

Université de Montréal

# **Modélisation de l'engagement et de la charge mentale de travail dans les Systèmes Tutoriels Intelligents**

par

Maher Chaouachi

Département d'Informatique et de Recherche Opérationnelle  
Faculté des Arts et des Sciences

Thèse présentée à la Faculté des Arts et des Sciences  
en vue de l'obtention du grade de Philosophiae Doctor (Ph.D.)  
en Informatique

Septembre 2014

© Maher Chaouachi, 2014



Université de Montréal  
Faculté des Arts et des Sciences

Cette thèse intitulée :

**Modélisation de l'engagement et de la charge mentale de  
travail dans les Systèmes Tutoriels Intelligents**

Présentée par :  
Maher Chaouachi

a été évaluée par un jury composé des personnes suivantes :

Jian-Yun Nie, président-rapporteur  
Claude Frasson, directeur de recherche  
Eugène Syriani, membre du jury  
Stefano A. Cerri, examinateur externe  
Carole Groleau, représentante du doyen



# Résumé

Les récents avancements en sciences cognitives, psychologie et neurosciences, ont démontré que les émotions et les processus cognitifs sont intimement reliés. Ce constat a donné lieu à une nouvelle génération de Systèmes Tutoriels Intelligents (STI) dont la logique d'adaptation repose sur une considération de la dimension émotionnelle et affective de l'apprenant.

Ces systèmes, connus sous le nom de Systèmes Tutoriels Émotionnellement Intelligents (STEI), cherchent à se doter des facultés des tuteurs humains dans leurs capacités à détecter, comprendre et s'adapter intuitivement en fonction de l'état émotionnel des apprenants. Toutefois, en dépit du nombre important de travaux portant sur la modélisation émotionnelle, les différents résultats empiriques ont démontré que les STEI actuels n'arrivent pas à avoir un impact significatif sur les performances et les réactions émotionnelles des apprenants. Ces limites sont principalement dues à la complexité du concept émotionnel qui rend sa modélisation difficile et son interprétation ambiguë.

Dans cette thèse, nous proposons d'augmenter les STEI des indicateurs d'états mentaux d'engagement et de charge mentale de travail. Ces états mentaux ont l'avantage d'englober à la fois une dimension affective et cognitive. Pour cela, nous allons, dans une première partie, présenter une approche de modélisation de ces indicateurs à partir des données de l'activité cérébrale des apprenants. Dans une seconde partie, nous allons intégrer ces modèles dans un STEI capable d'adapter en temps réel le processus d'apprentissage en fonction de ces indicateurs.

**Mots-clés** : Systèmes Tutoriels Intelligents, engagement mental, charge mentale de travail, adaptation en temps réel, émotions, électroencéphalographie (EEG).

# Abstract

Recent advances in cognitive science, psychology and neuroscience have shown that emotions and cognitive processes are closely intertwined. This fact has given rise to a new generation of Intelligent Tutoring Systems (ITS) whose adaptive logic is based on the consideration of the learner's emotional and affective dimension.

These systems, known as Emotionally Intelligent Tutoring Systems (EITS), seek to acquire the human tutors' ability in detecting, understanding and adapting to the learners' emotional state. However, despite the large body of work on emotional modeling, several empirical results showed that current EITS fail to have a significant impact on the learners' performance and emotional reactions. These limitations are mainly due to the complexity of the emotional concept which makes its modeling difficult and its interpretation ambiguous.

In this thesis we propose to increase EITS with mental state indicators of engagement and mental workload. These mental states have the advantage to include both affective and cognitive dimensions. To this end, we first present an approach to modeling these indicators from the learners' brain activity data. In the second part, we will integrate these models into an EITS able to adapt in real time the learning process according to these indicators.

**Keywords** : Intelligent Tutoring Systems, mental engagement, mental workload, real-time adaptation, emotions, electroencephalography (EEG).

# Table des matières

<b>CHAPITRE 1. INTRODUCTION .....</b>	<b>1</b>
1.1 CONTEXTE GÉNÉRAL .....	1
1.2 MOTIVATIONS.....	3
1.3 OBJECTIFS DE RECHERCHE .....	5
1.4 ORGANISATION DU DOCUMENT .....	6
<b>CHAPITRE 2. ÉTAT DE L'ART.....</b>	<b>9</b>
2.1 INTRODUCTION.....	9
2.2 HISTORIQUE .....	10
2.3 LES SYSTÈMES TUTORIELS INTELLIGENTS.....	13
2.3.1 <i>Le module d'interface.....</i>	<i>14</i>
2.3.2 <i>Le module expert.....</i>	<i>15</i>
2.3.3 <i>Le module apprenant.....</i>	<i>17</i>
2.3.4 <i>Le module tuteur.....</i>	<i>18</i>
2.4 AXES DE RECHERCHE ET LIMITES DES STI .....	19
2.5 INTÉGRATION DE LA DIMENSION ÉMOTIONNELLE DANS LES STI .....	22
2.5.1 <i>Les émotions et l'apprentissage.....</i>	<i>22</i>
2.5.2 <i>L'intelligence émotionnelle.....</i>	<i>23</i>
2.5.3 <i>L'informatique affective et les Systèmes Tutoriels Émotionnellement Intelligents .....</i>	<i>24</i>
2.6 REPRÉSENTATION DES ÉMOTIONS.....	26
2.6.1 <i>Le phénomène émotionnel.....</i>	<i>26</i>
2.6.2 <i>Approches de modélisation des émotions.....</i>	<i>28</i>
2.7 MOYENS DE DÉTECTION DES ÉMOTIONS DANS LES STEI.....	31
2.7.1 <i>L'auto-évaluation.....</i>	<i>31</i>
2.7.2 <i>Les mesures comportementales.....</i>	<i>34</i>
2.7.3 <i>Les mesures physiologiques.....</i>	<i>40</i>
2.8 STRATÉGIES D'ADAPTATION DANS LES STEI .....	43
2.9 PROBLÉMATIQUES DE RECHERCHE .....	48
2.9.1 <i>Vers l'intégration des indicateurs cérébraux d'états mentaux .....</i>	<i>52</i>
2.9.2 <i>Un système adaptatif selon l'état mental de l'apprenant.....</i>	<i>53</i>

<b>CHAPITRE 3. MODELISATION DE L'ENGAGEMENT MENTAL DES APPRENANTS .....</b>	<b>55</b>
3.1 INTRODUCTION.....	55
3.2 LA NOTION D'ENGAGEMENT MENTAL.....	56
3.2.1 <i>Modélisation de l'engagement dans les STEI .....</i>	<i>57</i>
3.3 CALCUL DE L'INDEX D'ENGAGEMENT MENTAL.....	59
3.3.1 <i>L'électroencéphalographie .....</i>	<i>60</i>
3.3.2 <i>Extraction de l'index d'engagement mental à partir des signaux EEG.....</i>	<i>62</i>
3.3.3 <i>Analyse de l'index d'engagement.....</i>	<i>64</i>
3.4 EXPERIENCE 1 : ÉTUDE DE L'INDEX D'ENGAGEMENT MENTAL DANS UN CONTEXTE D'APPRENTISSAGE	66
3.4.1 <i>Méthodologie et dispositif expérimental.....</i>	<i>67</i>
3.4.2 <i>Enregistrement des signaux EEG.....</i>	<i>71</i>
3.4.3 <i>Détection des émotions à partir des senseurs physiologiques.....</i>	<i>73</i>
3.5 RÉSULTATS EXPÉRIMENTAUX.....	76
3.5.1 <i>Analyse du comportement de l'index d'engagement par rapport aux performances....</i>	<i>77</i>
3.5.2 <i>Analyse de l'index d'engagement avec les indicateurs physiologiques émotionnels, le temps de réponse et les performances .....</i>	<i>78</i>
3.5.3 <i>Corrélations de l'indice d'engagement avec les signaux GSR et HR.....</i>	<i>80</i>
3.5.4 <i>Relation entre l'état émotionnel et l'index d'engagement.....</i>	<i>81</i>
3.6 DISCUSSION .....	83
3.7 CONCLUSION.....	84
<b>CHAPITRE 4. MODELISATION DE LA CHARGE MENTALE DE TRAVAIL DES APPRENANTS...87</b>	<b>87</b>
4.1 INTRODUCTION.....	87
4.2 LA CHARGE MENTALE DE TRAVAIL .....	89
4.2.1 <i>Modélisation de la charge mentale de travail .....</i>	<i>91</i>
4.3 EXPÉRIENCE 2 : MODÉLISATION DE LA CHARGE MENTALE DE TRAVAIL DANS UN CONTEXTE	D'APPRENTISSAGE .....
	93
4.3.1 <i>Protocole expérimental.....</i>	<i>95</i>
4.3.2 <i>Enregistrement des signaux physiologiques EEG.....</i>	<i>100</i>
4.3.3 <i>Mesures subjectives et objectives de la charge mentale de travail.....</i>	<i>101</i>
4.3.4 <i>Autres variables mesurées .....</i>	<i>103</i>
4.4 CONSTRUCTION DU MODÈLE PRÉDICTIF DE CMT .....	104

4.4.1	<i>Prétraitement des données EEG et extraction des attributs</i> .....	104
4.4.2	<i>Réduction de la dimensionnalité</i> .....	105
4.4.3	<i>Régression par processus gaussiens</i> .....	106
4.5	RÉSULTATS ET DISCUSSION .....	107
4.5.1	<i>Modélisation de la charge mentale de travail</i> .....	108
4.5.2	<i>Analyse de l'engagement mental</i> .....	115
4.5.3	<i>Relation entre les index cérébraux et les émotions</i> .....	117
4.6	CONCLUSION.....	119
<b>CHAPITRE 5. LE SYSTEME MENTOR</b> .....		<b>121</b>
5.1	INTRODUCTION.....	121
PARTIE I — PRESENTATION DU SYSTEME MENTOR.....		122
5.2	CAPTURE DES SIGNAUX EEG.....	126
5.2.1	<i>Dispositif matériel EEG</i> .....	126
5.2.2	<i>Acquisition des données EEG</i> .....	129
5.2.3	<i>Filtrage et transformation des données EEG</i> .....	129
5.3	LE MODE ENTRAINEMENT DE MENTOR.....	131
5.3.1	<i>Exercices d'entraînement cérébral</i> .....	132
5.3.2	<i>Collecte des données dans le mode Entraînement</i> .....	133
5.3.3	<i>Réduction de la dimensionnalité des données</i> .....	135
5.3.4	<i>Apprentissage du modèle de charge mentale de travail</i> .....	136
5.4	CALCUL DES INDEX CEREBRAUX.....	142
5.4.1	<i>L'index de charge mentale de travail</i> .....	142
5.4.2	<i>L'index d'engagement</i> .....	142
5.5	LE MODE AFFICHAGE DE MENTOR .....	143
5.5.1	<i>Le module d'analyse de MENTOR</i> .....	144
5.6	LE MODE APPRENTISSAGE DE MENTOR.....	146
5.6.1	<i>Le curriculum de MENTOR</i> .....	148
5.6.2	<i>Le module décisionnel de MENTOR</i> .....	150
5.6.3	<i>Fonctionnement du mode Apprentissage</i> .....	156
PARTIE II — EXPERIENCE 3 : ÉVALUATION DU SYSTEME MENTOR.....		157
5.7	METHODOLOGIE ET PROTOCOLE EXPERIMENTAL .....	158
5.7.1	<i>Participants</i> .....	160

5.7.2	<i>Déroulement de l'expérience.....</i>	160
5.7.3	<i>Enregistrement des émotions.....</i>	164
5.7.4	<i>Paramétrage du module d'analyse.....</i>	165
5.8	RESULTATS ET DISCUSSION.....	165
5.8.1	<i>Analyse des performances d'apprentissage.....</i>	165
5.8.2	<i>Analyse des mesures subjectives.....</i>	167
5.8.3	<i>Comparaison des réactions émotionnelles.....</i>	167
5.8.4	<i>Analyse du comportement des index cérébraux.....</i>	169
5.9	SYNTHESE ET CONCLUSION.....	171
<b>CHAPITRE 6. CONCLUSION.....</b>		<b>175</b>
6.1	CONTRIBUTIONS.....	175
6.2	TRAVAUX FUTURS.....	178
<b>BIBLIOGRAPHIE.....</b>		<b>180</b>
<b>ANNEXE A. REGRESSION PAR PROCESSUS GAUSSIENS.....</b>		<b>197</b>
<b>ANNEXE B. L'INSTRUMENT NASA TLX.....</b>		<b>200</b>
<b>PUBLICATIONS.....</b>		<b>201</b>

