

## Chapitre 6. L'intégration multimédia

<b>6.1. Description, dépicition et modèles mentaux: aperçus notionnels</b>	<b>139</b>
<b>6.2. Illustrations et compréhension de textes</b>	<b>142</b>
6.2.1. Illustrations et compréhension des descriptions d'objets complexes	143
6.2.2. Différentes illustrations favorisent différents aspects du modèle mental	144
6.2.3. La compréhension du texte illustré sollicite le calepin visuospatial	145
<b>6.3. Les facteurs matériels d'intégration multimédia</b>	<b>146</b>
6.3.1. Introduction	146
6.3.2. Effets de la disposition spatiale et temporelle	147
6.3.3. Effets de la cohérence référentielle	149
<b>6.4. Différences individuelles et développement des capacités d'intégration multimédia</b>	<b>150</b>
6.4.1. Le rôle des capacités de compréhension	150
6.4.2. Le rôle des capacités visuospatiales	151
6.4.3. Aptitude visuospatiale et gestion cognitive de l'activité	153
6.4.4. Le développement cognitif des capacités d'intégration multimédia	154
<b>6.5. Les animations graphiques complexes</b>	<b>155</b>
6.5.1. Quand la pratique résiste à la théorie	155
6.5.2. Animations et consignes de traitement: une étude exploratoire	157
6.5.2.1. Introduction et méthode	157
6.5.2.2. Résultats préliminaires	159
6.5.2.3. Discussion et perspectives	161
6.5.3. Quelques pistes de réflexion pour l'ergonomie des animations	162
<b>6.6. Synthèse et conclusion</b>	<b>164</b>

### 6.1. Description, dépicition et modèles mentaux: aperçus notionnels

Notre système cognitif semble traiter de façon distincte les informations verbales et les informations imagées. Cette distinction élémentaire ne peut pas surprendre si l'on se souvient que le langage n'est pas une forme de représentation immédiatement disponible aux humains à partir de leur équipement neuroanatomique. La maîtrise du langage oral demande plusieurs années d'éducation, et celle du langage écrit plusieurs années supplémentaires (et même de nombreuses années pour certains de ses aspects, ce dont j'espère avoir convaincu le lecteur dans les Chapitres 2, 3 et 4). La spécificité des traitements verbaux et imagés se maintient à l'âge adulte puisque, si l'on en croit la théorie de Baddeley (1986), la mémoire de travail serait un système composite, constitué de modules plus ou moins indépendants: un centre exécutif chargé des décisions opérationnelles (orientation de l'attention, attribution des ressources, contrôle) et deux sous-systèmes esclaves: la boucle phonologique et le calepin visuospatial. Ces deux sous-systèmes seraient spécialisés dans le maintien et le traitement de l'information verbale et imagée, respectivement. Une telle approche, de mieux en mieux validée par les données de la neuropsychologie, présente des conséquences importantes pour certaines notions classiques en psychologie, telle que celle d'attention. En effet, on ne peut plus dans une telle perspective considérer l'attention comme une dimension unitaire.

Je ne m'attarderai pas sur cette question, ni sur la controverse sur le format des représentations entre, d'une part, les propositionnalistes (pour qui grosso modo tout est verbal) et, d'autre part, les partisans de représentations analogiques ancrées dans l'expérience perceptive et/ou émotionnelle (pour un aperçu de l'état actuel de ce débat Cf. Schnotz, in press). J'en retiendrai seulement la thèse de Kintsch (1998, chapitre 2) selon laquelle il s'agit bien plus d'un débat sur les niveaux de description des représentations cognitives ("comment la science représente l'esprit") que d'un débat sur l'existence ou non de différents formats de représentation ("comment l'esprit représente le monde"). Il est en effet hautement probable que notre mémoire à long terme ressemble bien plus au bric à brac d'un chiffonnier qu'à l'étagère propre et bien garnie d'un linguiste. Ainsi, même un propositionnaliste avéré comme Walter Kintsch admet sans aucune difficulté l'existence des images mentales, dont il suggère simplement qu'on peut les décrire approximativement sous forme de propositions et simuler ainsi efficacement le fonctionnement du système cognitif sans avoir à manipuler différents formats de représentation.

Quoiqu'il en soit, l'acquisition du langage apporte à l'individu la possibilité de multiplier les formes de représentation d'un même référent. On peut ainsi nommer les objets et produire des images mentales d'objets dont on lit ou entend la dénomination (Denis, 1989). On sait depuis les travaux de Paivio et ses collègues dans les années 1970 que le double codage intentionnel est un puissant moyen de mémorisation. Les travaux récents de psychologues inspirés de la théorie des modèles mentaux (Johnson-Laird, 1983) sont venus considérablement affiner ce constat. Nous y reviendrons dans la suite de ce chapitre. Mais qu'en est-il des propriétés générales des représentations (externes) verbales et imagées?


Schnotz (in press) examine sous un angle sémiotique la distinction entre les représentations descriptives et dépicitives. Il définit les représentations descriptives comme des suites de symboles, dont le langage naturel est bien entendu le meilleur exemple (quoique pas le seul): les objets sont représentés par des noms, leurs propriétés par des adjectifs, leurs relations par des verbes et des prépositions. Les représentations dépicitives

sont quant à elles des "*configurations spatiales dans un espace fonctionnel qui représentent une configuration dans un autre espace*". Les représentations dépicatives reposent sur des analogies structurales: la "configuration cognitive" est similaire à la configuration réelle.

Il n'y a pas de différence, selon Schnotz, entre les images concrètes et les icônes logiques. Dans le premier cas, l'analogie consiste en la similarité perceptive des configurations (dessin d'une école), dans l'autre l'analogie repose sur des relations plus abstraites (la métonymie par exemple, dans le cas du panneau du code de la route représentant des silhouettes qui se tiennent la main). Cette position diffère de celle d'autres auteurs, comme Hegarty, Carpenter et Just (1996) qui établissent des distinctions sémiotiques entre les images (photos, dessins...), les schémas (dispositifs techniques) et les graphiques (camemberts, histogrammes...). Pour ces auteurs, ces trois types de dépications correspondent à des référents de nature différente, respectivement concret, abstrait et relationnel. La représentation repose sur l'isomorphie (similitude) dans le cas des images, des conventions graphiques dans le cas des schémas et des principes cartésiens dans celui des graphiques.

Pour Schnotz (in press) les représentations descriptives et dépicatives se distinguent sur un certain nombre de dimensions (Tableau 6.1).

Tableau 6.1. Les caractéristiques sémiotiques des représentations descriptives et dépicatives (d'après Schnotz, in press).

	DEPICIONS	DESCRIPTIONS
Exemple		Une maison Avec une porte, le toit est au dessus de la porte Une allée y conduit en zigzagant Elle me rappelle la maison des voisins quand j'étais enfant (etc.)
Codage des relations	Représentations intrinsèques Propriétés structurales inhérentes	Représentations extrinsèques Utilisent des symboles pour représenter les propriétés structurales
Puissance représentative	restreinte	plus grande (négation, disjonction)
Efficacité computationnelle	élevée	Plus faible (ex: géométrie)
Complétude	Plus complètes (intégrées)	Moins complètes (componentielles)
Robustesse	Plus robustes	Moins robustes
Cohésion	A priori plus cohésives	A priori moins cohésives

Tout d'abord, les représentations descriptives et dépicatives se distinguent par la façon dont elles codent les relations entre les parties, aspects ou constituants du référent. Dans les représentations descriptives, ces relations sont représentées explicitement par des symboles particuliers (verbes, adverbes, prépositions...). Elles semblent être construites de l'extérieur, ce qui permet de qualifier ces représentations d'extrinsèques. Dans le cas des

représentations dépicatives, les relations sont contenues implicitement dans la correspondance structurale (concrètes ou abstraites) entre l'image et l'objet qu'elle représente. On peut donc parler de représentations intrinsèques.

Représentations descriptives et dépicatives se distinguent ensuite par leurs propriétés computationnelles: les descriptions sont générales et abstraites, alors que les dépications sont plutôt concrètes et spécifiques (le mot chien peut désigner la catégorie générique des chiens alors que l'image sera toujours a priori celle d'un chien particulier). La "puissance représentative" des descriptions est donc plus élevée, en ce qu'elles permettent de représenter des généralités, des disjonctions ou des négations, ce qui n'est possible qu'avec des dépications de type particulier (par exemple, les symboles conventionnels). En revanche, les dépications possèdent un pouvoir computationnel plus élevé, c'est-à-dire qu'elles permettent de produire une plus grande variété d'inférences. Elles sont également plus complètes, dans la mesure où les différentes propriétés des objets y sont intrinsèquement liées (par exemple, la taille, la forme et l'orientation), alors que les descriptions relèvent toujours de sélections a priori. Elles sont plus robustes (enlever une partie d'une image ne remet pas nécessairement en cause son pouvoir représentationnel), et plus cohésives: le déplacement d'un point dans un graphique entraîne automatiquement la mise à jour de ses relations spatiales avec les autres objets, alors que le changement d'une proposition dans une phrase peut poser de gros problèmes de mise à jour (Cf. chapitre 1, sur la mise à jour des modèles mentaux textuels).

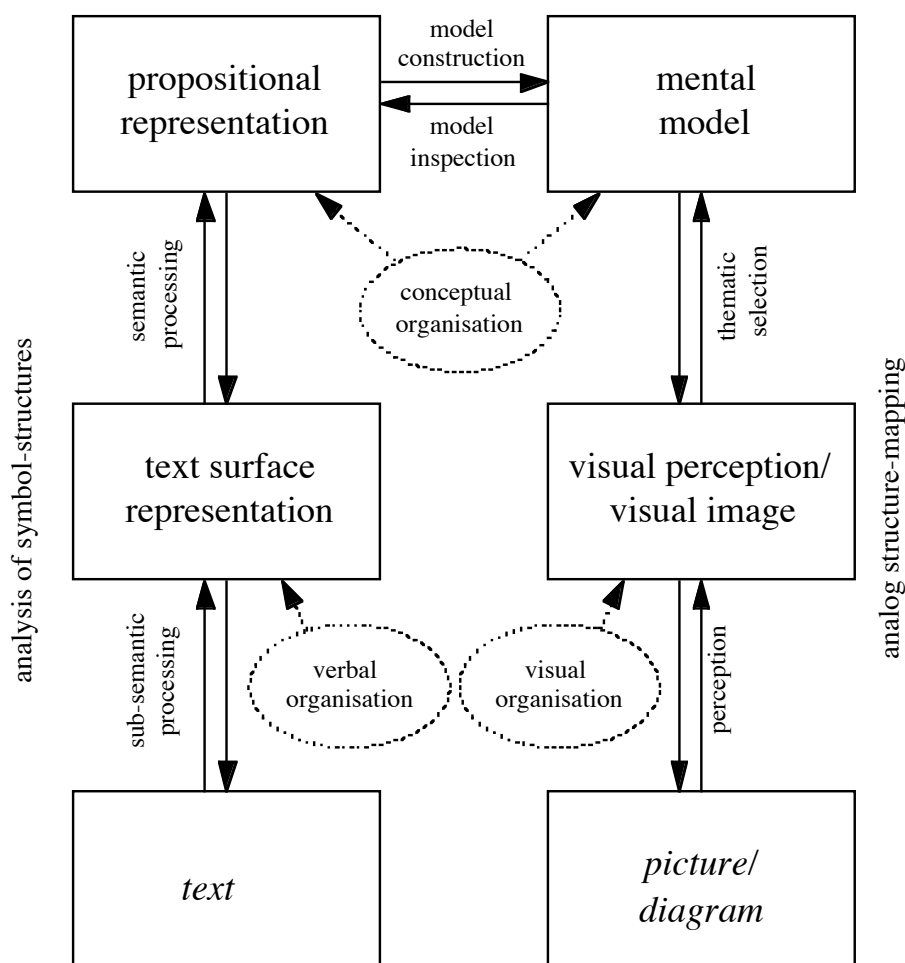


Figure 6.1. Le modèle d'intégration texte-image de Schnotz (in press).

