux élèves de développer leurs compétences de méta-mémoire. La méta-mémoire, c'est simplement la mémoire sur la mémoire : savoir ce que je sais, et ce que je ne sais pas. Plus un élève a une bonne méta-mémoire, plus il sait quelles sont ses lacunes, et quelles sont ses points forts. Cette méta-mémoire est fortement corrélée à la réussite scolaire ultérieure.

On remarquera que les notes ne permettent pas de développer ces capacités. Quand on a 5, ou 15, on sait qu'on est bon ou mauvais, mais on ne sait pas ce qu'on a raté.

Actuellement, cette évaluation formative est utilisée sous une forme faible d'évaluation formative : l'approche par compétence. Malheureusement, nombreux sont les enseignants réticents à cette forme d'évaluation.

Mais les recherches sur le sujet sont claires :

Diverses méta-analyses le prouvent : l'évaluation formative permet de mieux faire progresser les élèves.

- Black and Wiliam's (1998) : revue de 250 études sur le sujet ;
- Terrance Crooks (1988) : revue d'une partie de littérature scientifique sur le sujet.

## Motivation

Enfin, nous arrivons au dernier passage obligé dans toute pratique pédagogique qui se respecte : l'évaluation. Celle-ci peut prendre plusieurs formes, qui ont toutes des effets différents sur la réussite des élèves.

De ce point de vue les notes seules sont assez catastrophiques.

Pour comprendre pourquoi, il faut savoir que nombre d'élève travaillent pour la note, plus que par motivation pour la matière. C'est un fait connu de nombreux enseignants, qui considèrent que les notes sont une bonne carotte pour les élèves, qui les force à travailler. Combien d'entre eux croient encore que la note sert de salaire aux élèves... Mine de rien, même des enseignants traditionalistes médiatiquement connus en parlent : Brighelli parle d'élèves consommateurs de savoir, par exemple.

Et bien cela s'explique facilement : les notes sont une source de **motivation extrinsèque**. La motivation extrinsèque est celle qui pousse à faire quelque chose en échange d'une récompense.

Elle est à opposer à la **motivation intrinsèque**, qui est celle de nos intérêts, de nos passions, etc. Celle-ci est clairement meilleure en terme d'apprentissage, et est plus robuste sur le long terme.

Expérimentalement, la psychologie sociale a clairement montré que la motivation extrinsèque disparaît dès que les récompenses ne sont plus données, ou que l'élève s'y habitue : on observe des phénomènes d'extinction lors de la suppression des récompenses. Par contre, la motivation intrinsèque perdure longtemps, sous certaines conditions.

De même, cette motivation intrinsèque peut rapidement remplacer la motivation extrinsèque. Exemple : un élève intéresse par une matière ou un sujet finira par perdre son intérêt pour celle-ci si en fonction de ses notes.

Et enfin, ceux qui ne reçoivent pas les récompenses (les bonnes notes), et qui reçoivent des punitions (mauvaises notes) ont tendance à perdre l'intérêt et la motivation de travailler à force de mauvaises notes, par des phénomènes d'impuissance acquise. Bilan : ceux qui ne reçoivent pas de bonnes notes ont tendance à se désengager de leur travail et finissent par détester les matières dans lesquels ils sont mauvais.

Résultat : une inappétence pour ceux qui ne reçoivent pas les renforçateurs (notes faibles), ainsi que des comportements de recherche d'augmentation des notes basés sur la triche, le bachotage, et les trucs et astuces utiles pour avoir de bonnes notes ou la moyenne. Ce qui est franchement délétère sur la motivation et les résultats.

Les approches basées sur l'évaluation formative sont légèrement moins concernées par ce phénomène, ce qui est un argument de plus en faveur des approches par compétences.

Autre détail : si on veut motiver les élèves, il faut agir à la source de la motivation intrinsèque.

La motivation a diverses origines :

### ***Sentiment d'auto-efficacité personnelle***

Première origine : le **sentiment d'auto-efficacité personnelle**. En clair, pour motiver les élèves, il faut leur donner l'impression qu'il réussissent. On est motivé parce que l'on est bon dans la matière, pas l'inverse.

Cela passe évidemment par l'usage de pédagogies adaptées, qui font réussir les élèves. Mais pas seulement : l'usage de renforcement positif, de feedback, peut aussi aider.