## Liste des tableaux

Tableau 3.1 – Exemples de fréquences du signal EEG (Demos, 2005).....	62
Tableau 3.2 – Comparaison des performances entre les groupes G1 et G2.....	77
Tableau 5.1 – Exercices cognitifs réalisés durant le premier jour de l’expérience.....	161
Tableau 5.2 – Performances des apprenants dans les deux groupes avant et après le tutoriel. Les valeurs qui diffèrent significativement ont des indices différents ( $p < 0.05$ ).....	166
Tableau 5.3 – Statistiques descriptives des proportions des émotions pour chaque condition expérimentale .....	168



## Liste des figures

Figure 2.1 – Machine à enseigner de (Crowder, 1959).....	10
Figure 2.2 – Architecture d'un STI.....	13
Figure 2.3 – Exemples de dispositifs d'entrée-sortie non standard. ....	15
Figure 2.4 – Courbe de taille d'effet des différents moyens de tutorat.....	21
Figure 2.5 – Modèle « circumplex » de Russel, tirée de Feldman-Barrett & Russell (1998). .	30
Figure 2.6 – Exemples d'instruments d'auto-évaluation selon l'approche dimensionnelle des émotions : [a] grille d'affect de Russel (1986), [b] l'échelle « Self-Assessment Manikin scale ». ....	32
Figure 2.7 – Exemple de codage FACS d'une expression faciale qui exprime une légère confusion. L'image est tirée de l'expérience de Mcdaniel et ses collègues (2007).....	35
Figure 2.8 – Étapes de la reconnaissance des émotions à partir des expressions faciales .....	36
Figure 2.9 – Exemple de chaise utilisée pour la détection émotionnelle .....	38
Figure 2.10 – Exemple de senseurs physiologiques. [a] et [b] représentent respectivement un bracelet et un gant utilisés pour mesurer la réponse galvanique de la peau, [c] est un capteur de pression du volume sanguin et [d] représente le système EEG de NeuroSky. ....	41
Figure 2.11 – Interface du système de Prendinger et Ishizuka (2005).....	44
Figure 2.12 – Exemples de réactions émotionnelles exprimées par l'agent virtuel de Burleson (2006).....	45
Figure 2.13 – Dispositifs de détection des émotions utilisés dans le système Autotutor (D'Mello et al., 2007).....	47
Figure 3.1 – Exemple de signal EEG sur une durée d'une seconde .....	60
Figure 3.2 – Engagement mental positif sur une période d'une minute .....	65
Figure 3.3 – Dispositif expérimental de la première expérience .....	68
Figure 3.4 – Emplacements des canaux EEG .....	71
Figure 3.5 – Exemple d'un signal BVP.....	73
Figure 3.6 – Transposition des signaux physiologiques BVP et SC pour la détection des émotions.....	75

Figure 3.7 – Espace bidimensionnel valence/activation des émotions avec les 4 quadrants ....	76
Figure 3.8 – Changement de l’indice EEG d’engagement mental à travers les quadrants émotionnels Q1 et Q2.....	81
Figure 3.9 – Moyenne de l’indice d’engagement selon les quatre quadrants.....	82
Figure 4.1 – Interface de l’environnement de trigonométrie utilisé dans la deuxième expérience .....	98
Figure 4.2 – Modèle bidimensionnel valence/activation utilisé dans la deuxième étude .....	103
Figure 4.3 – Approche utilisée pour la construction du modèle de CMT .....	104
Figure 4.4 – Procédé de recouvrement pour la segmentation des données EEG.....	105
Figure 4.5 – Moyennes des estimations subjectives de la CMT (NASA_TLX) à travers les différents niveaux de difficulté des tâches d’entraînement cérébral : Empan Numérique (6 niveaux de difficulté), Empan Numérique Inversé (5 niveaux de difficulté) et Test de Logique et de Calcul Mental (3 niveaux de difficulté).....	109
Figure 4.6 – Corrélations bivariées entre le score NASA_TLX et les mesures objectives (performances et temps de réponse) .....	109
Figure 4.7 – Corrélations bivariées entre le score NASA_TLX et l’index EEG_CMT de charge mentale de travail dérivé par notre modèle.....	111
Figure 4.8 – Moyenne des mesures EEG_CMT sur chaque activité de l’environnement d’apprentissage.....	112
Figure 4.9 – Évolution de l’index EEG d’engagement mental à travers les activités d’apprentissage.....	116
Figure 4.10 – Moyennes des index cérébraux (à gauche l’index d’engagement et à droite la charge mentale de travail) à travers les quatre quadrants d’émotions.....	117
Figure 4.11 – Correspondance entre le modèle bidimensionnel valence/activation des émotions et les indicateurs cérébraux d’engagement et de charge mentale de travail. ....	118
Figure 5.1 – Aperçu du fonctionnement général du système MENTOR. ....	125
Figure 5.2 – Dispositif EEG Emotiv Epoc utilisé. De gauche à droite : le casque EEG sans fil, le récepteur USB et le kit d’hydratation des senseurs. ....	126
Figure 5.3 – Positionnement des 16 électrodes du casque EEG.....	127
Figure 5.4 – Suivi de l’impédance des électrodes à travers l’application Emotiv TestBench.	128

Figure 5.5 – Filtrage et transformation des données EEG : [a] Formation d’un bloc EEG et élimination du bruit. [b] Redécoupage des segments EEG. ....	130
Figure 5.6 – Procédure de collecte des données d’entraînement du modèle de charge mentale de travail. ....	134
Figure 5.7 – Réduction de dimensionnalité des données EEG dans le mode Entraînement : la première interface (en haut) illustre la matrice des bandes de fréquences de 1 Hz (Bins) extraites du signal EEG, la deuxième interface (en bas) illustre les composantes du signal EEG retenu par l’algorithme d’ACP .....	137
Figure 5.8 – Interface du mode Entraînement de MENTOR .....	141
Figure 5.9 – Interface du mode Affichage de MENTOR.....	145
Figure 5.10 – Architecture d’un STI augmentée par les indicateurs d’états mentaux .....	147
Figure 5.11 – Interface du mode Apprentissage du système MENTOR.....	157
Figure 5.12 – Déroulement du processus expérimental de la deuxième journée des expérimentations .....	162



# Chapitre 1

## Introduction

### 1.1 Contexte général

Les Systèmes Tutoriels Intelligents (STI) ont été développés dans la perspective de fournir un enseignement adapté en fonction des caractéristiques et des besoins de l'apprenant. Ces systèmes — qui représentent la forme la plus évoluée des environnements d'apprentissage — sont dotés de capacités de raisonnement et d'adaptation qui leur permettent de produire un comportement *intelligent* par rapport au déroulement du processus d'apprentissage. Toutes ces caractéristiques font que les STI sont de plus en plus utilisés dans les sociétés modernes comme supports ou alternatives aux environnements éducatifs traditionnels (tels que l'enseignement en classe), afin de répondre aux besoins sans cesse croissants des individus en matière de formation (Frasson, 1991; Koedinger et al., 1996; Symonds, 2001; VanLehn, 2011).

D'une manière générale, le degré d'intelligence d'un STI dépend principalement de l'approche utilisée par les concepteurs du système dans la modélisation des apprenants (Shute, 1995). En effet, le modèle apprenant contient toutes les connaissances sur l'apprenant qui serviront à orienter toute la logique d'adaptation à savoir le séquençement des activités d'apprentissage et l'assistance pédagogique (Brusilovsky et al., 1996; Nwana, 1990).

Les chercheurs dans le domaine des STI ont pendant longtemps privilégié l'aspect *purement cognitif* dans leurs modélisations de l'apprenant (Picard et al., 2004). Le parcours, les performances, les solutions, les erreurs, les actions, etc. sont méticuleusement analysés par les STI afin de déterminer le niveau de connaissances et de raisonnement des apprenants et d'adapter leurs interactions en conséquence (Stern et al., 1998). Toutefois, cette logique d'adaptation entièrement basée sur des critères cognitifs limitait considérablement la qualité d'enseignement des STI (Lane, 2006; VanLehn, 2011).

- En effet, cette façon de concevoir le fonctionnement des premières générations de STI réduisait le processus d'apprentissage à un simple processus d'acquisition et de traitement de connaissances et négligeait une dimension fondamentale de l'apprentissage humain qui est la dimension *émotionnelle et affective* (Phelps, 2006; Picard et al., 2004). Par exemple, une mauvaise réponse de la part d'un apprenant n'est pas automatiquement synonyme d'insuffisances en termes de connaissances, mais peut être également le résultat d'un manque de réflexion et d'investissement causé par un état émotionnel négatif tel que l'ennui (Baker et al., 2010; Rodrigo et al., 2008).

Ainsi, avec l'émergence d'un nouveau courant de recherche, connu sous le nom de *l'informatique affective* (Picard, 1997), qui s'intéresse à l'intégration d'une forme d'*intelligence émotionnelle* (Mayer et al., 1997) dans l'interaction homme-machine en général; de plus en plus de travaux ont commencé à prendre en compte certains aspects émotionnels et affectifs dans l'interaction du STI avec l'apprenant. On parle alors d'une nouvelle génération de systèmes capables de simuler le comportement d'un tuteur humain dans sa capacité à analyser et à prendre en compte intuitivement l'état émotionnel de l'apprenant lors de son enseignement. Ces systèmes — communément appelés Systèmes Tutoriels Émotionnellement Intelligents (STEI) (Ochs et al., 2004) ou Systèmes Tutoriels Affectifs (STA) — (Alexander et al., 2003; Nkambou, 2006) ont pour objectif d'améliorer non seulement les performances d'apprentissage, mais également l'*expérience émotionnelle* de l'apprenant à travers des stratégies d'intervention basées sur des considérations émotionnelles. Par exemple, à la manière d'un tuteur humain et contrairement un STI classique, un STEI peut proposer une aide à un apprenant confus même dans le cas où ses performances d'apprentissage sont bonnes (Forbes-Riley et al., 2009).

## 1.2 Motivations

Un des principaux objectifs d'un STEI est de modéliser les états émotionnels des apprenants (Ochs et al., 2004). De ce fait, les chercheurs ont recours à une panoplie de moyens de détection allant des simples questionnaires d'auto-évaluation qui interrogent directement les apprenants sur leurs états émotionnels (Baker et al., 2010; de Vicente et al., 2002 ; Graesser et al., 2007); à la détection d'émotions à partir d'indices comportementaux tels que les expressions faciales (McDaniel et al., 2007; Nkambou et al., 2004; Whitehill et al., 2008); en passant par l'utilisation de senseurs physiques capables de renseigner sur les émotions à partir de leurs manifestations physiologiques (Conati, 2002; Hussain et al., 2011; Kapoor et al., 2007; Mcquiggan et al., 2007b; Jraidi et al., 2013b).

En dépit des efforts constants visant à améliorer les capacités de détection émotionnelle dans les STEI, les techniques actuellement utilisées sont souvent inefficaces. Par exemple, les questionnaires d'auto-évaluations sont souvent biaisés par des considérations subjectives de la part de l'apprenant (Picard et al., 2005). La détection des émotions à partir des expressions faciales peut être inefficace dans la mesure où certaines expressions peuvent être ambiguës, voire indéchiffrables. De plus, certains apprenants sont introvertis et ne laissent pas transparaître leurs émotions (Russell et al., 2003). L'approche basée sur les senseurs physiologiques malgré la corrélation que peuvent avoir certains senseurs avec certaines manifestations émotionnelles, ne dispose pas de modèles génériques qui permettent de reconnaître d'une façon précise tous les types d'émotions (Arroyo-Palacios et al., 2008; Liao et al., 2006; Ward et al., 2003).