Ces distinctions sémiotiques étant posées, Schnotz défend l'idée que d'un point de vue psychologique, les représentations internes construites à partir de textes et d'images reposent sur des processus similaires. La compréhension du texte passe par l'extraction des propositions, mais n'est achevée que lorsque le sujet intègre les propositions au sein d'un modèle mental ou modèle de situation (Figure 6.1, Cf. également Chapitre 1).

Or, entre autres caractéristiques, les modèles mentaux sont analogiques, c'est-à-dire qu'ils comportent des correspondances structurales avec les "objets" ou situations représentées (Glenberg, Meyer & Lindem, 1987; Hegarty & Just, 1993). A ce niveau les correspondances ne se limitent plus à la simple similitude perceptive, elles peuvent être d'un niveau plus abstrait (par exemple la force, la vitesse ou la probabilité). De plus, nous l'avons vu au chapitre 1, les modèles mentaux sont génératifs et actifs: à partir d'un modèle mental constitué, le sujet peut extraire de nouvelles informations, qui guideront en retour l'extraction et l'intégration des nouvelles propositions.

On sait que l'encodage des images est un processus rapide, direct et relativement indépendant des connaissances. Cependant, selon Schnotz, pour "comprendre" une image il faut également mettre en œuvre des processus cognitifs d'ordre supérieur. Par exemple, l'examen de traits perceptifs particuliers peut conduire à la production de propositions qui codent certains aspects des images. Ces propositions en activent d'autres qui permettent l'interprétation de l'image à l'aide des connaissances antérieures. A ce stade, l'expertise dans le domaine est bien évidemment critique, comme l'ont montré par exemple les travaux de Lowe dans le domaine des cartes météorologiques (Lowe, in press; voir plus bas). Il peut donc y avoir mauvaise compréhension d'une image par défaut d'analyse conceptuelle, comme il peut y avoir défaut de compréhension d'un texte par défaut d'intégration au sein d'un modèle mental.

## 6.2. Illustrations et compréhension de textes

On sait depuis longtemps que l'image peut jouer un rôle positif dans la compréhension d'un texte, surtout dans le cas où le contenu du texte est peu familier. Dans l'expérience bien connue de Brandsford et Johnson (1972, Cf. Chapitre 1, Figure 1.1), la présentation préalable d'une image facilitait la compréhension orale d'un texte rédigé de façon confuse et abstraite. Brandsford et Johnson invoquaient l'information globale procurée par l'image, le fait qu'elle donne un contexte dans lequel les informations du texte seront intégrées, pour expliquer ces effets. Un examen attentif des effets révèle plusieurs autres phénomènes intrigants (Tableau 6.2).

Tableau 6.2. Compréhension auto-estimée et rappel des éléments d'un texte selon la condition de présentation (d'après Brandsford et Johnson; 1972, expérience 1).

	Sans image (1 lecture)	Sans image (2 lectures)	Image après	Image partielle	Image avant	Score max.
Compréhension (estimée)	2,30	3,60	3,30	3,70	6,10	7,00
Rappel	3,60	3,80	3,60	4,00	8,00	14

Ni la répétition du texte seul, ni la présentation de l'image après le texte, ni même la présentation des constituants de cette image ("image partielle") ne permettait d'atteindre le

même niveau de compréhension subjective et de rappel. Comment expliquer précisément ces effets? Comme le rappellent Glenberg et Langston (1992), les effets facilitateurs des illustrations n'ont reçu d'interprétation détaillée que très récemment, grâce notamment aux théories de la mémoire de travail et des modèles mentaux (Baddeley, 1986; Johnson-Laird, 1983; Cf. aussi Gyselinck, 1996). Les recherches menées depuis 10 ans ont permis de préciser l'effet des illustrations sur la compréhension des relations entre éléments d'un objet ou dispositif spatial dynamique (en mouvement), mais aussi dans certains cas sur la compréhension des constituants de ce dispositif. Elles ont également précisé les mécanismes en mémoire de travail responsable de la facilitation observée (Cf. Gyselinck & Tardieu, 1999, pour une revue récente).

### 6.2.1. Illustrations et compréhension des descriptions d'objets complexes

Hegarty et Just (1993) présentent deux expériences qui précisent les processus par lesquels l'illustration améliore la construction des modèles mentaux. Selon Hegarty et Just, la compréhension de textes descriptifs est particulièrement facilitée si le texte s'accompagne d'illustrations (schémas) qui montrent les relations importantes entre parties de l'objet ou aspects de la situation. Dans une première expérience, ils examinent la "contribution" respective d'un texte et d'un schéma à la compréhension d'un dispositif mécanique particulier: les systèmes de cordes et de poulies. Ils présentent à des étudiants de premier cycle trois problèmes successifs comportant chacun un paragraphe descriptif et un schéma. Les participants doivent étudier chaque problème en temps libre, puis répondre à 8 QCM de compréhension inférentielle sur les configurations. Certaines questions concernent des relations spatiales non explicites dans le texte (par exemple, savoir où est accrochée telle poulie) alors que d'autres concernent les relations cinématiques (savoir dans quel sens tourne telle poulie quand on tire sur telle corde). Les participants reçoivent soit un texte (condition T), soit un schéma (condition S), soit un texte accompagné d'un schéma (condition T+S). Par ailleurs les sujets sont répartis en fonction de leur aptitude dans le domaine mécanique (une dimension qui corrèle fortement avec les épreuves de rotation mentale).

Les questions sur les configurations sont légèrement mieux réussies dans la condition T+S que dans la condition S, la condition T donnant des résultats intermédiaires. On observe un effet d'aptitude seulement dans la condition T+S. Les questions cinématiques sont mieux réussies dans la condition T+S que dans les deux autres conditions. Enfin le temps d'étude est environ 4 fois plus court dans la condition S que dans la condition T, alors que la condition T+S donne des temps intermédiaires. Dans les trois conditions les sujets à forte aptitude mécanique obtiennent de meilleurs scores que les autres.

Hegarty et Just soulignent le "rendement" élevé de la condition S, qui donne un niveau de compréhension équivalent au texte seul pour un temps de traitement quatre fois moindre. Ceci pourrait refléter le caractère intrinsèque des dépicions qui permet de construire directement des modèles mentaux avec un certain degré de précision.

Dans la deuxième expérience, Hegarty et Just se penchent sur les étapes de la construction du modèle mental dans la condition T+S. Ils utilisent pour cela un dispositif d'enregistrement des mouvements oculaires. 10 sujets de niveau universitaire étudient les trois configurations T+S de l'expérience 1, présentées cette fois sur l'écran d'un ordinateur. L'analyse des mouvements oculaires montre plusieurs effets intéressants: d'abord, les sujets interrompent leur lecture en moyenne 6 fois par problème pour examiner le schéma.

Les interruptions se font le plus souvent aux frontières syntaxiques. Les consultations du schéma se répartissent en deux catégories: des consultations locales, d'une durée moyenne inférieure à la seconde, et des consultations globales, d'une durée de 4 secondes en moyenne. Au cours de ces dernières, les sujets font fréquemment des va-et-vient entre constituants du schéma, ce qui n'est pas le cas dans les consultations locales. De plus, le temps de consultation par constituant est deux fois plus élevé en cas de consultation globale.

Hegarty et Just en déduisent que la construction du modèle mental utilise en parallèle l'information textuelle et graphique, et se fait en deux phases: une phase locale au cours de laquelle les sujets construisent une représentation locale correspondant à une phrase et un ou deux constituants du schéma; une phase globale au cours de laquelle les sujets intègrent les parties du modèle mental et, probablement, en construisent les propriétés cinématiques. Les sujets à faible aptitude mécanique effectuent davantage de consultations locales, mais autant de consultations globales que les sujets à forte aptitude mécanique, ce qui suggère qu'ils sont moins à même de visualiser mentalement les constituants du système à partir du texte seul.

Plus généralement cette étude montre que l'illustration apporte au lecteur une sorte de mémoire externe, qui permet de visualiser les relations entre constituants, libérant ainsi des ressources en mémoire de travail. Dans ces études, la consultation du schéma apparaît fortement contrainte par la lecture du texte (consultation aux frontières syntaxiques). Par rapport au diagramme seul, le texte rend explicite certaines relations implicites voire invisibles dans le schéma, notamment ses propriétés cinématiques. Les sujets n'extraient donc pas spontanément toutes les propriétés d'une dépicition, ce que confirment d'ailleurs les temps d'étude plus faibles dans la condition S. Enfin, considérant la fréquence élevée des consultations en cours de lecture, Hegarty et Just suggèrent qu'une disposition rapprochée des textes et des graphiques est sans doute plus favorable à leur compréhension.

### 6.2.2. Différentes illustrations favorisent différents aspects du modèle mental

Les illustrations semblent surtout utiles pour comprendre les relations spatiales et dynamiques entre constituants d'un objet ou d'une situation complexe. L'objectif de Waddill et McDaniel (1992) est de montrer que les illustrations peuvent également renforcer la compréhension des constituants eux-mêmes. Dans une première expérience, ils proposent à 48 étudiants un texte descriptif sur les glaciers accompagné, selon les cas, de 7 illustrations de détail (par ex., dimensions d'un glacier), de quatre illustrations relationnelles (par ex., comparaison de deux types de glaciers), ou d'aucune illustration. Après l'étude du texte, les sujets en effectuent un rappel libre incident. Les résultats montrent que chaque type d'illustration augmente le rappel du type d'information qu'elle concerne (les temps d'étude des deux conditions avec illustration sont équivalents). Dans la seconde expérience Waddill et McDaniel répliquent ce résultat et montrent qu'il est indépendant de la présence d'expressions relationnelles dans le texte (ainsi, en effet, parce que...). Il observent en revanche une forte relation entre l'effet des illustrations et la capacité de compréhension en lecture des sujets (Cf. 6.4 plus bas).

L'effet des illustrations n'est donc pas cantonné à la seule représentation des relations. Selon leur nature, les illustrations peuvent favoriser différents aspects du modèle mental. Ceci correspond à différentes fonctions des illustrations dans la typologie de Levin (cité par Gyselinck, 1996): fonction de représentation pour certains aspects d'un objet (par

exemple des dimensions), ou fonctions d'organisation et de transformation pour les relations entre parties du référent.