## ***Attribution de la performance***

Diverses analyses statistiques ont montré que plus un élève croit que l'intelligence et la réussite scolaire n'est pas innée, plus il progresse. Après tout, rien de plus évident : plus on dit à un élève qu'il peut progresser et que sa situation n'est pas une fatalité, plus il progressera... Il fallait bien des études statistiques de haute volée pour en être sûr !

En conséquence, certains se sont dit qu'engrainer les élèves à croire en la non-innité de l'intelligence leur permettrait de progresser. C'est le principe de l'***attribution training***.

## **Bilan**

Cet article vous a appris beaucoup de choses, qu'il serait temps de résumer.

Un cours idéal est un cours structuré, dont le déroulement est toujours le même :

- une phase de mise en place ;
- des rappels ;
- un cours magistral ;
- une phase de pratique guidée ;
- une phase de pratique autonome ;
- une phase de *Feedback*.

Ce déroulement ressemble légèrement à un cours traditionnel. Mais au-delà de ces ressemblances de surface, les différences sont nombreuses, comme nous avons pu le voir.

La phase de mise en place :

- explique aux élèves ce qu'il vont apprendre durant le cours ;
- quels sont les sujets qu'ils vont aborder ;
- quel est le but de la leçon ;
- ce que les élèves seront capables de faire après la leçon ;
- etc.

La phase de rappel sert à réveiller les pré-requis, à les pré-activer.

Le cours magistral :

- est construit autour d'un plan hiérarchique, fortement basé sur des catégories ;
- est chunké, découpé en petites unités notionnelles assemblées progressivement ;
- sépare les connaissances théoriques des savoirs-faire ;
- est fortement basé sur la compréhension ;
- admet peu et relie toute nouvelle connaissance à des connaissances antérieures ;
- explique pourquoi les choses sont ce qu'elles sont ;
- contient beaucoup de connaissances et est assez complet ;

- donne de nombreux exemples et des cas d'utilisations variés ;
- et fait en sorte de ne pas interférer avec des apprentissages ultérieurs.

La pratique guidée :

- commence par des exemples totalement travaillés ;
- avant de laisser la place à des exemples partiellement travaillés ;
- et finit par laisser place à la pratique autonome.

La pratique autonome :

- utilise des exercices goal-free ;
- chunke ses exercices et les procédures à apprendre ;
- montre de nombreuses applications diverses d'un même concept ou d'une même procédure (effet de contexte) ;
- utilise la répétition d'élaboration et/ou de rappel ;
- utilise beaucoup de questions de compréhension ;
- demande aux élèves de faire des résumés, de ré-expliquer par eux-même ce qu'ils ont compris ;
- n'utilise pas le sur-apprentissage ;
- n'utilise pas de répétition de maintien.

Les supports pédagogiques :

- utilisent beaucoup de visuel, de schémas, d'images, etc ;
- mélangent informations visuelles et sonores ;
- évitent au maximum la redondance ;
- incorporent le texte au plus près des représentations visuelles correspondantes.

Les devoirs :

- sont fréquents ;
- sont assez rapides à faire, et ne doivent pas prendre trop de temps ;
- sont relativement faciles ;
- sont plus basés sur la répétition d'élaboration et de rappel que sur la répétition de maintien ;
- demandent de faire des résumés, des cartes mentales, des synthèses, des exposés, etc.

Pour ceux qui souhaiteraient se renseigner sur le sujet, je recommande la lecture du livre « Cognitive load theory », qui explique dans le détail la théorie de la charge cognitive, et qui montre quelques expériences de base de la théorie.

Je vous conseille aussi de regarder les ouvrages écrits par Alain Lieury, qui sont tous d'excellente facture, et méritent vraiment le coup d'œil.

TODO : améliorer partie sur la mémoire, et notamment sur la MLT

Plan :

- mémoire explicite
- sémantiques
- lexicale (différence décodage / sens, double codage et dyslexie, etc)
- mémoire implicite
- conditionnement (opérant et pavlov) : détailler
- habiletés motrices et cognitive (lecture) : ne pas détailler
- perceptive : ne pas détailler
- autres : voies réflexes

Fusionner partie transferts et connaissances d'une manière qui soit crédible, pour éviter les doublons (notamment avec les analogies. Ajouter trucs sur analogies (bien pour ressemblance, attention aux différences) → parler du transfert analogique.

- Apprentissage de la lecture
- Apprentissage de la numération
- Compréhension de texte
- apprentissage par cœur : dates, cartes, départements, etc.

Modèles de compréhension de texte, et schémas d'histoire + expériences tirée du premier chapitre de « le niveau monte », par Establet et al. - expériences sur textes de Marcel Pagnol.