Toujours dans le même objectif de modélisation émotionnelle dans les STEI, il n'existe pas de consensus clair ni de théorie précise sur quelles émotions faut-il modéliser dans un contexte d'apprentissage. Certains travaux, par exemple, privilégient la modélisation d'un seul état émotionnel (tel que la confusion, ou la frustration) (Grafsgaard et al., 2011; Kapoor et al., 2007), d'autres en revanche, modélisent un ensemble d'émotions (D'Mello et al., 2009; D'Mello et al., 2008; Forbes-Riley et al., 2008; Jraidi et al., 2013a).

De plus, bien qu'il soit établi que les émotions jouent un rôle primordial dans le processus d'apprentissage (Phelps, 2006), la relation qui existe entre les émotions et l'apprentissage n'est pas toujours linéaire (Graesser et al., 2007). En particulier, l'association automatique des émotions négatives avec un état négatif pour l'apprentissage n'est pas toujours valide. Certaines émotions qualifiées de négatives à priori (telles que la frustration, le stress ou la confusion) ne sont pas toujours synonymes de conditions émotionnelles défavorables à l'apprentissage. Par exemple, l'état de frustration qui est souvent considéré comme un état négatif pour l'apprentissage peut, dans certains cas, révéler une certaine envie de bien faire ainsi qu'une certaine motivation de la part de l'apprenant (Gee, 2004).

Ainsi les limites de la modélisation des émotions dans les STEI se situent non seulement au niveau de la précision des techniques de détection émotionnelles, mais également au niveau du choix et de l'interprétation des émotions à modéliser (Picard et al., 2004). Dans cette thèse, nous proposons de remédier à cela grâce à l'utilisation des *indicateurs d'états mentaux*. En particulier, nous proposons d'utiliser une approche qui se base sur l'analyse d'indices d'état d'*engagement* et de *charge mentale de travail* extrait à partir de l'activité cérébrale de l'apprenant et plus précisément l'ÉlectroEncéphaloGraphie (EEG). En effet, les états mentaux sont généralement considérés comme des états affectifs complexes ayant une dimension à la fois affective et cognitive, ce qui facilite considérablement leurs interprétations (el Kaliouby, 2005; Scherer, 2005; Sobol-Shikler, 2011). De plus, en dépit du fait que l'état d'engagement et la charge mentale de travail aient été très étudiés en interaction homme-machine, leur modélisation dans les STEI est plutôt rare.

Par ailleurs, malgré de nombreux travaux de recherche menés sur la modélisation des émotions, très peu de STEI ont été développés pour détecter *en temps réel* les émotions des apprenants. De plus, les différentes études empiriques visant à analyser l'impact des stratégies d'adaptation utilisées dans les STEI ne démontrent pas d'effet positif statistiquement significatif, ni sur les performances de l'apprenant ni sur ses réponses émotionnelles (Beale et al., 2009; Burleson et al., 2007; Kim, 2005; Mori et al., 2003; Woolf et al., 2009).

Dans cette thèse nous proposons de développer un système tutoriel capable de détecter et d'adapter en temps réel le processus d'apprentissage selon les indicateurs d'engagement et

de la charge mentale de travail. Nous cherchons également à démontrer qu'un tel système peut améliorer significativement l'expérience d'apprentissage aussi bien au niveau des performances cognitives des apprenants qu'en termes de satisfaction par rapport au déroulement de leur apprentissage.

### **1.3 Objectifs de recherche**

Notre objectif dans cette thèse est d'améliorer le fonctionnement des STEI en intégrant des indicateurs cérébraux d'engagement et de charge mentale de travail. Plus précisément, nous proposons de réaliser ces deux objectifs suivants :

1. Modéliser l'engagement mental et la charge mentale de travail des apprenants : notre premier objectif consiste à augmenter les STEI avec des indicateurs fiables et capables de quantifier le niveau d'engagement et de charge mentale de travail des apprenants à partir de leurs activités cérébrales.
2. Intégrer ces indicateurs d'états mentaux dans un STEI : notre deuxième objectif est de montrer qu'il est possible pour un STEI d'intégrer ces indicateurs d'engagement et de charge de travail en temps réel, mais aussi d'adapter, en temps réel également, le processus d'apprentissage en fonction de ces indicateurs.

Afin de réaliser notre premier objectif, nous avons réalisé deux études expérimentales. Dans la première étude, nous avons testé, dans un contexte d'apprentissage, la validité d'un index d'engagement mental développé dans le cadre des travaux de recherche menée dans la « National Aeronautics and Space Administration » (NASA) pour mesurer le niveau d'attention des pilotes, à partir des données EEG (Pope et al., 1995). Pour cela, nous avons d'abord vérifié si cet index d'engagement peut fournir une indication pertinente sur les performances et le comportement des apprenants. Ensuite, nous avons analysé la relation qui pourrait exister entre les variations de l'index d'engagement mental et les états émotionnels des apprenants.

Dans la deuxième étude expérimentale, nous avons proposé une approche pour modéliser la charge mentale de travail des apprenants à partir de leurs données EEG. Nous

avons validé et évalué ce modèle dans un environnement d'apprentissage et examiner sa relation avec l'index d'engagement mental ainsi qu'avec les états émotionnels.

Pour réaliser notre deuxième objectif, nous avons développé un système tutoriel que nous appellerons le système MENTOR (« MENTAL tuTOR ») capable d'évaluer en temps réel l'état mental des apprenants en fonction d'indicateurs d'engagement et de charge mentale de travail et d'adapter le déroulement de la session d'apprentissage en conséquence.

Nous avons réalisé une troisième étude expérimentale dédiée à l'évaluation de notre système. L'objectif de cette étude a été de vérifier si l'intégration des indicateurs mentaux dans la logique d'adaptation d'un STI pouvait avoir un impact positif sur les performances d'apprentissage des apprenants. Nous avons également vérifié l'influence de cette intégration sur l'expérience de l'apprenant vis-à-vis du déroulement de son apprentissage. Et enfin, nous avons également examiné si un tel système était capable d'agir et de corriger les difficultés mentales de l'apprenant.

## **1.4 Organisation du document**

Le chapitre suivant de cette thèse sera consacré à l'analyse de la littérature. Nous commençons par présenter le cadre général de cette thèse, à savoir les STI, en décrivant l'historique de leur développement, les principaux modules qui les composent ainsi que leurs limites. Nous présentons par la suite l'importance de la dimension émotionnelle dans l'apprentissage et l'apparition de la nouvelle génération de STEI. Nous exposons également les principales approches utilisées par les STEI pour la modélisation des émotions ainsi que pour l'adaptation. Nous concluons ce chapitre par une présentation des limites des STEI actuels et de l'approche que nous allons utiliser pour y remédier.

Dans le chapitre 3, nous présentons le concept de l'engagement mental ainsi que les différentes approches utilisées pour la modéliser. Nous montrons également la technique que nous allons utiliser pour l'intégration de l'index d'engagement à savoir l'électroencéphalographie. Nous exposons par la suite la technique de calcul et d'interprétation

de cet index. Nous enchainons par la description de nos premières études expérimentales ainsi que les principaux résultats trouvés.

Dans le chapitre 4, nous présentons le concept de charge mentale et les principales approches permettant de la mesurer. Nous décrivons par la suite notre approche de modélisation de la charge mentale de travail ainsi que la deuxième étude expérimentale réalisée pour la valider. Nous présenterons ensuite les principaux résultats et conclusions obtenus.

Dans le chapitre 5, nous procédons à la description de notre système MENTOR. Nous présentons ses principaux modes de fonctionnement ainsi que sa logique d'adaptation en fonction des indicateurs d'engagement et de charge mentale de travail. Nous dévoilons ensuite notre troisième étude expérimentale réalisée pour tester l'efficacité de notre système et nous exposons les principaux résultats trouvés.

Dans la conclusion de cette thèse, nous résumons les contributions et les limites de nos recherches avant de présenter les travaux futurs que nous envisageons de réaliser afin d'améliorer notre approche.



## Chapitre 2

# État de l'art

### 2.1 Introduction

Ce chapitre est consacré à l'analyse de la littérature. Nous commencerons par présenter un bref historique de l'évolution des systèmes informatiques d'apprentissage en débutant par les toutes premières « machines à enseigner », puis les systèmes d'Enseignement Assisté par Ordinateur (EAO), jusqu'à arriver aux Environnements Interactifs pour l'Apprentissage Humain (EIAH), où se situent les Systèmes Tutoriels Intelligents (STI). Nous décrirons ensuite les différents modules qui assurent le fonctionnement d'un STI, à savoir : le module apprenant, le module expert, le module tuteur et le module d'interface.

À partir de la quatrième section du chapitre, nous aborderons les principaux axes et travaux de recherches entrepris pour faire évoluer ces systèmes. Ceci va nous amener vers l'intégration du concept de l'intelligence émotionnelle ainsi que de l'informatique affective dans les Interactions Homme-Machine, et plus particulièrement les environnements d'apprentissage. Nous verrons que la considération de ce facteur émotionnel a conduit à une deuxième génération de STI : les Systèmes Tutoriels Émotionnellement Intelligents (STEI).

Nous nous attarderons ensuite, à partir de la sixième section, sur les approches de modélisation et de détection des émotions dans les STEI. Puis, nous enchaînerons par une

revue des principales stratégies utilisées par ces systèmes pour adapter leur apprentissage selon l'état affectif de l'apprenant.

Nous clôturerons ce chapitre par les problématiques de recherche que nous tenterons de pallier dans le cadre de cette thèse.

## 2.2 Historique

Le principe d'adaptation dans les systèmes informatiques a été développé sous divers angles en vue d'améliorer l'utilisation des machines par les humains, et ce, dans plusieurs domaines d'application (transport, prise de décision, conception, communication, etc.). Toutefois, la mise en place de systèmes adaptatifs pour l'apprentissage et la formation par ordinateur a constitué un véritable saut qualitatif dans le déroulement de cette activité.



Figure 2.1 – Machine à enseigner de (Crowder, 1959)

En effet, les premiers systèmes informatiques d'apprentissage étaient conçus sous la forme de « machines à enseigner », qui fournissaient un enseignement programmé et linéaire. Leurs fonctionnements se résumaient typiquement, à présenter séquentiellement un ensemble de connaissances décomposées en sous-éléments, et à poser des questions suivies immédiatement de leurs réponses. Un exemple de tels systèmes est illustré dans la figure 2.1.

Progressivement, les recherches commencent à donner plus d'initiative à l'apprenant, et à envisager les programmes éducatifs plutôt comme des outils qui accompagnent son processus d'apprentissage. Les systèmes commencent à intégrer des mécanismes qui orientent l'apprenant dans ses choix des activités d'apprentissage, ainsi que des stratégies de redressement des erreurs commises. Un aspect plus algorithmique vient substituer le caractère linéaire des premiers programmes éducatifs avec l'apparition de ces systèmes d'Enseignement Assisté par Ordinateur (EAO). Cette catégorie de systèmes, appelés également didacticiels ou tuteurs-guides, vise à consolider les connaissances des apprenants en utilisant une gamme d'activités généralement sous forme de situation/problème. La réponse de l'apprenant est analysée par le système afin de décider, soit de continuer sur la même activité, soit de se brancher sur une autre activité.

La conception des systèmes d'EAO était principalement centrée sur la diffusion des informations à enseigner sur la forme de scénarii à transmettre à l'apprenant (dans un seul sens) ; l'impact de ces méthodes sur l'apprenant était ignoré. L'évolution de l'EAO en un Enseignement Interactif Assisté par Ordinateur (EIAO) a permis de recentrer les systèmes éducatifs plus sur les problèmes liés à l'apprentissage. L'accent est mis sur le processus de traitement de l'information, et l'apprentissage est considéré comme la résultante d'échanges interactifs entre l'apprenant et son environnement. L'EIAO considère l'apprenant comme un membre actif dans le processus d'apprentissage, qui agit et réagit relativement à l'environnement. Les systèmes commencent à adopter des stratégies d'accompagnement individualisées basées sur des paramètres et modèles cognitifs qui leur permettront de diriger leurs interactions avec l'apprenant. Le perfectionnement de ces systèmes avec l'augmentation de leurs capacités à raisonner et à adapter leurs comportements en fonction du processus d'apprentissage leur a permis d'acquérir un degré d'intelligence. On parle alors de systèmes d'Enseignement Intelligemment Assisté par Ordinateur.<sup>1</sup>

Aujourd'hui, le concept intégrateur d'Environnement Informatique ou Interactif pour l'Apprentissage Humain (EIAH) est utilisé pour désigner un environnement qui vise à susciter

---

<sup>1</sup> Le sigle EIAO est utilisé à la fois pour désigner l'Enseignement Intelligemment Assisté par Ordinateur et l'Enseignement Interactif Assisté par Ordinateur.

ou favoriser l'apprentissage humain en produisant des actions ou des réactions pertinentes et adéquates (Tchounikine, 2002).