Dans d'autres cas les illustrations corrigent des problèmes d'organisation du modèle mental inhérents à la structure linéaire des textes. Ainsi Glenberg et Langston (1992) étudient le décalage qui existe parfois entre l'ordre d'énonciation des informations dans un texte et le déroulement temporel de la situation décrite par le texte. C'est le cas des descriptions de procédures: même les étapes simultanées des procédures sont nécessairement représentées l'une après l'autre dans un texte. Glenberg et Langston ont démontré qu'après avoir étudié un texte accompagné d'un schéma illustrant la simultanéité de deux étapes d'une procédure, les réponses des sujets à des questions de compréhension reflètent la plus grande "proximité" de ces étapes dans leur modèle mental. En revanche, après la lecture du texte seul, les réponses reflètent l'ordre d'énonciation des étapes de la procédure dans le texte.

### 6.2.3. La compréhension du texte illustré sollicite le calepin visuospatial

Quels sont, précisément, les processus cognitifs responsables de l'intégration des éléments textuels et graphiques au sein d'un modèle mental? Pour Glenberg et ses collègues (Glenberg & Langston, 1992; Kruley, Sciana & Glenberg, 1994), la réponse est assez simple: notre système cognitif est doté d'une mémoire de travail à plusieurs constituants, dont l'un est spécialisé dans le maintien et le traitement des informations visuospatiales: c'est le calepin visuospatial (CVS; Baddeley, 1986). Or, un modèle mental est par définition une structure analogique qui représente directement les relations spatiales entre éléments d'un objet ou d'une situation (Glenberg, Meyer & Lindem, 1987). Le "lieu" naturel d'instanciation du modèle mental durant la compréhension d'un texte illustré serait donc par hypothèse le CVS.

Kruley, Sciana et Glenberg (1994) entreprennent d'apporter des éléments empiriques décisifs à l'appui de cette hypothèse. Ils utilisent pour cela un paradigme de double tâche. Le sujet a pour tâche principale de regarder une figure pendant l'écoute d'un texte explicatif (par exemple, le fonctionnement d'un volcan). Avant la présentation de ces informations, on lui demande de mémoriser une configuration de points sur une matrice à 16 cases. Après la présentation du texte illustré, il doit dire si une nouvelle matrice est identique ou non à celle initialement présentée. Ce type de tâche secondaire fortement visuospatiale sollicite le CVS, et devrait donc interférer particulièrement avec le traitement d'un texte illustré. C'est ce que montre une première expérience: les performances des sujets à la tâche secondaire sont plus faibles en présence du texte illustré. En revanche, par rapport à une condition contrôle, la compréhension du texte illustré ne semble pas plus affectée par la tâche visuospatiale concurrente que ne l'est la compréhension du texte sans illustration. Par ailleurs, dans cette expérience comme dans les trois qui suivent, Kruley et al. démontrent une augmentation de la compréhension en présence de l'illustration, ce qui montre qu'un traitement supplémentaire a bien eu lieu dans cette condition.

Dans l'expérience 2, Kruley et al. essaient de provoquer une perturbation de la compréhension en demandant aux sujets de donner la priorité à la tâche secondaire (qui devient donc "tâche principale"). Ils observent un léger effet dans ce sens, mais les sujets ne semblent pas capables d'ignorer l'information contenue dans l'image. Kruley et al. pensent que lorsque des éléments de l'image sont dénommés dans le texte entendu, l'attention du sujet se porte sur ces constituants de manière quasi-irrépressible. De plus ils suggèrent un phénomène d'attitude: les questions de compréhension posées après le texte ressemblent aux tests que les sujets ont l'habitude de passer; ils tendent donc à y prêter

plus d'attention qu'à l'épreuve des matrices de points qui n'est associée à aucune attente scolaire.

Lorsque la présentation du texte ne s'accompagne pas d'une contrainte de compréhension, on n'observe plus de relation entre la présence d'une image et la tâche secondaire (expérience 3). Les sujets réussissent aussi bien à garder en mémoire une configuration de points lorsque la présentation et le test sont séparés par un texte illustré que lorsqu'elles sont séparées par un texte seul. Enfin, une quatrième expérience montre que la présence d'une illustration n'a pas d'incidence sur la mémorisation concurrente d'une séquence de chiffres, censée mobiliser la boucle phonologique et non pas le CVS.

Ces résultats suggèrent fortement que le CVS est le "lieu" des traitements intégratifs supplémentaires qui se déroulent lors de la compréhension d'un texte illustré. Toutefois, ce type de phénomène est très difficile à établir expérimentalement. Dans l'étude de Kruley et al., par exemple, la "démonstration" du rôle du CVS passe à plusieurs reprises par l'acceptation de l'hypothèse nulle, ce qui oblige les auteurs à utiliser un raisonnement indirect faisant appel à la puissance statistique des données. Par ailleurs, une étude de Gyselinck, Ehrlich, Cornoldi, de Beni et Dubois (2000) réalisée dans des conditions voisines montre elle aussi une nette amélioration de la compréhension en présence d'illustrations, mais aucune interaction entre la présence d'une illustration et la performance à une tâche secondaire visuospatiale. Selon les éléments fournis par ces auteurs dans leur discussion, une autre étude en cours montrerait, elle, de tels effets. On voit donc que les hypothèses sur la "mécanique cognitive" à l'œuvre dans l'intégration texte-image n'ont pas encore trouvé une assise empirique très large.

En résumé, les illustrations favorisent objectivement la compréhension, particulièrement celle des relations spatiales ou dynamiques au sein d'un référent complexe. Toutefois, en fonction des types d'illustrations on peut observer une facilitation sélective de la mémorisation des détails ou des relations. La lecture du texte et de l'illustration semblent s'effectuer en parallèle, du moins chez des sujets adultes (modèle de compréhension pas à pas de Hegarty et ses collègues). Les lecteurs identifient d'abord les constituants locaux du modèle mental, puis ils en construisent les propriétés spatiales et dynamiques globales. Enfin certaines données empiriques semblent confirmer l'idée que l'intégration des textes et des graphiques sollicite préférentiellement le calepin visuospatial de la mémoire de travail qui serait, selon Glenberg et ses collègues, le "lieu" cognitif de la construction des modèles mentaux. Mais ces résultats demandent encore confirmation.

### **6.3. Les facteurs matériels d'intégration multimédia**

#### **6.3.1. Introduction**

Les images peuvent donc faciliter la compréhension d'un texte, c'est un fait avéré et compatible avec les théories actuelles de la mémoire de travail. Que ce soit de façon explicite ou - le plus souvent - inconsciente, les concepteurs d'ouvrages techniques ou pédagogiques font un usage parfois immodéré de ce principe, en proposant des documents dans lesquels textes et images fusent littéralement de tous les côtés, au mépris d'un autre principe bien connu des psychologues: la capacité de traitement limitée du système cognitif. Les choix de présentation des textes et des images, tant au plan de la disposition spatiale, de la contiguïté temporelle, ou bien, dans les systèmes audiovisuels, de la complémentarité modale, sont déterminants. Or la présentation simultanée de textes et d'images peut entraîner une augmentation de la charge cognitive nécessaire pour extraire

l'information de ces deux sources au point de saturer la capacité de traitement du sujet (Sweller, Chandler, Tierney et Cooper, 1990). Il est donc important de s'intéresser aux facteurs matériels qui influencent la compréhension de documents illustrés et notamment la contiguïté spatiale, la complémentarité modale (en jouant sur la présentation orale ou écrite des textes), et la cohérence référentielle.

### 6.3.2. Effets de la disposition spatiale et temporelle

L'efficacité de sources d'informations hétérogènes (par ex: texte + image) dépend des conditions de présentation de ces sources. On doit à Richard Mayer et ses collègues une série d'articles extrêmement informatifs sur ce sujet (Cf. Mayer, 1997, pour une revue). Mayer et al. ont notamment formulé deux grands principes d'intégration multimédia: le principe de contiguïté et le principe de modalité (Mayer & Anderson, 1992; Mayer & Gallini, 1990; Moreno & Mayer, 1999; voir aussi Sweller et al., 1990).

Le principe de contiguïté (Mayer & Anderson, 1992), aussi appelé "effet de partage de l'attention" (Sweller et al., 1990), prévoit qu'une source hétérogène d'informations (texte+images) est d'autant plus efficace au plan de la compréhension que ses éléments sont présentés de façon contiguë. Cette contiguïté peut-être définie selon des critères spatiaux et temporels. La contiguïté spatiale est caractérisée par la proximité physique des éléments de la source qui désignent les mêmes référents, proximité qui peut encore être renforcée au moyens de signes explicites (étiquettes, flèches, etc.). La contiguïté temporelle concerne l'agencement temporel des différentes sources d'informations. La compréhension est généralement améliorée si les différents stimuli sont présentés simultanément ou dans une fenêtre temporelle restreinte. Ce principe concerne surtout les présentations multimodales, c'est-à-dire à la fois visuelles et auditives.

Le principe de complémentarité modale concerne précisément l'utilisation simultanée de différents canaux perceptifs. Il prévoit que des sources d'informations hétérogènes seront mieux mémorisées si elles sollicitent les deux canaux (visuel + auditif) plutôt qu'un seul. Par exemple, dans l'étude de Mousavi, Low et Sweller (1995, cités par Moreno & Mayer, 1999), les sujets étudient des fonctions géométriques représentées graphiquement. Les graphiques sont accompagnés de commentaires soit écrits, soit oraux, soit les deux. La présentation orale favorise la compréhension, que les différentes parties du message soient présentées simultanément (expérience 1 et 2) ou successivement (expériences 3 et 4). L'effet de complémentarité modale se manifeste également lorsque les deux parties du message sont verbales (texte écrit + texte oral).

L'étude de Moreno et Mayer (1999) a pour but de dissocier les effets de la contiguïté spatiale et de la complémentarité. Ils proposent à 132 étudiants de premier cycle une animation graphique représentant le processus de formation des éclairs<sup>15</sup>. Cette animation, présentée sur un écran d'ordinateur est accompagnée soit d'un commentaire oral (O), soit d'un commentaire écrit présenté au dessous de la zone graphique (ES), soit d'un commentaire écrit présenté à l'intérieur de la zone graphique, près du référent (EI, Figure 6.2).

---

<sup>15</sup> Bien que cette étude utilise une animation informatique, je la considère plus à sa place dans une section sur l'effet des illustrations en raison de la simplicité du matériel utilisé. Une justification complémentaire de ce choix - moins anodin qu'il n'y paraît - est présentée dans la section 6.5 consacrée aux animations "complexes".

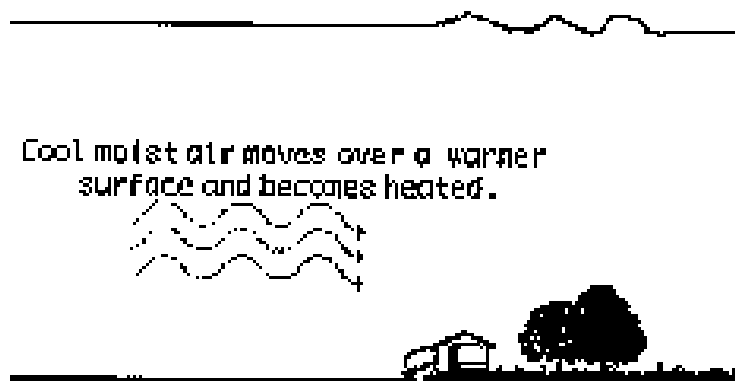


Figure 6.2. Matériel utilisé par Moreno et Mayer (1999, condition EI)

Dans les trois conditions, l'animation dure trois minutes, et la durée de présentation du commentaire écrit est équivalente à celle du commentaire oral. L'hypothèse est que la performance pourrait s'améliorer si l'on réduit l'effort nécessaire pour intégrer les éléments visuels et graphiques (Cf. l'étude de Hegarty & Just, 1993, ci-dessus). On s'assure au moyen d'un pré-test que les participants ont très peu de connaissances du domaine. Les sujets doivent étudier l'une des versions de l'animation, puis effectuer trois post-tests: production d'une explication de la formation des éclairs (rappel), résolution de quatre questions inférentielles (transfert), et appariement de quatre expressions verbales avec les éléments graphiques correspondants (ex: "*entourez le front froid*").

La version O entraîne de meilleures performances que les versions ES et EI aux trois post-tests; la version EI est supérieure à la version ES pour les tests de rappel et de transfert. Ceci est compatible à la fois avec le principe de complémentarité modale et avec le principe de contiguïté spatiale, du moins pour le rappel et la production d'inférences.

L'expérience 2 a pour but de séparer les effets de contiguïté spatiale et temporelle. En effet dans l'expérience 1, si l'image et le commentaire oral peuvent effectivement être traités simultanément, ce n'est pas le cas du commentaire écrit qui demande un va-et-vient entre texte et graphique, même dans la condition intégrée (Hegarty & Just, 1992). Six conditions sont étudiées: les conditions O et ES de l'expérience 1 plus quatre conditions avec présentation successive de chacune des sources: commentaire oral suivi de l'animation (OA), l'inverse (AO), commentaire écrit suivi de l'animation (EA) ou l'inverse (AE).

De nouveau, on observe un effet de complémentarité modale: les groupes avec commentaire oral (O, OA, AO) obtiennent de meilleurs scores que les groupes avec commentaire écrit sur les trois post-tests. Il n'y a pas d'effet d'ordre pour les versions successives. Il n'y a pas non plus d'effet de contiguïté temporelle pour les versions orales: la présentation simultanée O donne les mêmes résultats que les présentations successives OA et AO. En revanche, cet effet est observé pour les versions écrites, lors du test d'appariement: les sujets qui ont reçu la version ES obtiennent de moins bons scores que ceux ayant vu les versions EA ou AE. Ceci suggère un effet de partage de l'attention lorsque la présentation est simultanée.

L'absence d'effet de contiguïté est contradictoire avec les résultats précédents de Mayer et Anderson (1992), qui indiquaient une facilitation en présentation intégrée. Cette

différence peut s'expliquer par une organisation différente du matériel dans les deux études: dans l'étude de Moreno et Mayer, animations et commentaires sont présentés en alternance, ce qui permet aux sujets de garder en mémoire de travail la première partie de l'information jusqu'au moment de traiter l'autre. Dans l'expérience de Mayer et Anderson, le sujet voit d'abord toutes les informations d'une catégorie, puis toutes celles de l'autre.

L'intégration spatiale est également importante dans le cas de systèmes donnant accès à des textes secondaires à partir d'un texte principal (Black, Wright, Black & Norman, 1992). Dans le domaine de l'infographie, Bétrancourt et Caro (1998) ont mis en évidence une économie de traitement lorsque des commentaires verbaux sont intégrés à une figure complexe sous la forme "d'escamots", des étiquettes qui une fois sélectionnées entraînent l'apparition d'une fenêtre de texte superposée à la figure. Cette technique semble réduire la complexité perceptive de la figure et permettre de meilleures stratégies de prise d'information.

### 6.3.3. Effets de la cohérence référentielle

Pour que l'image joue son rôle facilitateur de la compréhension du texte, il faut que les deux sources soient cohérentes l'une par rapport à l'autre, c'est-à-dire qu'elles fassent référence aux mêmes objets. En termes de modèle mental, il faut que les représentations descriptives et décriptives contribuent au même modèle, qui sera alors enrichi en vertu de la spécificité de chaque source (Hegarty & Just, 1993; Waddill & McDaniel, 1992). Dans la plupart de ces études, les textes et les diagrammes sont soigneusement sélectionnés ou construits de façon à assurer la meilleure cohérence référentielle. Dans le cas contraire, on peut s'attendre à ce que l'effort réalisé pour établir les points communs entre image et texte interfère avec la construction d'un modèle mental.

Les travaux de Merlet (Merlet & Gaonac'h, 1995; Merlet, 1998) ont apporté une démonstration directe du phénomène de surcharge cognitive inhérent à une mauvaise intégration texte-image. Merlet a travaillé à partir d'un produit multimédia du commerce destiné à l'enseignement de l'anglais langue étrangère. Ce produit propose, entre autres exercices, l'écoute de dialogues illustrés par des images simples représentant les acteurs du dialogue et les situations évoquées dans ce dialogue. Merlet a proposé à des étudiants de niveau intermédiaire (2e année universitaire) l'un de ces dialogues avec ou sans les images d'accompagnement. Les résultats ont montré que la présence d'images, loin de faciliter la compréhension, entraînait au contraire une augmentation des difficultés par rapport au groupe contrôle. L'auteur a proposé une interprétation de ce phénomène en termes de charge cognitive: les étudiants de niveau intermédiaire sont contraints de consacrer une attention importante au traitement des unités linguistiques élémentaires (les mots du texte entendu). Dans ces conditions, l'ajout d'informations imagées, qui peuvent difficilement être ignorées, entraîne une surcharge cognitive attestée par des pauses plus nombreuses lors de l'écoute, des retours en arrière, et de moins bonnes performances en compréhension.

Un exemple concret suffit à illustrer ce phénomène (Figure 6.3).



Figure 6.3. Un exemple de texte illustré dans l'étude de Merlet (1998).

NB. Le dialogue qui accompagne l'image est le suivant: **1. Jane** : By the way, you didn't watch that programme on BBC2 last night, did you ? On Australia ! **2. Sharon** : No, I've far too much homework mid-week to watch anything. What a shame, I'd have loved that !

Dans cet exemple, l'image représente deux jeunes filles en train de discuter. Les commentaires spontanés des personnes à qui l'on demande de quoi elles parlent incluent: "elle s'est achetée une télé", "elle aime lire", "je suis fatiguée, j'ai besoin d'un café". Le sens littéral du message ("au fait, tu n'as pas vu ce reportage sur BBC2 hier soir, sur l'Australie?", "non, j'ai beaucoup trop de devoirs en milieu de semaine, quel dommage, ça m'aurait plu.") est rarement deviné. Les mots "non pertinents" activés ("café", "acheter"...) peuvent potentiellement interférer avec la compréhension.

En conclusion, la simple présence d'informations multiples ne garantit donc pas la compatibilité cognitive entre l'hypermédia et l'apprenant. Pour assurer cette compatibilité, les informations doivent être présentées de façon intégrée, et ce d'un point de vue spatial, temporel et référentiel.

#### 6.4. Différences individuelles et développement des capacités d'intégration multimédia

##### 6.4.1. Le rôle des capacités de compréhension

Waddill et McDaniel (1992, expérience 2) examinent l'effet différentiel des illustrations en fonction des capacités de compréhension des sujets. Selon ces auteurs les bons comprennent se caractérisent par leur capacité à mettre en relation les éléments du texte (via notamment la production d'inférences). Ils proposent quatre hypothèses quant à l'interaction entre capacité de compréhension et présence des illustrations:

- Selon l'hypothèse de compensation générale, les bons comprennent effectuent spontanément la mise en relation des éléments du texte. Ils n'ont donc pas besoin de l'information complémentaire apportée par les illustrations. Ces dernières seraient en revanche particulièrement utiles aux mauvais comprennent, car elles les dispensent de construire les relations à partir du texte.

- Selon l'hypothèse de compensation spécifique, les bons compreneurs se focalisent sur les relations, mais pourraient négliger les éléments de détail du texte. Les illustrations relationnelles leur seraient inutiles mais pas les illustrations de détail. L'inverse serait observé pour les mauvais compreneurs, centrés sur les détails mais incapables d'extraire les informations relationnelles.

- L'hypothèse d'enrichissement général postule que les illustrations ne sont utiles qu'aux individus qui maîtrisent les processus nécessaires à leur traitement. Les bons compreneurs profiteraient de l'information ajoutée par les illustrations alors que les mauvais compreneurs peineraient à élaborer une représentation mentale que l'information vienne d'un texte ou d'une image.

- L'hypothèse de l'enrichissement spécifique prévoit que les bons compreneurs bénéficieraient des illustrations compatibles avec leur niveau de traitement privilégié, à savoir les illustrations relationnelles, alors que les mauvais compreneurs tireraient avantage des seules illustrations de détail.

Waddill et McDaniel testent ces hypothèses au moyen d'une expérience dans laquelle ils demandent à des étudiants bons, moyens ou mauvais compreneurs de lire un texte sur les glaciers présenté soit seul, soit accompagné d'illustrations de détail, soit accompagné d'illustrations relationnelles. La compréhension est mesurée au moyen d'une évaluation subjective et d'un rappel libre. Les bons et moyens compreneurs qui ont reçu des illustrations relationnelles estiment avoir mieux compris que ceux qui ont reçu des illustrations de détail ou pas d'illustration du tout. En revanche, la condition de présentation n'influence pas les évaluations des mauvais compreneurs. Les rappels des bons compreneurs contiennent davantage d'informations relationnelles et de détails lorsqu'ils ont vu un texte accompagné d'illustrations relationnelles. Seul le taux d'informations relationnelles augmente dans cette condition chez les sujets de niveau intermédiaire. Enfin chez les mauvais compreneurs, la présentation d'illustrations de détail augmente le rappel des détails. Ces résultats sont compatibles avec l'hypothèse d'enrichissement sélectif. Les illustrations favoriseraient la mise en œuvre des processus "disponibles" chez les sujets. Elles permettraient un codage profond d'informations que les sujets ne codent pas normalement assez profondément pour les rappeler, bien qu'ils en aient la capacité. En d'autres termes, les illustrations serviraient de signal pour la mise en œuvre de traitements profonds.

Il est intéressant de noter que cette vision est complémentaire à celle proposée par Hegarty et Just (1993) dans leurs expériences sur les poulies: ils faisaient du texte un système de signaux déclenchant la consultation du schéma, alors que Waddill et McDaniel défendent l'hypothèse inverse. Ces deux interprétations pourraient être valides chacune dans des domaines de connaissances différents, à moins que les deux types de signaux puissent fonctionner en parallèle, le sujet se trouvant ainsi guidé à la fois dans son traitement du texte et de l'illustration mais toujours à condition qu'il maîtrise les processus correspondants.

#### 6.4.2. Le rôle des capacités visuospatiales

Les individus ne sont en effet pas tous égaux devant les combinaisons de textes et de graphiques. Hegarty et Sims (1994) ont réalisé une série d'expériences mettant en évidence le rôle de l'aptitude visuospatiale dans la résolution de problèmes mécaniques simples à partir de phrases et de schémas. Les tâches étaient similaires aux problèmes de cordes et de poulies utilisés dans l'étude de Hegarty et Just (1993) rapportée plus haut.