On peut distinguer trois catégories majeures d'EIAH :

- *Les hypermédias* : systèmes plurifonctionnels destinés principalement à guider l'apprentissage en facilitant l'accès aux ressources pédagogiques. Ils englobent des technologies permettant la conception et la mise en page de présentations multimédias destinées à Internet ou Intranet. Les hypermédias adaptatifs permettent d'adapter le contenu ou la navigation en fonction des objectifs, préférences et connaissances de l'apprenant (Brusilovsky et al., 1996).
- *Les micromondes* : systèmes qui simulent des situations d'apprentissage en permettant à l'apprenant de faire évoluer progressivement ses connaissances en effectuant des actions physiques et mentales. L'apprenant est capable d'explorer à travers des opérations élémentaires et avec le minimum possible de contraintes un domaine de connaissances. L'apprentissage en utilisant les micromondes se fait à travers la manipulation et la découverte, mais aussi à travers les possibilités qu'offre le système à l'apprenant pour revenir sur ses actions, les comprendre et les corriger.
- *Les Systèmes Tutoriels Intelligents (STI)* : systèmes adaptatifs qui utilisent un ensemble de techniques permettant de fournir un enseignement qui tient compte des diversités des apprenants. Les STI disposent de capacités de raisonnement leur permettant d'ajuster et de contrôler l'apprentissage. Les systèmes peuvent entreprendre des interventions pour présenter, expliquer ou adapter le contenu pédagogique. Ils peuvent aussi détecter les erreurs ou les problèmes d'apprentissage, et déclencher les stratégies de correction requises le cas échéant. Les STI regroupent l'ensemble de technologies qui permettent d'assister les deux intervenants de l'apprentissage (apprenant et enseignant).

Nous décrivons dans la section suivante les principales composantes de l'architecture des STI ainsi que les principaux axes et travaux de recherches entrepris pour faire évoluer ces systèmes.

## 2.3 Les Systèmes Tutoriels Intelligents

Le développement des STI a commencé vers la fin des années 70 — début des années 80, dans le but d'améliorer les performances de l'EAO qui peine à rivaliser avec la qualité de l'enseignement donné par un tuteur humain. En effet, constitué d'un ensemble de programmes rigides, l'EAO se contentait principalement de générer des problèmes, présenter les solutions et évaluer les performances des apprenants. L'insuffisance de ces systèmes relativement à la complexité du processus d'apprentissage était évidente et le besoin d'évoluer vers des systèmes plus intelligents coïncidait avec une vague de progrès dans le domaine de l'intelligence artificielle (IA), l'éducation et la psychologie cognitive favorisant ainsi la migration vers les STI.

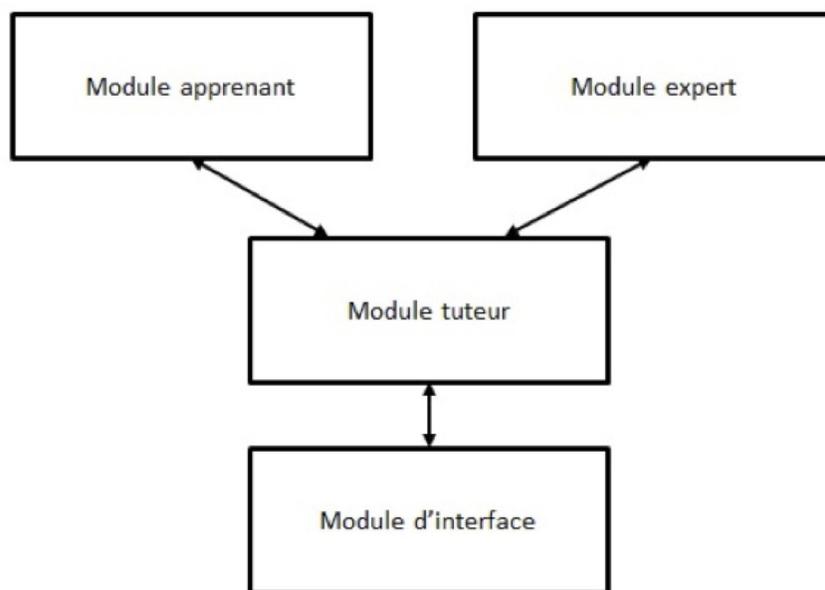


Figure 2.2 – Architecture d'un STI

Les STI visent à produire un comportement intelligent qu'on qualifierait « d'un enseignement de qualité » s'il est effectué par un humain (Psothka et al., 1988a). L'intégration des techniques IA permet aux STI de fournir un enseignement individualisé et de connaître quoi enseigner, à qui enseigner et surtout comment l'enseigner (Psothka et al., 1988a). Le terme Intelligent Tutoring Systems (ITS) a été cité pour la première fois en 1982. Bien qu'il existe plusieurs définitions, on peut voir les STI comme étant « un système informatique capable

d'adapter l'enseignement et d'interagir avec l'apprenant sans intervention humaine » (Pspotka et al., 1988b).

Il existe plusieurs architectures de STI. La plus répandue (figure 2.2) se compose essentiellement de 4 modules interdépendants où chaque module représente une entité abstraite qui traite certaines fonctionnalités du STI (Wenger, 1987). Côté implémentation, les frontières et les interactions entre ces modules ne sont pas explicitement précisées : les fonctionnalités d'un module peuvent être réparties sur plusieurs entités logicielles. Nous présentons dans ce qui suit une description de ces modules.

### **2.3.1 Le module d'interface**

La conception des STI sépare généralement le noyau fonctionnel du système — partie responsable du raisonnement et de l'adaptation — de l'interface partie qui englobe un ensemble de moyens d'*entrées* et de *sorties*. Le module d'interface représente la composante de communication du STI. Il prend en charge la liaison entre la représentation interne du système et l'apprenant ainsi que la forme finale de présentation des connaissances (Wenger, 1987).

L'importance de créer des interfaces de qualité provient de l'objectif même de l'Interaction Homme-Machine qui vise une meilleure « collaboration » entre les deux parties. Dans les systèmes actuels, le concept d'interface intelligente est de plus en plus mis en avant. Les interfaces sont de plus sensibles aux besoins et au comportement des apprenants. Elles anticipent leurs choix et préférences, prédisent leurs actions, filtrent les informations dont ils ont besoin (ou les présentent sous différents modes), génèrent des explications, engagent des dialogues, etc.

En plus des moyens d'entrée sortie standards (c.-à-d. clavier et souris), les interfaces intelligentes actuelles utilisent davantage de moyens non standard d'interaction plus adaptés à la tâche. Selon D. Engelbart, inventeur de la souris, « Si la simplicité est le seul critère valide, on serait resté aux tricycles et l'on passerait jamais à la bicyclette ». Les STI intègrent de plus

en plus de technologies sophistiquées et complexes (dispositifs, capteurs, environnements 3D, etc.) afin d'enrichir leurs interactions avec l'apprenant.

La figure 2.3 illustre trois exemples de dispositifs d'entrée sortie non standards. À gauche, le gant de données *Cyberglove*, qui capte et envoie des informations en provenance de 22 articulations ; utilisé par exemple pour le développement de systèmes éducatifs pour les sourds et malentendants (Adamo-Villani et al., 2007). Au centre, le système *Construct3D* pour l'enseignement de la géométrie (Kaufmann et al., 2007). L'apprenant utilise un stylo et un panel sans fil pour interagir avec le système. La partie gauche de cette même figure présente un exemple de ce que voit l'apprenant à travers le visiocasque. À droite, un exemple de lunettes contenant des capteurs de détection du mouvement des yeux utilisés pour suivre l'intérêt des apprenants par rapport au contenu éducatif présenté sur l'écran (D'Mello et al., 2012b).



Figure 2.3 – Exemples de dispositifs d'entrée-sortie non standard.

### 2.3.2 Le module expert

Appelé aussi curriculum, ce module définit l'expertise du domaine. Il contient une description et une codification des connaissances et des compétences à acquérir.

Il existe plusieurs manières de représenter les connaissances du domaine pour un STI. On peut citer les réseaux sémantiques, les graphes conceptuels, ou les graphes conceptuels multicouches nommés « curriculum knowledge transition network » (CKTN).

Le module expert doit également contenir des modèles permettant de raisonner sur les connaissances de l'apprenant. Ces modèles permettent au STI de déterminer les différences

qui existent entre les actions (ou les réponses) de l'apprenant avec celles attendues par le système, et ainsi de prendre les mesures d'adaptation nécessaires. Plusieurs modèles ont été utilisés pour développer cette fonction essentielle pour les STI :

- Le modèle de boîtes noires : cette approche utilise les actions de l'apprenant pour évaluer sa méthode de raisonnement sur le domaine. Les STI qui utilisent cette approche sont appelés *tuteurs réactifs*. Ils peuvent évaluer les solutions, corriger les erreurs, donner des instructions ; mais ils ne fournissent pas aux apprenants des explications sur leurs erreurs.
- Le modèle de boîtes de verres : cette approche se base sur la mise en place d'un système expert qui intègre des règles sur le domaine de connaissance ainsi que des règles pédagogiques. Le STI peut alors comparer le raisonnement de l'apprenant par rapport à l'expert et générer des explications relatives à ses erreurs et difficultés d'apprentissage.
- Le modèle cognitif : encouragés par les avancées réalisées dans les sciences cognitives, plusieurs chercheurs ont proposé d'intégrer cette dimension dans les STI. Ces modèles qui se basent sur un ensemble de règles et d'hypothèses sur la cognition humaine ont permis le développement de plusieurs STI capables de comprendre et de simuler le raisonnement d'un apprenant. On peut citer le *tuteur cognitif* PAT (Koedinger, K. R., & Sueker, E. L. F., 1996) qui se base sur le modèle « active control of thought-relational » (ACT-R) (Anderson, 1993), le système STEVE (Elliott et al., 1999) qui se base sur l'architecture cognitive « state operator and result » (SOAR) ou encore Cognitive Constructor (Samsonovich et al., 2008) qui se base sur une architecture cognitive inspirée de la biologie « Biologically Inspired Cognitive Architecture » (BICA).

En plus du suivi et de l'analyse du raisonnement et des connaissances de l'apprenant, ce module intègre aussi la génération de questions, réponses, explications relatives à la solution de l'expert, ainsi que des normes d'évaluation des performances des apprenants.

### 2.3.3 Le module apprenant

Comme son nom l'évoque, ce module renseigne sur toutes les informations pertinentes relatives à l'apprenant. Appelé également module de *diagnostic* ou *modèle de l'apprenant*, il permet d'apporter une mesure des connaissances de l'apprenant et une évaluation de sa progression.

La modélisation de l'apprenant représente un enjeu majeur dans la conception d'un STI, dans la mesure où toute la logique d'adaptation et d'individualisation de l'enseignement dépend de la nature et de la qualité des informations contenues dans ce module. La modélisation de l'apprenant permet alors au STI de prendre en compte les spécificités de chaque apprenant afin de favoriser son apprentissage. Les informations qui sont contenues dans ce module peuvent être considérées comme les paramètres des mécanismes d'interaction du système. De ce fait, la modélisation de l'apprenant a été un thème récurrent dans les STI et a fait l'objet de plusieurs évolutions et innovations visant à intégrer et extraire plusieurs facettes qui ont trait aux connaissances, compétences et comportement de l'apprenant (Webber et al., 2002).

Les informations contenues dans le modèle de l'apprenant peuvent être classées en deux catégories à savoir : les informations spécifiques au domaine et les informations indépendantes du domaine (Brusilovsky et al., 1994).