Les sujets (étudiants de 1er cycle universitaire) recevaient des séries d'items comportant chacun un dessin du système de poulies et de cordes, et une phrase énonçant le mouvement d'un des éléments (par ex: "*la poulie du bas tourne de gauche à droite*"). Les sujets devaient vérifier la justesse de l'assertion par rapport au dessin (étant donné une flèche indiquant le mouvement d'un autre élément point de départ du système). Hegarty et Sims proposent un modèle de tâche en 6 étapes:

1. Lire et comprendre le texte.
2. Inspecter et comprendre le diagramme.
3. Construire des connexions référentielles entre texte et diagramme.
4. Inférer le mouvement de l'élément critique (par exemple, la poulie du bas).
5. Comparer le mouvement inféré avec celui décrit par l'assertion.
6. Répondre.

L'hypothèse est que si les capacités visuospatiales jouent un rôle dans ce type de problème, alors on devrait observer une différence en termes de vitesse et de précision des réponses entre sujets à forte ou faible capacité. De plus, cette différence devrait s'accroître à mesure que le problème posé demande davantage d'inférences visuospatiales (par exemple, juger une assertion sur un élément éloigné du point de départ). Enfin, la différence ne devrait concerner que des inférences portant sur les aspects cinématiques du système, par opposition aux aspects statiques ("*la poulie de gauche est accrochée au plafond*"), car ces derniers ne demandent pas d'inférence visuospatiale.

Ces trois hypothèses sont vérifiées dans l'expérience 1. Les sujets à faibles capacités visuospatiales (test du "papier plié" ou Paper Folding Test) font plus d'erreurs surtout pour les vérifications portant sur des éléments éloignés du point de départ et pour les questions portant sur les aspects cinématiques. Il n'y a pas de différence de temps de réponse, ce qui suggère que les sujets utilisent des stratégies similaires, mais de façon plus ou moins précise. La seconde expérience réplique ces résultats, et montre que les sujets à forte ou faible capacité visuospatiale ont les mêmes patterns de mouvements oculaires: ils décomposent le problème en autant d'éléments qu'il y a de constituants du système à animer mentalement pour juger l'assertion (Hegarty & Just, 1993). La troisième expérience vérifie que ce sont bien les étapes strictement visuospatiales de la tâche (étapes 2 et 4 du modèle) qui sont affectées par les différences individuelles. En effet, en utilisant une variante du problème dans laquelle l'énoncé verbal est remplacé par une flèche montrant le mouvement supposé de l'élément "critique", les auteurs observent une diminution du temps de réponse et une augmentation des réponses correctes, mais pas d'interaction avec l'aptitude visuospatiale.

Quel est le mécanisme sous-jacent à ces résultats? Il se situe selon Hegarty et Sims au niveau du fonctionnement d'un des constituants de la mémoire de travail, le calepin visuospatial (Baddeley, 1986). Soit les sujets à faibles capacités visuospatiales ont un empan visuospatial plus faible (calepin à capacité plus réduite), soit ils sont moins efficaces dans les opérations de transformation mentale de l'image (visualisation du sens de rotation). Comme le soulignent les auteurs, ces deux hypothèses sont sans doute indissociables car il pourrait y avoir un compromis capacité-complexité des traitements dans le fonctionnement de la mémoire de travail.

### 6.4.3. Aptitude visuospatiale et gestion cognitive de l'activité

Hegarty et Steinhoff (1997) étudient l'effet d'une prise de notes sur la résolution de problèmes basés sur des représentations graphiques (engrenages et poulies dans l'expérience 1, leviers dans l'expérience 2)<sup>16</sup>. Un exemple typique de ce type de problème consiste à déterminer le sens de rotation d'une des roues dentées d'un engrenage quand le sens de rotation d'une autre roue est indiqué (Figure 6.4.).

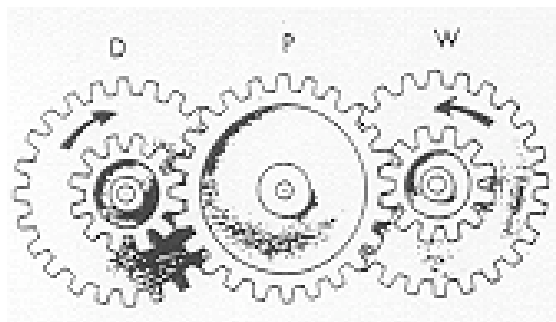


Figure 6.4. Exemple de matériel utilisé par Hegarty et Steinhoff (1997)

Ce type de tâche demande que le sujet comprenne le diagramme, qu'il en effectue une animation mentale, et qu'il en déduise le sens de rotation de la roue test. En ce qui concerne la phase d'animation mentale, les travaux antérieurs d'Hegarty ont montré que les sujets tendent à décomposer un graphique complexe en unités plus simples, puis à construire mentalement le mouvement de chacune des unités, à partir de la roue de départ (Hegarty & Just, 1993). Le temps de résolution du problème est donc proportionnel au nombre d'unités intermédiaires à animer avant d'arriver à l'unité critique. En soulageant la mémoire de travail, la prise de notes pourrait faciliter la résolution du problème, surtout chez les sujets ayant de faibles capacités de visualisation, dont on sait qu'ils sont moins performants dans ce type de tâche (Hegarty & Sims, 1994, voir ci-dessus). Toutefois, prendre des notes suppose que le sujet ait conscience (a) du fait que la charge en mémoire de travail risque de lui poser problème, (b) du fait que les notes peuvent aider à alléger cette charge et (c) du moment et des modalités de prise de notes les plus favorables.

Dans l'expérience 1, 48 étudiants doivent résoudre 40 problèmes de vérification de mouvement dont 20 portent sur des engrenages et 20 sur des poulies. Dans chaque catégorie, la moitié des problèmes concerne deux éléments proches dans le dispositif alors que l'autre moitié porte sur des éléments distants. Les sujets sont répartis en deux groupes: avec ou sans prise de notes. Par ailleurs dans chaque groupe on distingue les sujets selon qu'ils ont de fortes ou faibles capacités visuospatiales (VS+ et VS-, d'après la médiane des scores au Paper Folding Test).

Les sujets prennent des notes sur environ un tiers des problèmes en moyenne. 80% des notes prises indiquent la direction du mouvement de l'un des éléments du système. Le taux de prise de notes varie en fonction du type et de la complexité du problème: de 6% pour des problèmes de poulies à éléments proches à 51% pour des problèmes d'engrenages à éléments distants. Le taux d'erreurs est plus élevé pour les problèmes à éléments distants, surtout chez les sujets VS-. Contrairement aux hypothèses, la prise de notes n'a globalement aucun effet sur la performance des sujets. Toutefois on

<sup>16</sup> Je remercie Elsa Eme pour m'avoir signalé l'existence de cet article.

constate d'importantes variations dans le taux de prise de notes (sans qu'il soit lié à la capacité visuospatiale). Les sujets VS- qui prennent le plus de notes ont un taux d'erreurs plus faible. La quantité de notes n'a pas d'incidence sur la performance des sujets VS+.

La deuxième expérience a pour but de vérifier que l'absence d'effet de la prise de notes chez les VS+ n'est pas liée à un effet plancher. A partir d'un ensemble d'items plus difficiles, les auteurs répliquent les résultats de la première expérience, ce qui suggère que chez les sujets à fortes capacités spatiales, le maintien en mémoire de travail des informations intermédiaires n'est pas un problème.

La capacité visuospatiale et les compétences métacognitives semblent donc avoir une influence distincte sur la performance des sujets dans ces tâches de vérification de mouvement. La prise de notes semble utile aux sujets à faibles capacités visuospatiales, mais cette prise de notes varie d'un sujet à l'autre. On ne peut donc conclure avec certitude que la prise de notes est la cause de l'amélioration des performances. En effet, comme le soulignent Hegarty et Steinhoff, il se peut aussi qu'une dimension sous-jacente (la métacognition) soit responsable à la fois de l'augmentation des notes et d'une meilleure performance (par exemple via de meilleures stratégies d'exploration visuelle du schéma).

#### 6.4.4. Le développement cognitif des capacités d'intégration multimédia

On sait peu de choses sur le développement des capacités d'intégration des textes et des graphiques chez l'enfant. La sagesse populaire veut que, comme chez l'adulte, l'image joue un rôle facilitateur dans la compréhension de textes par les enfants. Il est à peu près certain que l'image présente un caractère plus attractif que le texte et qu'elle favorise ainsi l'implication de l'enfant dans la tâche de compréhension. Les données sur les effets de l'image sur les processus de construction des modèles mentaux sont plus rares mais quelques études réalisées sur ce sujet montrent de nettes tendances développementales et une corrélation - déjà! - entre certaines dimensions cognitives et la capacité à mettre en relation les informations des deux sources. Miller et Pressley (1987) ont par exemple montré que la présentation d'images facilite la production d'inférences concernant les informations implicites, par ex. "le travailleur creuse un trou" (inférence: avec une pelle) chez des enfants de 6 à 7 ans, mais pas chez ceux de 4 à 5 ans (preschoolers). Il faut donc attendre (coïncidence?) l'acquisition du code écrit pour que les enfants apprennent à rattacher les traits figuratifs (et/ou les idées qui leur viennent à l'esprit) avec les formes linguistiques cooccurrentes. Pressley, Cariglia-Bull, Deane et Scheinder (1987) montrent qu'une consigne d'imagerie favorise l'apprentissage de phrases chez des enfants de 6 à 11 ans, mais particulièrement chez ceux qui présentent les compétences cognitives les plus développées (par exemple l'empan de mémoire de travail).

Guercin (in press) a montré que dans une tâche de description d'itinéraire, des enfants de 8 ans ne bénéficiaient pas d'une présentation de l'itinéraire sous forme d'une description verbale et d'un plan, par rapport à la présentation séparée de l'une de ces deux sources. Les enfants plus âgés (10 et 12 ans) semblent mieux à même d'intégrer les deux sources pour élaborer un modèle mental de l'itinéraire: leur temps d'étude et de dessin de l'itinéraire sont inférieurs à ceux de la condition "texte seul" (mais toujours supérieurs à ceux de la condition "image seule").

Enfin, Jean-Michel Boucheix et ses collaborateurs ont comparé l'effet de représentations statiques ou animées pour la compréhension de phénomènes mécaniques ou physiques chez des élèves d'école primaire. Ils ont constaté des interactions entre le type de support (papier, ordinateur) et plusieurs variables individuelles: expérience

préalable des enfants dans l'usage du support, niveau scolaire, et style d'apprentissage (Boucheix, communication personnelle, 6/10/2000; Noël, 1999). Les animations graphiques permettent d'améliorer la compréhension de systèmes d'engrenages, mais surtout dans la condition où l'enfant règle lui-même le rythme de défilement des images.

Comme la capacité de recherche d'informations (Cf. Chapitre 4), il est probable que la capacité à intégrer les textes et les images se construit progressivement sur la base d'apprentissages plus élémentaires. La capacité très limitée de l'enfant à produire des inférences lexicales pourrait être une source de contraintes redoutables. Ces inférences sont souvent requises pour identifier les relations entre les mots qui désignent des objets (l'enfant, la rue, la maison, la poste) et des actions mettant en scène ces objets (aller de la maison à l'école puis tourner à droite sur la rue de la poste). Les autres sources de difficultés recensées dans le domaine de la recherche d'informations (mémoire de travail et gestion des ressources cognitives) sont probablement tout aussi déterminantes.