Les informations spécifiques au domaine permettent au STI d'évaluer les connaissances et les compétences de l'apprenant, en se basant sur des indicateurs de performance (comme les résultats ou les notes) ou sur les différentes interactions que l'apprenant opère avec l'environnement. Il s'agit de construire son profil cognitif et d'avoir une représentation fiable de ses connaissances et son raisonnement. Pour cela, plusieurs méthodes étroitement liées au module expert ont été considérées. La méthode de *recouvrement* (« *overlay* »), par exemple, consiste à considérer les connaissances de l'apprenant comme étant incluses dans celles de l'expert, et en fonction du parcours et des résultats d'apprentissage, celles-ci augmentent jusqu'à atteindre – idéalement – les connaissances de l'expert. La méthode *différentielle*, qui ressemble à la méthode de recouvrement, sépare, dans sa conception, les connaissances déjà

acquises par l'apprenant, des connaissances à acquérir. Les *modèles cognitifs* (Anderson et al., 1990) ont eux aussi été utilisés — comme pour le module expert d'ailleurs — pour identifier les besoins et les caractéristiques cognitives de l'apprenant.

Les informations indépendantes du domaine sont en rapport avec l'apprenant lui-même : ses objectifs, ses intérêts, ses traits de caractère, son humeur, ses aptitudes ou même son sexe et son âge. Ces attributs sont fournis par l'apprenant lui-même ou par des questionnaires préétablis appuyés par des théories psychologiques. Ces informations servent à mieux cerner les préférences de l'apprenant et donc à permettre au système d'élaborer sa stratégie pédagogique en conséquence.

### **2.3.4 Le module tuteur**

Ce module définit les interventions du STI visant à aider et accompagner l'apprenant dans le processus d'apprentissage. Appelé également modèle d'interaction, il a pour but d'appliquer les différentes stratégies pédagogiques, éducatives et psychologiques qu'utilise l'enseignant humain dans son interaction avec l'apprenant.

Ce module se base dans ces décisions sur les informations et les connaissances existantes dans le module apprenant et le module expert, afin de concevoir d'une manière adaptative un plan d'enseignement des éléments pédagogiques (Wenger, 1987). Ce plan détermine les concepts clés à enseigner parmi ceux proposés par le module expert, ainsi que la manière de les enseigner. Il contrôle subséquemment la masse et la forme d'informations à fournir. Selon Wenger (1987), le module tuteur contient les réponses aux questions suivantes :

- *Pourquoi intervenir ?* Ce module établit les objectifs précis de chaque intervention du système. Orientés par des considérations pédagogiques et stratégiques, les objectifs des interventions peuvent être par exemple d'encourager l'apprenant ou encore de lui fournir une connaissance.
- *Quand intervenir ?* Il s'agit de définir le moment opportun d'intervention. Décider du moment d'intervention est une tâche délicate dans la mesure où ceci présente un risque de dérangement pour l'apprenant. Ce modèle peut par exemple déterminer le moment

d'intervention en fonction des erreurs de l'apprenant. Sa stratégie peut alors consister à prévoir les erreurs et empêcher leurs occurrences. Le système peut aussi intervenir après l'erreur immédiatement ou en laissant un moment de réflexion à l'apprenant pour rectifier son erreur évitant ainsi le risque de le décourager.

- *Comment intervenir ?* Plusieurs stratégies pédagogiques ont été développées dans les STI. On peut citer la stratégie du compagnon (Chan, 1995) développée dans le but de créer un compagnon virtuel qui assiste l'apprenant dans les différentes tâches d'apprentissage. Ce compagnon peut jouer le rôle d'un perturbateur (Aïmeur et al., 2001), en contredisant l'apprenant et en remettant en question ses réponses le forçant ainsi à mieux réfléchir. Le compagnon peut aussi prendre le rôle du tuteur en posant des questions à l'apprenant afin de l'aider à mieux organiser ses connaissances lors de la réponse. Dans les deux cas, les notions de coopération et de compétition jouent un rôle dans le renforcement du processus d'apprentissage.

Le module tuteur peut aussi utiliser des stratégies d'interventions basées sur le guidage ou le coaching qui se déclenchent pour orienter l'apprenant lorsqu'il s'éloigne de la solution. Il peut aussi engager un processus d'interaction ciblée pour aider l'utilisateur à apprendre de ses erreurs ou de son blocage, en utilisant des stratégies élaborées du type « parlons-en » ou « expliquez-moi ».

## **2.4 Axes de recherche et limites des STI**

Il est très fréquent pour les chercheurs dans le domaine des STI de procéder en deux étapes (Lane, 2006). Dans la première étape, une observation des facteurs et des événements qui influencent le processus d'apprentissage est établie. Le but est d'analyser le comportement de l'apprenant, sa manière de penser, de raisonner, de généraliser, d'agir, de se tromper, de rectifier ses erreurs, d'encoder les connaissances, etc. Dans la deuxième étape, les chercheurs essayent d'intégrer les patrons observés dans les STI et de renforcer leurs capacités à fournir une interaction plus ciblée et un contenu plus adapté aux apprenants.

Généralement, les STI sont utilisés soit en remplacement des moyens traditionnels d'apprentissage (tel que l'enseignement classique en classe, ou à travers les livres), soit pour aider l'apprenant à s'exercer. Dans le premier cas, les apprenants découvrent le contenu de la matière pour la première fois à travers le système. Dans le deuxième cas, les apprenants connaissent déjà la matière et utilisent le système pour améliorer leurs niveaux de compréhension et leurs connaissances (Lane, 2006). Les recherches dans les STI penchent plus vers cette deuxième façon d'envisager et de concevoir les systèmes. Les STI sont davantage considérés comme des outils de support à l'apprentissage, présentant des activités et un contenu permettant aux apprenants d'améliorer, de développer et de tester leurs connaissances d'une manière intelligente et interactive.

Bien que les STI soient capables de prodiguer aux apprenants un enseignement de qualité et d'avoir un impact significatif sur leurs connaissances, ces systèmes peinent à rivaliser avec les tuteurs humains (Anderson, 1993). Selon Lane (2006), les tuteurs humains surclassent toutes autres formes d'enseignement. Il rapporte que le tuteur humain arrive à avoir une taille d'effet  $d=2$  sur le gain d'apprentissage, relativement à un enseignement en classe et sans tutorat.<sup>2</sup> En revanche, les meilleurs STI atteignent une taille d'effet  $d=1$  toujours par rapport à un enseignement en classe sans tutorat. La taille d'effet de l'EAO, quant à lui, ne dépasse pas  $d=0.4$ . Cette même constatation a été reprise par VanLehn (2011) dans sa synthèse de résultats de différentes études comparatives établies entre l'enseignement humain, les STI et l'EAO. La figure 2.4, traduite du rapport de VanLehn (2011), récapitule les différentes tailles d'effet des formes d'enseignement sur le gain d'apprentissage.

La supériorité des tuteurs humains relativement aux STI est également apparente dans leurs capacités à dépasser la dimension cognitive dans leurs interactions avec l'apprenant pour

---

<sup>2</sup> La taille d'effet est une mesure du degré de relation entre deux variables. Le résultat  $d$  du test de Cohen est une mesure commune de la taille d'effet en unité de déviation standard. Dans les recherches en éducation, une valeur de  $d$  égale à 1 sigma correspond approximativement à une amélioration d'un grade littéral. Pour deux échantillons de moyennes respectives  $M_1$  et  $M_2$ , et de déviations standard  $S_1$  et  $S_2$ , la taille d'effet  $d = (M_1 - M_2) / \sqrt{(S_1^2 + S_2^2) / 2}$  (D'Mello et al, 2012).

prendre en compte des considérations d'ordre émotionnelles et affectives. En effet, les humains disposent d'une faculté naturelle qui leur permet de décoder, juger et exploiter les réactions émotionnelles d'un apprenant pour avoir une vision plus objective du processus d'apprentissage et adapter intelligemment leurs enseignements (Graham et al., 1996; Zimmerman, 2000). Les premières générations de STI ne disposent pas de ce *haut niveau* d'interaction. Leurs fonctionnements reposaient principalement sur des modèles et des mécanismes de simulation du raisonnement humain, qui faisait d'eux des moteurs d'analyse de la logique des apprenants (Picard et al., 2004). Cette conception des STI limitait considérablement l'efficacité de leurs interactions avec les apprenants dans la mesure où l'apprentissage humain est un processus complexe qui ne peut pas être réduit à un simple traitement cognitif de l'information. Dans son analyse de l'apprentissage humain, Phelps (2006) affirme que « l'examen des fonctions cognitives sans considération du contexte social, émotionnel et motivationnel résultera en une compréhension limitée ».

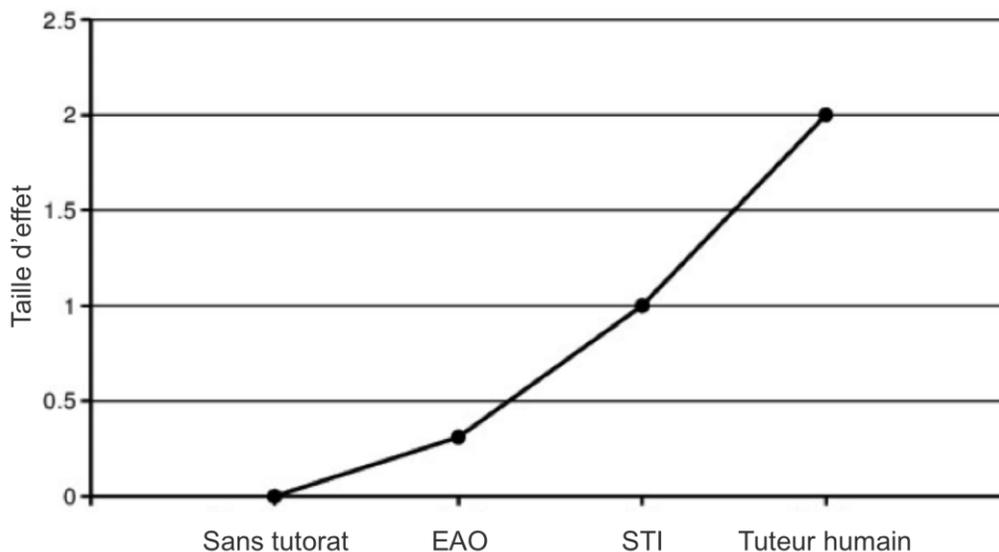


Figure 2.4 – Courbe de taille d'effet des différents moyens de tutorat

Les recherches dans le domaine des STI ont de plus en plus pris conscience de la limite des approches exclusivement cognitives et du rôle fondamental que peuvent jouer ces facteurs émotionnels dans le processus d'apprentissage (Picard, R. W., 1997).

## 2.5 Intégration de la dimension émotionnelle dans les STI

### 2.5.1 Les émotions et l'apprentissage

Longtemps négligées dans le domaine de l'apprentissage, les émotions ont intégré les STI à la suite des récents travaux de recherche en neurosciences, psychologie et sciences cognitives, qui ont mis en évidence la relation *étroite* et *complexe* qui existe entre la cognition et les émotions (Ashby et al., 1999; Cytowic, 1989; Mandler, 1984). En effet, les émotions sont omniprésentes dans le comportement d'un individu et exercent une influence fondamentale sur ses fonctions cognitives : elles régissent sa perception, dominant son intuition et sa prise de décision; stimulent son inspiration et ses capacités créatives, etc. (Damasio, 1994; Isen et al., 1987; Pekrun, 1992; Picard, 1997).

Les mécanismes de mémorisation, par exemple, associent le plus souvent l'état émotionnel de l'individu à l'information à retenir (Pekrun, 1992). Cette association influence la qualité de rétention de l'information dans la mesure où les émotions positives permettent de mémoriser les éléments plus longtemps qu'un état neutre. De plus, certaines émotions comme l'anxiété ou le stress intense peuvent carrément bloquer le processus de mémorisation (Goleman, 1996). Ces mêmes émotions peuvent également, gêner ou monopoliser l'attention d'un individu. Les apprenants deviennent alors incapables de canaliser leur concentration sur les tâches d'apprentissage et vont plutôt se focaliser sur la cause de leurs émotions ce qui peut perturber leur concentration (Pekrun, 1992).

Les stratégies de réflexion et de traitement de l'information sont elles aussi affectées par les émotions. Les émotions positives permettent une stratégie moins systématique et plus heuristique dans le traitement de l'information. Plusieurs règles de simplification sont alors sollicitées ce qui améliore considérablement le raisonnement de l'individu (Isen, 2000). En revanche, les émotions négatives engagent un processus beaucoup plus systématique dans lequel la personne compte peu sur ses connaissances préexistantes et se focalisent davantage sur les détails (Fiedler, 1991, 2001; Pekrun, 1992). Ce constat sur la relation étroite entre les émotions et la réflexion est également partagé par Marvin Minsky (2007) qui affirme que « Lorsque nous changeons d'état émotionnel, nous changeons notre façon de réfléchir ».

Certaines émotions sont plus favorables à certains types de traitements cognitifs que d'autres, mais les émotions positives ne sont pas inéluctablement les meilleures. En effet, la présence d'émotions négatives comme la confusion (dans le cas où un individu éprouve un blocage ou des difficultés à comprendre des concepts) peut augmenter la probabilité de l'occurrence d'un apprentissage profond dans lequel l'individu finira par apprendre et maîtriser parfaitement ces concepts (Burlison et al., 2004; D'Mello et al., 2014). En revanche, l'occurrence d'émotions positives lorsqu'un individu est dans une 'impasse cognitive' et n'arrive pas à comprendre certains concepts peut être le signe d'un manque d'intérêt ou d'implication par rapport à la situation actuelle.

Ainsi, il est très fréquent de parler de conditions émotionnelles favorables à certaines situations plus que d'autres. Dans de bonnes conditions émotionnelles, les apprenants expriment un plus grand intérêt et une meilleure appréciation des activités d'apprentissage ainsi qu'une meilleure persévérance dans la réalisation de leurs objectifs même en cas d'échecs. L'influence négative, quant à elle, produit l'effet inverse avec un manque d'implication par rapport à la réalisation des tâches, un découragement rapide face aux difficultés et une tendance à décrocher rapidement (Pekrun, 1992; Picard, 1997).

## **2.5.2 L'intelligence émotionnelle**

Comme nous l'avons précédemment mentionné, le manque de considération de la composante émotionnelle fondamentale pour l'apprentissage humain a pour longtemps fait défaut au fonctionnement des STI. En effet, les STI ne disposaient pas de cette faculté naturelle que les tuteurs humains utilisent afin de percevoir et de s'adapter en fonction des états émotionnels des apprenants. Par exemple, un tuteur qui pose une question à un apprenant peut spontanément apercevoir de la confusion ou de l'incertitude chez ce dernier même si la réponse fournie est correcte. Ces différentes constatations d'ordre émotionnelles peuvent ainsi influencer la suite de la stratégie d'enseignement. L'aptitude à déchiffrer les émotions, déterminer leurs causes, estimer leurs répercussions et évaluer comment les influencer (ou les réguler) constitue une forme d'intelligence, appelée *intelligence émotionnelle* (Salovey et al., 1990). Cette faculté — qui varie d'un individu à un autre et qui se développe avec le temps et

l'expérience — peut même être un facteur déterminant dans le succès d'un individu sur le plan personnel et professionnel (Goleman, 1995).

Le concept d'intelligence émotionnelle a soulevé un grand intérêt dans la communauté des STI vu le potentiel que peut avoir le développement de cette forme d'intelligence dans ces systèmes. De plus, l'avènement d'un nouveau courant de recherche connu sous le nom de *l'informatique affective* (Picard, 1997) a permis l'apparition d'une nouvelle génération de systèmes tutoriels dont les mécanismes d'adaptations sont davantage enrichis avec des heuristiques et des considérations émotionnelles. Cette tendance a complètement changé la relation de ces systèmes avec l'apprenant. On parle de Systèmes Tutoriels Émotionnellement Intelligents (STEI) (Ochs et al., 2004), de systèmes tutoriels affectifs (Alexander et al., 2003), ou encore de STI sensibles à l'affect (D'Mello et al., 2005).

### **2.5.3 L'informatique affective et les Systèmes Tutoriels Émotionnellement Intelligents**

L'informatique affective est un courant de recherche dérivé de l'Interaction Homme-Machine (IHM) qui a permis de mettre en place les fondements requis pour faire valoir la dimension émotionnelle dans les systèmes interactifs actuels. Introduit par Picard en 1997, ce domaine de recherche vise à améliorer le fonctionnement des systèmes en les rendant davantage conscients et proactifs par rapport à l'état émotionnel de leurs utilisateurs. Trois principaux thèmes de recherche peuvent être distingués en informatique affective (Picard, 1997) :

1. La reconnaissance des émotions : il s'agit d'étudier, de développer et d'améliorer les techniques et les dispositifs matériels permettant aux machines d'extraire des indices sur l'état émotionnel de l'utilisateur. Cet aspect fondamental en informatique affective vise donc à renforcer les capacités des machines à identifier et évaluer les émotions des utilisateurs et ainsi de créer un canal de communication émotionnelle de *l'utilisateur vers la machine*.
2. L'expression des émotions : ce thème s'intéresse à modéliser les processus émotionnels dans des systèmes informatiques. Il s'agit de permettre aux machines de déclencher

naturellement et d'une manière appropriée des réactions émotionnelles au cours de leurs interactions avec les utilisateurs. Ces émotions sont généralement transmises aux humains par des robots sociaux ou des avatars graphiques conçus pour avoir un comportement émotionnel crédible et cohérent ce qui permet, par conséquent, d'établir un canal de communication émotionnelle de *la machine vers l'utilisateur*.

3. La prise en compte des émotions : ce thème aborde les questions relatives à l'intégration de la dimension émotionnelle dans le fonctionnement des machines. L'objectif est de trouver les meilleures stratégies permettant aux systèmes d'interagir intelligemment en tenant compte du facteur émotionnel afin d'améliorer la satisfaction, l'engagement et les performances des utilisateurs.

Les avancées réalisées durant ces dernières années dans le domaine de l'informatique affective ont marqué un changement significatif dans la relation homme-machine dans plusieurs domaines, en particulier celui de l'apprentissage. L'intégration des principes de l'informatique affective a en effet permis l'émergence d'une nouvelle génération de systèmes d'apprentissage connus sous le nom de Systèmes Tutoriels Émotionnellement Intelligents (STEI) (Ochs et al., 2004; Sarrafzadeh et al., 2004). Ces systèmes sont définis comme étant des STI capables d'analyser le comportement émotionnel des apprenants et d'adapter leurs stratégies d'enseignement en conséquence. Les STEI se distinguent donc par leurs capacités à mettre en avant la dimension émotionnelle dans leurs stratégies d'enseignement et à dépasser le cadre strict d'analyse de raisonnement qui caractérise le fonctionnement classique d'un STI.

La logique d'interaction des STEI est principalement basée sur des stratégies de régulation émotionnelle dont l'objectif principal est de maintenir ou d'amener l'apprenant vers un état émotionnel favorable à l'apprentissage. En effet, cette approche, souvent adoptée intuitivement par les tuteurs humains, cherche à augmenter les performances d'apprentissage en jouant sur l'influence que peuvent exercer les émotions sur les capacités cognitives d'un individu (Ochs et al., 2004).

Généralement, les STEI utilisent des techniques issues de l'informatique affective pour reconnaître automatiquement les états émotionnels des apprenants au cours de leur

apprentissage. Cette étape est fondamentale pour le fonctionnement d'un STEI dans la mesure où toute la suite de la stratégie d'adaptation du système est basée sur la précision et la fiabilité de cette procédure de reconnaissance émotionnelle. Dans la deuxième étape d'adaptation, le système doit disposer d'une stratégie émotionnelle qui lui permet de déterminer quelle est la meilleure *intervention* à initier.

Nous décrivons dans ce qui suit les principaux moyens de reconnaissance émotionnelle ainsi que les principales stratégies d'adaptation employées par les chercheurs en STEI.

## **2.6 Représentation des émotions**

Afin de reconnaître et s'adapter aux émotions il est important de pouvoir les définir et les modéliser. Les recherches dans les STEI ont été largement influencées par les théories psychologiques dans la conception et la modélisation du phénomène émotionnel. Nous présentons dans les deux prochaines sections les principales approches de conceptualisation des émotions ainsi que les principaux moyens de reconnaissance.

### **2.6.1 Le phénomène émotionnel**

Bien que les émotions soient connues de tous, il est difficile de trouver un consensus sur leurs définitions ou leurs origines. Kleinginna et al. (1981) en compte plus d'une centaine de définitions. De Descartes à Darwin en passant par Spinoza ou Hume, philosophes, psychologues et scientifiques se sont toujours penchés sur ce phénomène complexe et mystérieux. En effet, les émotions déclenchent chez l'individu plusieurs processus tels qu'une expression corporelle et faciale, des sensations, une évaluation, une prédisposition à l'action, une adaptation, ce qui complique leur définition (Frijda, 1986). Dans le domaine de l'Interaction Homme-Machine et des STI en particulier, la revue de la littérature scientifique montre que les termes émotion, affect sont souvent employés d'une manière interchangeable ce qui peut porter à la confusion et rend parfois difficile la comparaison des méthodes et résultats de recherches (Zimmermann et al., 2003). Pour clarifier la différence entre ces deux concepts, nous faisons la distinction suivante :

- **Émotion.** Une émotion est généralement envisagée comme étant un processus de réaction qu'un organisme produit suite à l'interprétation d'un évènement significatif (Scherer, 2000). La nature de la signification de cet évènement (par exemple l'apparition d'une menace) caractérise alors la nature de l'émotion. Le degré de signification de l'évènement tel qu'il est perçu par l'individu détermine l'ampleur ou l'intensité de l'émotion.
- **Affect.** L'affect est un concept plus global et plus abstrait que les émotions. Il se rapporte plutôt à l'expérience inconsciente et non contrôlable des émotions (Shouse, 2005). L'affect peut avoir une grande influence sur le comportement, l'attitude et les choix d'un individu à long terme (Petty et al., 1988). Par exemple, une personne peut décider de voter pour un candidat politique, car il a été inconsciemment touché sur le plan affectif par la vidéo de sa campagne électorale qui met en avant l'importance de la famille. Ainsi, l'individu peut être inconscient de son affect contrairement à ses émotions qu'il peut percevoir et parfois contrôler.

Le phénomène émotionnel est donc une expérience intense d'une courte durée (quelques secondes ou minutes) et dont l'individu en est conscient (Zimmermann et al., 2003).

D'un point de vue neuropsychologique, trois principales régions du cerveau sont impliquées dans le traitement des émotions : le thalamus, le système limbique et le cortex. Les stimulus émotionnels sont reçus par le thalamus qui les relaie en parallèle vers le système limbique (voie émotionnelle courte et directe) ainsi que vers le cortex (voie émotionnelle longue) pour qu'ils reçoivent un traitement de plus haut niveau (LeDoux, 1999). La voie courte entre le thalamus et le système limbique correspond au déclenchement des *émotions primaires* comme la peur. Celles-ci provoquent chez l'individu une réaction impulsive, innée ou acquise et quasi instantanée. Le processus de réaction est activé avant de faire une évaluation cognitive du stimulus ou de l'évènement. En revanche, les *émotions secondaires*, telles que la frustration, résultent de l'acheminement des informations vers le système limbique, après qu'elles ont été traitées par le cortex. Les émotions secondaires sont plus complexes, mais plus rationnelles, car une évaluation cognitive de la situation est établie (Damasio, 1994).

La majorité des études sur les émotions dans le contexte de l'Interaction Homme-Machine s'intéressent principalement aux émotions secondaires. L'intégration de la dimension émotionnelle dans les systèmes informatiques se base sur des techniques de modélisation largement influencées par les approches de classification émotionnelle théoriques existantes en psychologie. Nous présentons dans la sous-section suivante les principales approches.

## **2.6.2 Approches de modélisation des émotions**

Quelle que soit la théorie utilisée pour définir le concept d'émotion, on distingue généralement deux principales approches pour modéliser les émotions dans les machines à savoir l'approche catégorielle (ou discrète) et l'approche dimensionnelle (ou continue).