## **6.5. Les animations graphiques complexes**

### **6.5.1. Quand la pratique résiste à la théorie**

Les progrès des systèmes informatiques permettent maintenant de réaliser à des coûts raisonnables des produits d'accompagnement scolaire à base de représentations graphiques animées. Comme les rédacteurs de manuels scolaires, les concepteurs multimédia font un usage immodéré de ces graphismes animés, sur la base d'une croyance générale en l'efficacité des animations pour la compréhension des phénomènes représentés.

Il est vrai que les travaux sur l'intégration texte-image fournissent plusieurs arguments théoriques en faveur de l'emploi des animations graphiques. On sait par exemple que la compréhension d'un phénomène dynamique à partir d'une image statique est liée à la capacité visuospatiale de l'élève. Certains individus sont plus lents et moins précis lorsqu'ils doivent imaginer mentalement le mouvement d'un objet (Cf. les exemples donnés dans les sections 6.4.2 et 6.4.3.). La visualisation des mouvements ou des transformations pertinents au moyen d'animations graphiques pourrait permettre de palier ces déficits. De plus les animations présentent un caractère attrayant qui peut être exploité pour attirer et maintenir l'attention de l'élève sur les informations importantes. D'un autre côté la présence ou l'excès d'animations pourrait aussi détourner l'attention de l'élève sur d'autres aspects de l'information, par exemple le texte.

Conformément à la situation générale dans les sciences de l'information, les publications qui affirment l'efficacité des animations (à partir d'arguments parfois douteux) sont infiniment plus nombreuses que les études scientifiques visant à en évaluer l'influence réelle sur les comportements des apprenants. Les quelques expériences publiées récemment ne peuvent toutefois qu'inciter à la prudence. Par exemple, Wright, Milroy et Lickorish (1999) examinent l'effet d'animations graphiques sur la mémorisation de courts textes d'intérêt général (Histoire et géographie de l'Angleterre). Au sein de chaque texte des illustrations de type "quoi?" (montrer un objet ou une notion), "quand?" (chronologie) et "Où?" (carte de géographie) sont proposées. Dans l'expérience 1, les illustrations sont accessibles à partir de boutons (liens hypertextes) placés à côté de certains mots-clés du texte. Les illustrations sont soit des images fixes, soit des animations dérivées d'images fixes, mais qui n'apportent pas d'information supplémentaire. Par exemple, dans la version

animée de la chronologie, les événements s'affichent progressivement sur l'axe chronologique, alors que dans la version fixe ils s'affichent d'un seul coup.

Les participants (60 adultes volontaires rémunérés) peuvent solliciter les illustrations en cliquant sur l'astérisque. Le texte est alors remplacé par l'illustration, au bas de laquelle se trouve un bouton "retour". Dans l'expérience 1, Wright et al. comparent trois conditions: une condition sans illustration, une condition avec illustrations fixes, et une condition avec illustrations animées. Seules 20% des participantes négligent de consulter au moins une illustration dans la condition animée, contre 60% dans la condition fixe. Il n'y a pas de différence quant à la compréhension du texte, estimée par une série de questions factuelles.

Dans l'expérience 2, les auteurs rajoutent une seconde variable de présentation: le moment où le graphique est proposé au lecteur. Dans la condition "avant" le graphique est proposé avant l'affichage du texte, alors que dans la condition "pendant" il est proposé en cours de lecture (comme dans l'expérience 1). De plus, les auteurs renforcent la consigne de consulter les graphiques. Le taux de non consultation est alors très faible, sauf dans la condition "statique-pendant". Cette fois la comparaison des conditions "statique-pendant" et "animé-pendant" montre une baisse significative de la compréhension en condition "animé-pendant" par rapport aux trois autres. Les auteurs en concluent que la présentation de graphiques animés pendant la lecture d'un texte attire effectivement l'attention des lecteurs (plus de consultations que pour des graphiques fixes), mais peut interférer avec la mémorisation des textes. Toutefois il est clair que dans cette étude le mode de présentation utilisé (pages de textes et d'illustrations qui se substituent les unes aux autres) est en contradiction avec le principe de contiguïté spatiale de Mayer et al. (Moreno & Mayer, 1999, Cf. section 6.3.2) ce qui pourrait suffire à expliquer l'interférence: les lecteurs ne peuvent mettre en place une stratégie de construction pas à pas du modèle mental (Hegarty & Just, 1993).

Mais d'autres résultats expérimentaux sont venus confirmer l'impression que les animations ne remplissent pas le rôle qu'on attend d'elles. Schnotz (in press) a utilisé une représentation graphique du globe terrestre pour expliquer le principe du changement de date. Les sujets peuvent effectuer différentes manipulations sur cette représentation, comme assigner une heure particulière à une ville particulière, ou effectuer des rotations contrôlées. Cette présentation augmente la mémorisation de détails (heure relative de différentes villes du globe), mais diminue la compréhension profonde (questions requérant une rotation mentale), par rapport à une condition où les sujets utilisent une image fixe.

Schnotz propose deux interprétations non exclusives du phénomène: d'une part, le passage d'une image fixe à une animation fait perdre le bénéfice des points de repères sur lesquels une correspondance texte-image peut être établie (NB. dans la perspective de Schnotz il y a mise en correspondance de texte et d'images même en l'absence de texte explicite, en raison des changements de formats qui caractérisent la construction des modèles mentaux. Cf. section 6.1). D'autre part, le fait de jouer une animation dispense l'élève de se représenter mentalement le mouvement correspondant. L'élève passerait donc implicitement d'un statut de compreneur-acteur à un statut de témoin-spectateur, ce qui est incompatible avec la construction de connaissances.

Lowe (in press) a lui aussi pour objectif d'améliorer la compréhension profonde à l'aide d'animations, cette fois dans le domaine des phénomènes météorologiques. Dans son expérience, des étudiants plus ou moins expérimentés dans le domaine utilisent une carte météorologique dynamique montrant la progression des fronts sur 24 heures, dans le but

de produire des prédictions exactes sur l'évolution des conditions dans les 24 heures suivantes. Contrairement aux attentes, l'utilisation de l'animation ne profite nullement aux sujets débutants. Les comptes rendus qu'ils écrivent présentent séparément les caractéristiques de la situation (listes de caractéristiques) sans parvenir à les intégrer. Lowe note que l'animation rend certains traits graphiques particulièrement saillants, ce qui a pour effet d'attirer l'attention des élèves, mais nuit à leur compréhension d'ensemble des phénomènes.

On le voit, les effets objectifs des animations tels qu'attestés dans des expériences scientifiques, ne sont jusqu'à présent pas aussi positifs qu'on le souhaiterait. Il semble qu'il faille choisir avec le plus grand soin les aspects des phénomènes qui doivent être animés, et avec plus grand soin encore la façon précise dont ces phénomènes sont animés. Peut-être la question des animations est elle également liée à la stratégie pédagogique d'ensemble des concepteurs. Par exemple, l'utilisation des animations paraîtrait peu utile dans un système d'enseignement basé sur l'apprentissage de faits: le texte appris par cœur est alors plus efficace. En revanche, dans une pédagogie basée sur le développement des capacités de raisonnement, l'utilisation d'environnements riches en animations, dans le cadre de situations-problèmes ouvertes, pourrait trouver une justification (à condition, encore une fois, que ces animations soient pertinentes et pas seulement décoratives; Choplin, Galisson & Lemarchand, 1999).

## 6.5.2. Animations et consignes de traitement: une étude exploratoire<sup>17</sup>

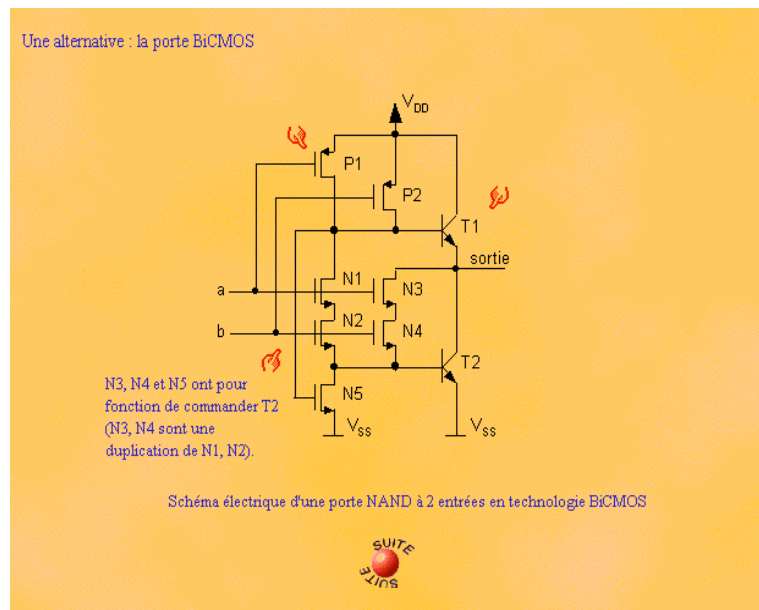
### 6.5.2.1. Introduction et méthode

Dans le cadre d'une coopération entre le Laboratoire Langage et Cognition et le Département d'Innovation Pédagogique de l'Ecole Nationale Supérieure des Télécommunications, nous avons entrepris une série d'études sur la conception de produits multimédia pour l'enseignement de l'électronique. Ces études consistent entre autres à tester expérimentalement l'impact de différents choix de conception (rassemblés au sein du cahier des charges de l'équipe de développement multimédia) et de mise en situation pédagogique du produit sur la compréhension des notions d'électronique par des étudiants de niveau post-bac.

Une première étude (Rouet, Choplin et Dubois, in prep.) avait pour but d'évaluer l'impact d'une stratégie de conception centrée sur les animations graphiques. Deux articles multimédias, portant respectivement sur le domaine des additionneurs et des transistors, ont été utilisés dans cette étude. Chaque article présente en 8 à 10 minutes des notions théoriques et pratiques, avec l'objectif d'aider l'élève à se représenter l'utilisation de ces notions en termes de compromis (par exemple, compromis surface-performance dans le cas des additionneurs). Deux versions de chaque article ont été développées. La première (V+) inclut des illustrations graphiques significatives, animées, et placées avant les explications verbales. La V+ privilégie aussi un contrôle par l'utilisateur du déroulement de la séquence multimédia. Dans la seconde version (V-) les images significatives sont supprimées ou remplacées par de simples illustrations, et l'interactivité est diminuée. Les Figures 6.5 a et b présentent un exemple d'écran multimédia en V+ et V-, respectivement.

---

<sup>17</sup> Je remercie Dominique Degrugillier, Marité Milon et les élèves volontaires de l'ENST, de l'INT et de l'ENSTB pour leur contribution à cette expérience.



Une alternative : la porte BiCMOS

Dans une porte BiCMOS, l'étage de sortie est constitué de deux transistors bipolaires montés en push-pull (quand l'un est passant, l'autre est bloqué). Le circuit logique de base CMOS est conservé et augmenté de quelques transistors afin de commander correctement ce push-pull de sortie.

Par rapport à une porte CMOS standard, la porte BiCMOS est capable de commander en sortie une capacité d'une valeur beaucoup plus forte (100 à 1000 fois), mais elle a l'inconvénient d'être au moins 4 fois plus grosse.

Figure 6.6. Un exemple d'écran multimédia en version V+ (6.6a, haut) et V- (6.6b, bas).  
NB. D'autres différences entre V+ et V-, liées à la dynamique et à l'interactivité, ne sont pas représentées sur cet exemple (Rouet, Choplin & Dubois, in prep.).