### **3.6.2.1 L'approche discrète ou catégorielle**

L'approche discrète est fondée sur le principe que le processus émotionnel peut être réduit à un nombre restreint d'émotions appelées *émotions de bases*. Ces émotions sont considérées comme innées et universelles ; elles sont donc partagées et exprimées de la même façon par tous les humains (Ekman, 1982, 1999; Ortony, 1990; Panksepp, 1992; Tomkins, 1984). Les autres émotions, plus complexes, sont alors vues comme des dérivés ou des combinaisons de ces émotions de base (Plutchik, 1980). Toutefois, il n'existe pas de consensus quant à la nature ni quant au nombre de ces émotions de base, qui peut varier de deux à vingt-deux émotions selon les différentes théories.

Parmi les modèles de l'approche discrète, on trouve celui de Plutchik (1980) qui propose un ensemble de huit émotions de base regroupées par paires opposées : joie-tristesse, acceptation-dégout, peur-colère et surprise-anticipation. Par analogie avec les nuances de couleurs, les émotions plus complexes seraient donc des dérivées ou des nuances d'une ou de plusieurs émotions de base. Par exemple, une émotion de base comme la tristesse, accompagnée d'une autre émotion de base comme la surprise, peut dériver une émotion plus complexe : la déception. Celle-ci peut dériver à son tour une autre émotion complexe qui est le mécontentement et ainsi de suite (Plutchik, 1980).

Ekman quant à lui propose une classe contenant six émotions de base qui sont la joie, la surprise, la colère, la peur et le dégoût en se basant sur les analyses des expressions faciales (Ekman, 1982 ; Ekman, 2004; Ekman et al., 1978 ; 2003 ). Selon cette théorie, chacune de ces émotions de base déclenche chez les individus un prototype d'expression faciale inné et universel. Les travaux entrepris par Paul Ekman ont permis la mise en place d'un descriptif complet des mouvements faciaux musculaires, appelé « Facial Action Coding System » (FACS), permettant de définir et d'identifier ces émotions de base. Par exemple, le dégoût se caractérise par un plissement de la peau vers le haut, un abaissement des coins externes des lèvres et une ouverture de la lèvre inférieure.

Le modèle OCC proposé par Ortony Clore, et Collins (1988) définit vingt-deux émotions de base avec leurs conditions de déclenchement (Ortony et al., 1988). Ces émotions sont regroupées en classes appelées également « clusters ». Chaque classe contient les émotions susceptibles d'être déclenchées par le même type d'évènement. Par exemple, la classe « détresse », dont la condition de déclenchement est l'occurrence d'un évènement non désirable, regroupe des émotions telles que la tristesse et le bouleversement.

### **3.6.2.2 L'approche dimensionnelle ou continue**

Par opposition à l'approche discrète, l'approche dimensionnelle ne se limite pas à un nombre fini d'émotions de base. Selon cette approche, il existe un nombre infini d'émotions qui peuvent être représentées dans un espace affectif constitué par un ensemble de dimensions élémentaires et quantifiables (Barrett et al., 1998). Les dimensions les plus couramment utilisées pour modéliser les émotions selon l'approche dimensionnelle sont : la *valence* (appelée également *plaisir*) et *l'intensité* (appelée également *éveil* ou *activation* « *arousal* »).

La dimension de valence caractérise la disposition d'un organisme à émettre un comportement d'évitement ou d'approche par rapport à l'évènement déclencheur de l'émotion. La valence d'une émotion peut varier du *positif* (agréable) au *néгатif* (désagréable). La dimension de l'intensité correspond au degré d'activation corporelle provoquée par le déclenchement de l'émotion. Cette dimension permet de définir l'émotion en fonction du

niveau de réaction de l'organisme qui peut aller d'une réaction de *faible* intensité (désactivation) à une réaction de *forte* intensité (activation).

En plus de ces deux dimensions qualifiées de *primaires*, certains théoriciens ajoutent également la dimension de *dominance* (Mehrabian, 1996). Cette dimension décrit la capacité de l'organisme à avoir un contrôle ou une influence sur la situation émotionnelle. La dominance d'une émotion peut varier du non-contrôle jusqu'au contrôle ou la maîtrise complète (« coping »). D'autres dimensions telles que la *fréquence* de l'émotion (échelle allant de rare à fréquente), ou sa *durée* (échelle allant de courte à longue) existent, mais elles sont très rarement utilisées dans la modélisation émotionnelle.

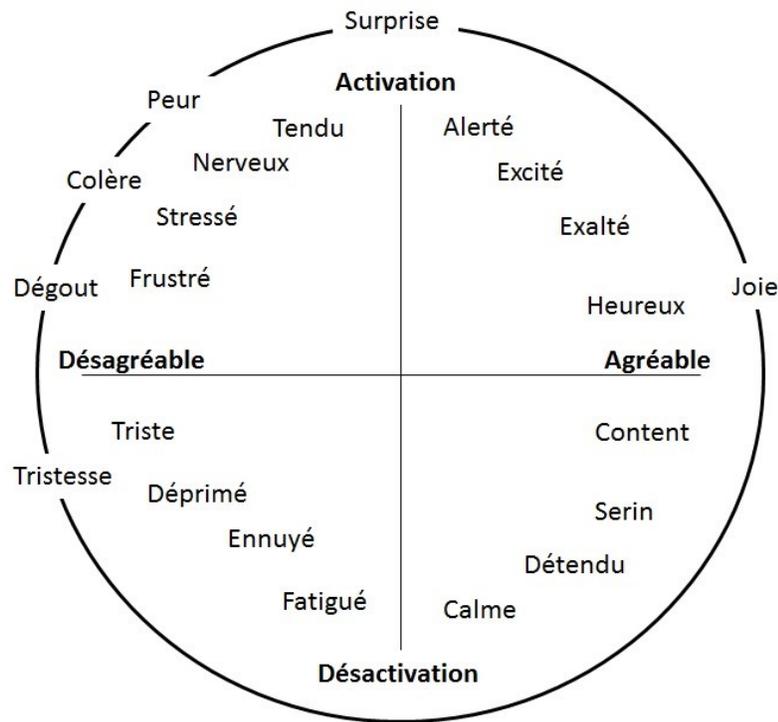


Figure 2.5 – Modèle « circumplex » de Russel, tirée de Feldman-Barrett & Russell (1998).

Parmi les modèles les plus utilisés de l'approche dimensionnelle, on peut citer le modèle « circumplex » de Russel (1980) (Barrett et al., 1998; Russell, 1980). Ce modèle également repris et validé dans les travaux de Lang (Lang, 1995) propose de représenter les émotions dans une structure circulaire bidimensionnelle dans laquelle se croisent les

dimensions de valence et d'intensité de manière à former quatre oppositions bipolaires. La figure 2.5 illustre ce modèle. Chaque émotion est située en fonction de sa description selon la dimension de valence (axe horizontal du modèle) et d'intensité (axe vertical du modèle).

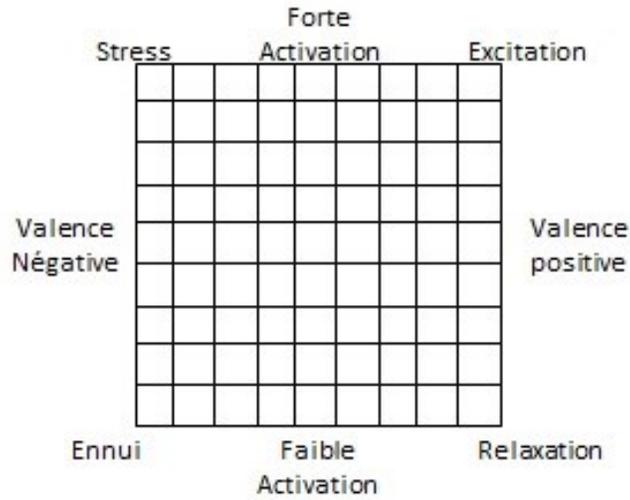
## **2.7 Moyens de détection des émotions dans les STEI**

Les moyens de détection des émotions dans les STEI peuvent être classés en trois principales catégories : l'auto-évaluation, les mesures comportementales et les mesures physiologiques. Dans l'auto-évaluation, l'apprenant est amené à décrire lui-même ses émotions. Les deux dernières mesures se font à partir d'un ensemble d'indices corporels. En effet, les émotions peuvent être révélées à travers un ensemble d'indices physiques apparents et observables (mesures comportementales) tels que les expressions faciales, la tonalité de la voix, les gestes, etc.), mais également à travers des indices qui ne sont pas directement observables (mesures physiologiques) tels que la fréquence cardiaque, le volume de la pression sanguine, la fréquence de la respiration, etc. (Picard, 1997).

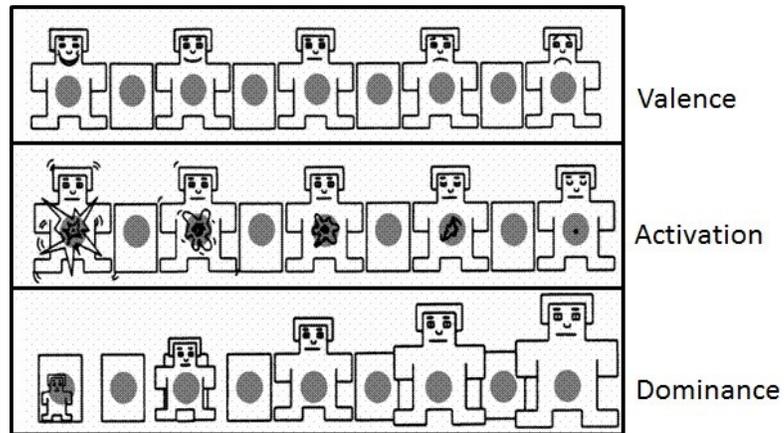
Nous présentons dans ce qui suit les principaux moyens de mesure dans chacune de ces trois catégories.

### **2.7.1 L'auto-évaluation**

Appelée également moyen subjectif de mesure émotionnelle, le principe de cette catégorie de mesure repose sur l'utilisation de questionnaires d'auto-évaluation (« self-report ») qui vont permettre à l'apprenant de faire état lui-même de ses émotions. Dans les STEI, les instruments d'auto-évaluation utilisés peuvent varier de la simple question à choix multiples aux questionnaires psychologiques complets. Ces derniers contiennent généralement plusieurs questions avec un protocole d'évaluation plus précis ce qui permet à l'apprenant de mieux juger ses propres émotions.



[a]



[b]

Figure 2.6 – Exemples d’instruments d’auto-évaluation selon l’approche dimensionnelle des émotions : [a] grille d’affect de Russel (1986), [b] l’échelle « Self-Assessment Manikin scale ».

Certains instruments d’auto-évaluations sont conçus pour modéliser les émotions selon l’approche discrète. On trouve alors un ensemble d’émotions cibles qui sont visées par un questionnaire et en fonction des réponses apportées par l’apprenant, une émotion finale sera déterminée. On peut citer comme exemple l’échelle des émotions différentielles (« Differential emotions scale ») (Izard, 1993) qui permet de différencier entre dix émotions ou encore le

questionnaire d'émotions académiques (« Academic Emotions Questionnaire ») (Pekrun et al., 2011) qui permet de discriminer entre huit émotions.