Vingt six élèves volontaires issus de trois écoles d'ingénieurs ont participé à deux séances de travail par groupes de 2 à 9 élèves. Lors de la première séance les élèves passaient plusieurs questionnaires destinés à évaluer leurs connaissances et leur capacité visuospatiale. Lors de la deuxième séance (de une à trois semaines après) les élèves étudiaient successivement les deux articles multimédias, l'un en V+, l'autre en V-. Les articles étaient présentés sur l'écran d'un ordinateur de type PC. Le croisement des domaines et des versions ainsi que l'ordre de passation des deux domaines étaient contrebalancés. Par ailleurs, la moitié des élèves recevait une consigne de mémorisation (groupe M) alors que l'autre recevait une consigne de résolution d'un problème dont l'énoncé leur était fourni (groupe P). Les deux groupes étaient équilibrés quant au sexe, à l'école d'origine et aux différents pré-tests. La période d'étude se divisait en deux parties: d'abord, l'élève devait visionner l'ensemble de l'article en évitant les retours en arrière; puis

il pouvait réviser certaines parties de l'article dans la limite d'un temps total d'étude de 20 minutes.

Après l'étude de chaque article-lien, l'élève répondait à un questionnaire de connaissances factuelles et devait résoudre un problème de conception à partir des concepts présentés dans l'article. Les questions et l'énoncé de problème étaient les mêmes que lors du pré-test. De plus on demandait à l'élève de donner son appréciation sur chaque article-lien et de comparer les deux articles-liens.

### 6.5.2.2. Résultats préliminaires

Le questionnaire factuel et les réponses au problème ont été notés sur une échelle en 10 points. Une procédure de double codage en aveugle a permis de vérifier la fiabilité de la grille de codage employée. Les résultats du problème sont présentés dans la Figure 6.6. L'augmentation des scores du pré-test au post-test est très importante ( $F(1,24)=112,62, p=0$ ). Les élèves ont pu trouver dans les articles-liens matière à fournir des réponses de bien meilleure qualité que celles qu'ils pouvaient donner quelques semaines auparavant, lors du pré-test. On observe une interaction phase x tâche ( $F(1,24)=7,21, p<.05$ ). Lors du pré-test les groupes M et P donnent des réponses équivalentes. En revanche, lors du post-test les élèves du groupe P donnent des réponses de meilleure qualité ( $F(1,24)=10,62, p<.01$ ).

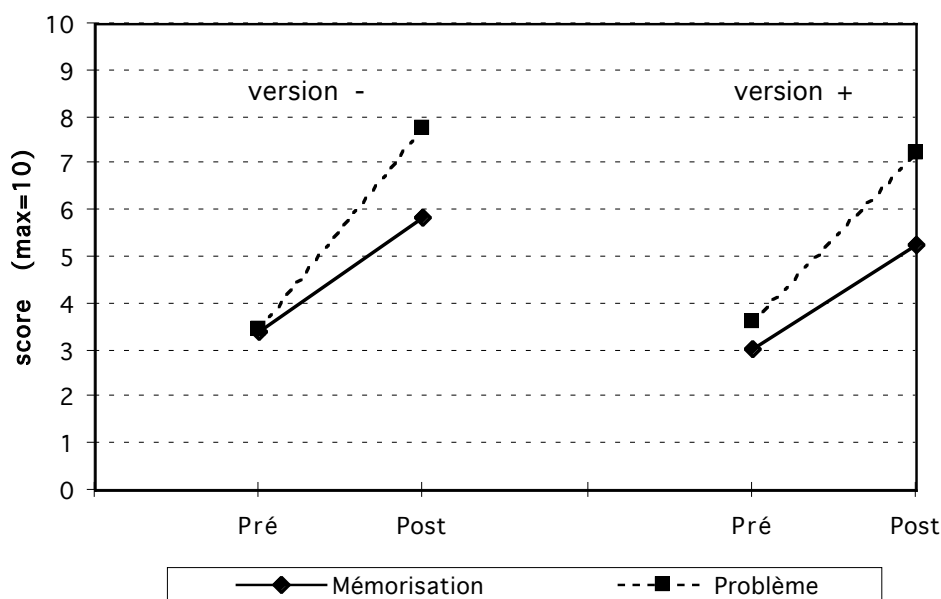


Figure 6.6. Scores de résolution de problème au pré- et post-test selon la consigne et la version étudiée.

En ce qui concerne le questionnaire factuel, on observe une très forte augmentation des scores du pré- au post-test, ceci dans toutes les conditions ( $F(1,20)=378,65, p=0$ ). Les facteurs "tâche" et "version" n'ont pas d'effet significatif. Cependant il existe une interaction entre la tâche et la phase de test ( $F(1,20)= 4,49, p<.05$ ). Cette fois, l'augmentation est légèrement plus forte pour les élèves du groupe M que pour ceux du groupe P. Une comparaison partielle indique que lors du post-test les participants du groupe M ont des résultats plus élevés que ceux du groupe P ( $F(1,20)= 4,40, p<.05$ ). La

version du logiciel est sans effet significatif sur les performances en compréhension, qu'il s'agisse de mémorisation factuelle ou de résolution de problème.

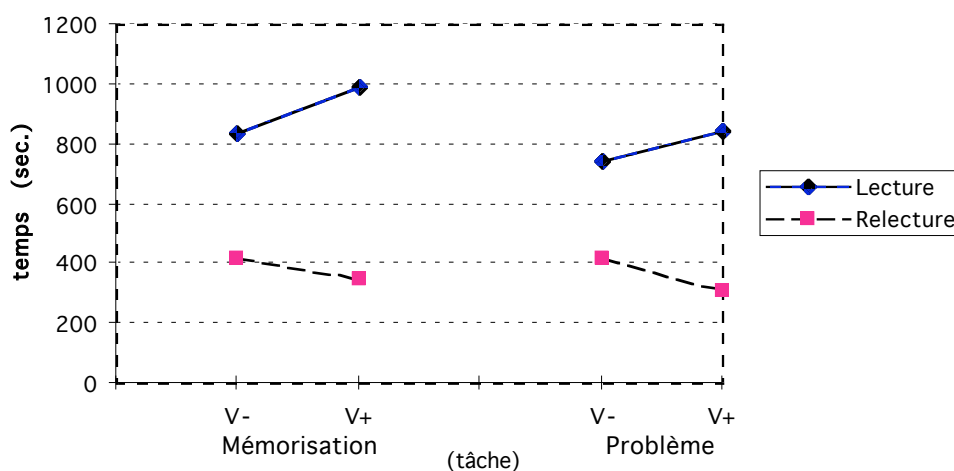


Figure 6.7: Répartition des temps d'étude initiale et de révision selon la consigne et la version.

La Figure 6.7 montre la répartition des temps d'étude entre la phase initiale et la relecture, dans les différentes conditions. Une analyse de la variance pour plan mixte avec la tâche comme facteur intergroupe et la version comme facteur intragroupe a été réalisée sur la lecture initiale de l'article. La V+ entraîne un allongement du temps de l'ordre de deux minutes ( $F(1,24)=12,70$ ,  $p<.01$ ). Les élèves du groupe M semblent avoir passé plus de temps à la lecture initiale que ceux du groupe P, mais la différence n'est que marginale ( $F(1,24)=3,62$ ,  $p<.10$ ).

Nous avons effectué une analyse plus détaillée des parcours réalisés par les élèves durant la phase initiale et la relecture. Ces parcours sont complexes et sujets à d'importantes différences individuelles, c'est pourquoi il est difficile d'en dégager des tendances générales. Cependant, la Figure 6.8 révèle plusieurs phénomènes intéressants.

Concernant le temps de lecture initiale (Figure 6.8a), on note que c'est l'étude des éléments de contenu, et surtout du plan P20, qui prend le plus de temps. On constate sur ce graphique que l'allongement du temps de lecture de la V+ se distribue sur tous les constituants. Lors de la relecture (Figure 6.8b), le profil des temps est très différent dans les groupes M et P. Le groupe M révisé principalement les plans P10 et P20, alors que le groupe P révisé plutôt la fin de l'article, notamment la conclusion. Les élèves ont donc appliqué une stratégie de relecture assez étroitement liée aux demandes de la tâche.

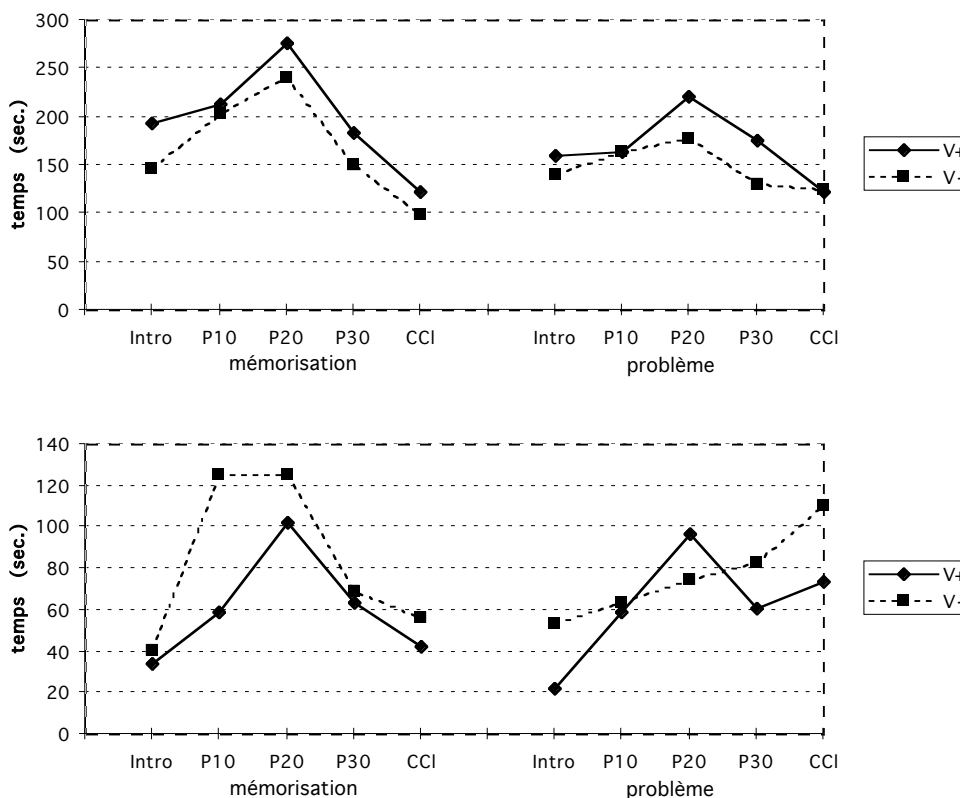


Figure 6.8: répartition des temps de lecture (4a, haut) et relecture (4b, bas) sur les différents constituants des articles-liens.

Toujours dans le groupe P, il faut noter la différence profonde entre les profils de la V+ et de la V-. En V- le temps de relecture croît presque linéairement avec l'ordre chronologique des plans. En V+ les étudiants consacrent plus de temps à P20, et relativement moins à l'introduction et à la conclusion. Cette centration sur le contenu, sans doute plus riche mais aussi plus lourd à traiter, au détriment des éléments de réflexion que sont l'introduction et la conclusion constituent peut-être un début d'explication à l'absence d'effet de la version + sur la qualité des réponses au problème.

### 6.5.2.3. Discussion et perspectives

L'absence d'effet des animations (V+) sur la compréhension vient s'ajouter à une série d'études récentes qui ont observé des résultats nuls voire négatifs des animations sur la compréhension (Wright et al., 1999; Schnotz, in press; Lowe, in press 2). Il faut prendre garde à ne pas en tirer de conclusions hâtives: il ne s'agit que d'une absence d'effet, ce qui en soi ne prouve rien. Toutefois, on peut risquer quelques conjectures quant aux origines de cette absence d'effet. D'abord, il est possible que des effets existent, mais qu'ils soient extrêmement faibles par rapport à la capacité d'apprentissage des sujets. On a pu apprécier l'augmentation considérable des scores entre le pré-test et le post-test. Compte tenu de l'intervalle entre les séances 1 et 2 (une à trois semaines séparées dans ce dernier cas par des congés scolaires) et des conditions de passation (petits groupes sans feedback explicite avant la fin de l'étude), il est peu probable qu'un facteur historique (communication entre élèves, voire prise d'informations dans d'autres sources entre les deux séances) soit seul responsable des effets. Les élèves ont donc trouvé dans les articles,

quelle que soit la version, matière à répondre aux questions de façon très correcte, quoique pas parfaite. Ensuite, il se peut que les effets des animations ne soient pas bien mesurés par nos post-tests. Seuls quelques items concernaient le fonctionnement de dispositifs représentés dans des animations. Des analyses complémentaires en cours suggèrent que la V+ a effectivement amélioré la compréhension pour ces items. Cependant, on trouve aussi des items dont la performance a été largement diminuée en version V+, sans doute parce que l'information visée par ces items était mieux présentée en V-. Ceci nous ramène à la nécessité d'articuler très soigneusement les objectifs d'apprentissage et les moyens multimédias (ou pas) mis en œuvre pour les atteindre.

La consigne d'étude a eu un effet significatif sur la performance des élèves. Le groupe mémorisation a davantage amélioré son score aux questions factuelles que le groupe problème, alors que ce dernier a obtenu de meilleurs résultats lors de la résolution de problème. Cet effet peut paraître trivial, d'un point de vue pédagogique. Toutefois, il faut le rapprocher d'absences d'effets des consignes parfois rapportées dans d'autres études (Rouet, Coirier et Favart (subm.), et section 3.4.3.2). L'effet important des consignes suggère que nous avons affaire à des étudiants dotés d'une bonne capacité d'autorégulation ce que confirme l'ajustement étroit entre consigne et parties de l'article effectivement relues par les élèves. L'absence d'interaction entre version et consigne d'étude signifie sans doute qu'il n'y a pas congruence entre les informations mises en avant dans la V+ et le type d'information requise pour résoudre le problème.

Enfin, la limitation du temps d'étude a pu avoir une influence sur les résultats. On a observé que les utilisateurs de la V+ passaient plus de temps sur la lecture initiale; il leur en restait donc moins pour réviser. De plus la répartition des temps d'étude montre que par rapport à la V-, les sujets semblent attirés vers les sections de l'article où se trouvent les animations les plus complexes. Ce résultat est lui aussi en faveur d'un problème de non congruence entre objectifs d'étude et type d'information proposée dans la V+.

Au total cette étude est le premier résultat d'une collaboration entre une équipe d'ingénierie pédagogique des médias et un laboratoire de psychologie cognitive, qui a reçu en 2000 le soutien du ministère de la recherche dans le cadre de l'action "Nouvelles Technologies et Cognition" du programme Cognitique. Les thèses en cours de Véronique Dubois (Paris V) et Hervé Potelle (Poitiers), également soutenues par le programme Cognitique, viseront à examiner plus en détail les processus cognitifs à l'œuvre dans l'apprentissage multimédia.

### 6.5.3. Quelques pistes de réflexion pour l'ergonomie des animations

Les études réalisées jusqu'à présent n'ont pas permis de prouver un effet positif des animations sur l'apprentissage. Cette absence d'effet constatée dans au moins quatre études récentes contraste fortement avec les effets généralement positifs des illustrations d'autant que les concepteurs de ces animations sont ici des psychologues spécialisés dans l'étude des processus d'apprentissage.

Pour tenter de comprendre d'où vient le problème, je suggère de comparer les hypothèses formulées dans le cadre d'études comme celles de Hegarty, Mayer et leurs collègues avec celles étudiées dans les expériences sur les animations. Dans le premier cas, les auteurs s'appuient sur des modèles de traitement de l'information extrêmement détaillés. A partir d'hypothèses précises sur les modalités de construction de modèles mentaux, ils dérivent des principes de construction des images dont ils prouvent l'efficacité dans des tâches elles aussi soigneusement définies.

Cette approche très analytique et qui prend la psychologie comme point de départ (et non pas comme simple instrument de validation) me semble la mieux à même d'aboutir à des résultats positifs. L'étude de Moreno et Mayer (Moreno & Mayer, 1999), que j'ai classée de façon impropre dans les études sur les illustrations (Cf. section 6.3.2) en est un bon exemple. Ce sont bien des animations informatiques qui sont utilisées dans cette étude: des séries de planches montrant pas à pas le déroulement d'un phénomène complexe au moyen de représentations graphiques évolutives. Ces animations sont très simples, presque minimales, mais elles relèvent de l'application de principes d'énonciation et d'intégration précis, et seulement de ceux-là.

Une autre illustration de cette approche se trouve dans les études de Mireille Bétrancourt et ses collègues (Bétrancourt, 1999; Bétrancourt & Bisseret, 1998; Bétrancourt & Tversky, subm.). Bétrancourt a étudié minutieusement les processus d'intégration de texte au sein d'images fixes et animées. Dans une première série d'expériences, Bétrancourt et Bisseret (1998) ont testé une technique d'intégration qu'ils appellent "escamot" dans la conception de documents multimédia pour l'apprentissage de notions techniques dans le domaine du bâtiment. Cette technique consiste à cacher les commentaires textuels d'une figure complexe (par ex., un dispositif d'évacuation de fumées) derrière des étiquettes verbales simples (par ex., le nom d'une partie du dispositif). Le sujet peut cliquer sur ces étiquettes, ce qui entraîne l'apparition du texte correspondant au sein d'une fenêtre. La technique des escamots réduit la densité informative du document, met en relief l'information verbale, et permet à l'utilisateur de structurer son exploration de la figure.

Bétrancourt et Bisseret observent en effet que cette technique augmente la mémorisation des étiquettes et des relations étiquettes-figure, par rapport à une présentation intégrée complète de l'information sur un seul écran, elle même supérieure à une présentation séparée de l'image et des commentaires (principe d'intégration spatiale; Moreno & Mayer, 1999). Dans une seconde expérience, Bétrancourt et Bisseret répliquent ces résultats avec un matériel à contenu procédural (sélection d'un diamètre de conduite en fonction du volume et de la vitesse d'évacuation de fumées). Bien qu'il ne s'agisse pas à proprement parler d'animation, cette étude montre comment l'application méthodique de principes de gestion de la charge cognitive et d'intégration texte-graphique permet de concevoir des dispositifs interactifs efficaces pour la présentation d'informations sur écran.

Dans une autre étude, Bétrancourt (1999) a montré que la mise en séquence de trois schémas représentant le fonctionnement du système de régulation de la température chez l'homme influence la façon dont les sujets rappellent ultérieurement ce phénomène, ainsi que la production d'inférences: une mise en séquence "procédurale" qui montre successivement les flux d'informations impliqués dans la régulation de la température (informations électriques, électrochimiques et chimiques) favorise la production d'inférences procédurales, alors que la mise en séquence "fonctionnelle" (selon les groupes d'organes impliqués: organes-cibles, glandes endocrines et système nerveux) favorise les inférences fonctionnelles. Ces deux présentations dynamiques tendent à donner des résultats meilleurs qu'une présentation statique, mais les différences ne sont pas toutes statistiquement significatives. Enfin, Bétrancourt et Tversky (subm.) montrent que l'ordre de présentation séquentielle des éléments d'un tableau ou d'une carte influence l'ordre de rappel, mais pas la quantité d'éléments rappelés. L'ordre de présentation influencerait l'organisation du modèle mental mais pas sa richesse ni son degré d'intégration (en tout cas pas ici).

## 6.6. Synthèse et conclusion

Le langage et les images sont les deux principaux moyens par lesquels des situations, objets ou concepts peuvent être représentés et communiqués de façon stable. D'un point de vue sémiotique, langage et images ne possèdent pas les mêmes propriétés. Entre autres différences, le langage donne naissance à des représentations descriptives générales mais sélectives et fragiles, alors que les images permettent des représentations décriptives plus spécifiques, mais complètes et robustes (Schnotz, in press). Ces représentations peuvent être utilisées de façon complémentaire pour favoriser la construction de modèles mentaux. A la complémentarité sémiotique des représentations visuelles et verbales s'ajoute la complémentarité modale dans le cas d'un emploi du langage oral. Pour observer un gain objectif en termes de compréhension, le texte et l'image doivent être référentiellement cohérents, et leur présentation matérielle doit respecter certains principes d'intégration (Moreno & Mayer, 1999). Il faut de plus que l'image remplisse une fonction déictique précise. Elle s'avère efficace pour la représentation des relations spatiales ou fonctionnelles entre constituants d'un référent complexe, mais peut également favoriser la mémorisation des détails (Waddill & McDaniel, 1992).

Cependant, les individus ne sont pas tous égaux devant les documents multimédia. Certaines dimensions cognitives au premier rang desquelles la capacité visuospatiale sont responsables de différences individuelles importantes dans la compréhension de ces documents. A une hypothèse généreuse qui voudrait que l'image compense les déficits des élèves les plus faibles, l'évidence doit nous conduire à préférer une hypothèse d'enrichissement sélectif selon laquelle les images profitent d'abord aux élèves imageurs. Selon certains auteurs, elles leur permettent de mettre en œuvre des processus d'intégration qu'ils n'appliquent pas spontanément. L'image peut ainsi aider des élèves ayant peu de connaissances ou ayant de faibles capacités inférentielles. La question de la facilitation de la compréhension chez les élèves faiblement imageurs est quant à elle encore largement ouverte.

Enfin, les premiers travaux sur les animations graphiques n'ont pas donné de résultats probants. Il semble y avoir un hiatus entre des travaux qui utilisent des cadres conceptuels précis permettant de développer des animations minimalistes mais efficaces, et ceux qui privilégient une approche globale, mettant en jeu des systèmes d'informations plus riches et plus complexes, et qui n'ont pas démontré de bénéfices clairs en termes de compréhension ou d'apprentissage. Parmi les problèmes encore à étudier, il faut noter celui du contrôle par l'utilisateur du déroulement des séquences (autorégulation du flux d'informations). Ce contrôle semble indispensable mais, comme d'ailleurs la prise de notes, il est probablement dépendant des capacités métacognitives du sujet (Hacker, 1999). Avec ou sans contrôle, les séquences dynamiques complexes entraînent un accroissement important de la charge cognitive ce qui devrait amener les concepteurs à ne les utiliser que de façon parcimonieuse et dans les cas où l'animation apporte une information précise et essentielle (Schnotz, in prep.).