D'autres instruments d'auto-évaluation opèrent selon l'approche dimensionnelle des émotions. L'évaluation se fait alors en fonction des dimensions émotionnelles — principalement la valence et l'activation — et ne vise pas une émotion en particulier. Parmi les moyens d'auto-évaluation dimensionnelle les plus simples et les plus utilisés, on peut citer la technique d'évaluation bipolaire des dimensions émotionnelles, dans laquelle l'apprenant décide uniquement si ses émotions sont de nature positive ou négative (pour la valence) et si elles ont une forte ou faible intensité (pour l'activation) (Mehrabian et al., 1974). On peut également citer la grille d'affect (« affect grid ») de Russel (Russell et al., 1989) qui se base sur le même principe d'évaluation dimensionnelle et qui utilise pour cela une grille 9x9 représentant l'espace affectif bidimensionnel formé par la valence et l'activation (voir figure 2.6 [a]). L'apprenant est alors amené à mettre une croix dans la zone qui traduit au mieux ses émotions et plus cette croix se rapproche du centre de la grille plus les émotions de l'apprenant sont neutres. Bradley et Lang (Bradley et al., 1994) proposent eux aussi d'utiliser le même principe d'évaluation, mais en remplaçant la grille d'affect par une échelle graphique contenant des figurines et en ajoutant, en plus de la dimension d'activation et de valence, la dimension de dominance. Cette échelle appelée « Self-Assessment Manikin scale » utilise neuf degrés pour chaque dimension, dont cinq figurines principales et quatre états intermédiaires (voir figure 2.6 [b]). L'apprenant devra alors choisir pour chaque dimension une figurine dont les traits ou le format se rapprochent le mieux à son état émotionnel. La combinaison des trois choix représente alors l'état émotionnel final.

Les instruments d'auto-évaluation sont très fréquemment utilisés dans les STEI (Baker et al., 2010 ; Beal et al., 2005 ; D'Mello et al., 2010 ; de Vicente et al., 2002 ; Graesser et al., 2007). La popularité de ces techniques provient principalement de la simplicité de leur principe, la rapidité avec laquelle elles peuvent être administrées et la facilité avec laquelle elles peuvent être intégrées dans les systèmes. Cependant, ces techniques présentent également divers inconvénients. En effet, la procédure d'auto-évaluation peut perturber la dynamique de la session d'apprentissage, car elle nécessite que l'apprenant interrompe ses activités pour

répondre au questionnaire (Picard et al., 2005). Si le système procède à une évaluation fréquente des émotions, cela peut devenir fastidieux pour l'apprenant et affecter sa concentration. En revanche, si cette évaluation est rare, le système risque de ne pas être à jour par rapport à l'état émotionnel de l'apprenant. De plus, certaines personnes peuvent être gênées à l'idée de révéler directement leurs émotions (Brody, 1999). D'autres éprouvent certaines difficultés à comprendre ce qu'ils ressentent exactement, ou encore, se préoccupent du fait que le système percevra leurs états, ce qui pourrait biaiser leurs évaluations (Picard et al., 2004).

Dans les études sur les émotions, les instruments d'auto-évaluation sont généralement combinés avec d'autres moyens comportementaux ou physiologiques de détection émotionnelle. Certains chercheurs ne recommandent pas de baser les STEI uniquement sur le principe de l'auto-évaluation (de Vicente et al., 2002). Toutefois, les valeurs issues de l'auto-évaluation sont souvent considérées comme des valeurs de référence « ground truth » pour les autres moyens de détection.

### **2.7.2 Les mesures comportementales**

Contrairement à l'auto-évaluation, cette catégorie de mesures identifie les émotions des apprenants d'une manière indirecte, et ce à travers l'analyse d'indices émotionnels contenus dans les expressions faciales, la posture, la voix ou encore la façon avec laquelle les apprenants interagissent avec le STEI.

Le recours à la caméra pour identifier les émotions est une pratique très courante dans cette catégorie, vu que la majorité des machines en sont équipées (El Kaliouby et al., 2005; McDaniel et al., 2007; Nkambou et al., 2004; Whitehill et al., 2008). Ces systèmes se basent sur une codification des contractions musculaires du visage sous la forme d'unités d'action appelées en anglais « Action Unit » (AU) où chaque unité définit le mouvement spécifique d'un ou de plusieurs muscles précis. Parmi les systèmes de codification les plus connus et les plus utilisés dans les STEI, on peut citer le codage « Facial Action Coding System » (FACS) d'Ekman (1978) qui utilise 46 unités d'action avec leurs descriptifs détaillés (par exemple, AU 10 correspond à une remontée de la partie supérieure de la lèvre). Le système ajoute également

une lettre parmi 6 lettres possibles (de A jusqu'à E) pour décrire l'intensité de l'action musculaire (la lettre B par exemple correspond à un mouvement léger de l'unité d'action). Le déclenchement de certaines unités d'action est alors révélateur de l'apparition d'une émotion spécifique. McDaniel et ses collègues (2007) ont pu établir une corrélation entre l'apparition de la joie et la confusion chez les apprenants lors de la session d'apprentissage et certaines unités d'action. Par exemple l'abaissement et le rapprochement des sourcils (AU 4), l'existence d'une tension au niveau des paupières (AU 7) et une absence d'étirement au niveau du coin des lèvres (AU 12) peut révéler la présence d'une certaine confusion chez l'apprenant (voir figure 2.7).



Figure 2.7 – Exemple de codage FACS d'une expression faciale qui exprime une légère confusion. L'image est tirée de l'expérience de McDaniel et ses collègues (2007).

Le fonctionnement de ces STEI qui utilisent un système de codage repose généralement sur l'utilisation d'un ensemble d'algorithmes qui vont dans un premier temps localiser le visage de l'apprenant dans l'image. Ensuite, les différents éléments du visage (yeux, bouche, sourcils, etc. sont détectés) et un ensemble de points spécifiques dans ces éléments sont analysés afin d'identifier les unités d'action. Finalement, les différentes unités d'action serviront comme des prédicteurs pour reconnaître certaines émotions. La figure 2.8 illustre ces étapes.

Bien que la reconnaissance des émotions à partir des expressions faciales soit une technique répandue, plusieurs faiblesses sont à signaler. Tout d'abord, seul un ensemble restreint d'émotions discrètes peut être détecté avec cette technique. En outre, certaines personnes sont peu expressives et parviennent à bien contrôler leurs réactions faciales. Donc

pour certaines personnes l'absence d'expression ne veut pas dire une absence d'émotion. De plus, certaines expressions faciales peuvent signifier l'inverse de ce qu'une personne ressent. Par exemple, un sourire détecté n'est pas automatiquement synonyme d'émotions positives et peut exprimer une certaine déception. De plus, malgré l'utilisation d'algorithmes de plus en plus sophistiqués, l'utilisation d'une simple caméra peine à détecter des émotions complexes, mais très récurrentes dans l'apprentissage comme l'ennui, l'intérêt, le stress ou la frustration (McDaniel et al., 2007; Picard et al., 2005). Certains chercheurs proposent d'utiliser des caméras infrarouges ou thermiques en vue de cibler des émotions spécifiques telles que le stress ou la frustration (Kapoor et al., 2001). Cependant, ces techniques restent limitées à une seule émotion et nécessitent d'être combinés avec d'autres moyens de détection émotionnels pour avoir de bonnes performances (Gunes et al., 2010).

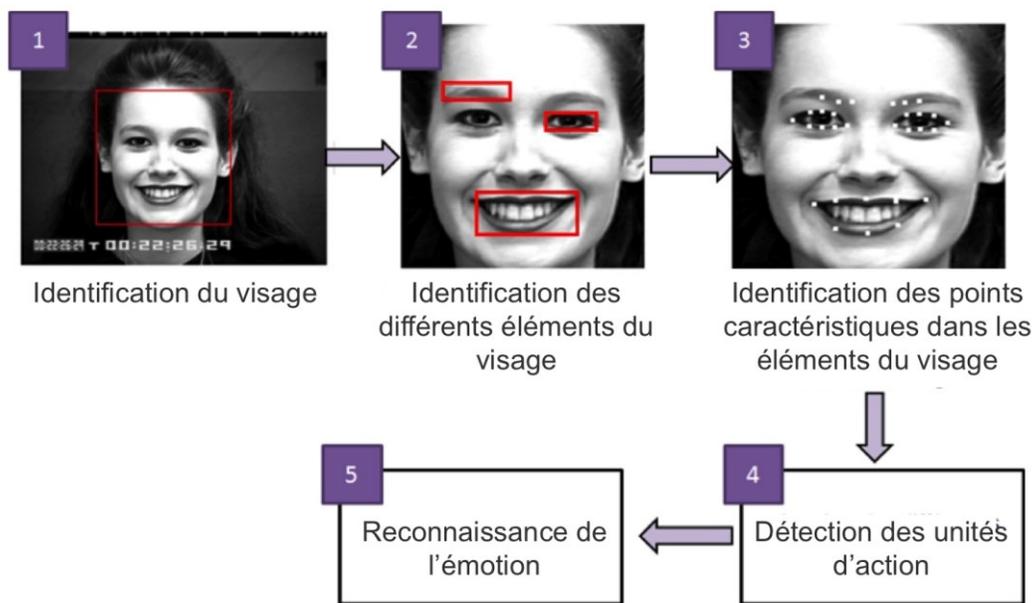


Figure 2.8 – Étapes de la reconnaissance des émotions à partir des expressions faciales

Tout comme les expressions faciales, la voix constitue un vecteur important d'informations affectives (Picard, 1997). Les émotions du locuteur peuvent être transmises d'une manière directe (linguistique) dans le message lui-même ou d'une manière indirecte (paralinguistique) dans la manière avec laquelle il est transmis. Il est généralement difficile d'extraire automatiquement et d'une manière précise les émotions d'un individu uniquement à

partir d'une analyse linguistique de son message (Gunes et al., 2010). En effet, le choix des mots ainsi que leurs teneurs affectives peuvent énormément varier d'une personne à une autre, mais aussi d'un contexte à un autre. L'extraction des émotions à partir de la voix se fait donc le plus souvent à travers une analyse paralinguistique (Ang et al., 2002; Ball et al., 2000; Litman et al., 2004, 2006). Des indices prosodiques et acoustiques (tels que la fréquence de vibration, l'intensité, le débit de parole, la durée de certains segments vocaux, la proportion des silences, etc.) sont analysés afin de renseigner sur les émotions de l'apprenant. Par exemple, une augmentation de l'intensité et de la fréquence de vibration de la voix d'un apprenant peut exprimer un certain énervement. Parmi les STEI qui se basent exclusivement sur l'analyse de la voix pour la détection émotionnelle on peut citer le système ITSPOKE réalisé par Litman et ses collègues (2004), qui utilisent 33 caractéristiques prosodiques et acoustiques pour distinguer entre trois états émotionnels à savoir : l'état positif, l'état négatif et l'état neutre.

L'analyse des gestes et de la posture constitue également un moyen de détection émotionnelle très utilisé dans les STEI. Dans une étude sur la reconnaissance des émotions, Coulson (2004) a établi que l'on peut reconnaître les six émotions discrètes (colère, dégoût, peur, tristesse et surprise) à partir de la posture d'une manière aussi précise qu'une reconnaissance à partir de la voix, et que certaines postures peuvent exprimer des émotions tout aussi efficacement que des expressions faciales (Coulson, 2004). La posture et les mouvements corporels présentent l'avantage d'être plus spontanés et moins contrôlables que les expressions faciales ou la voix (D'Mello et al., 2009). Un apprenant intéressé aura tendance à spontanément se rapprocher de l'écran. Se gratter la tête après une question peut exprimer une certaine réflexion et de la concentration. S'il commence à faire trembler ses genoux, cela peut être le signe d'un stress. Un haussement spontané des épaules peut signifier une incertitude. Un mouvement soudain en arrière peut exprimer une surprise. En revanche, s'il s'affaisse sur sa chaise et s'éloigne de l'écran, cela peut indiquer un ennui ou un désintérêt.

La première approche pour analyser de la posture de l'apprenant peut se faire en utilisant une caméra avec des algorithmes de suivi du positionnement et des mouvements des parties du corps (Asteriadis et al., 2009; Lu et al., 2005; Mitra et al., 2007). Asteriadis et ses collègues

ont pu, par exemple, suivre l'attention des apprenants lors de leurs interactions avec un environnement d'apprentissage de la lecture en analysant conjointement le positionnement des mains, le mouvement de la tête et la direction du regard (Asteriadis et al., 2009). La deuxième manière pour analyser la posture et les gestes des apprenants est une approche très populaire dans les STEI : elle consiste à utiliser une chaise équipée avec des détecteurs placés dans un coussin très mince (Arroyo et al., 2009; D'Mello et al., 2009; D'Mello et al., 2007; Mota et al., 2003). Ce coussin est placé sur l'assise et le dossier de la chaise (voir figure 2.9), et permet de donner une carte contenant les différentes pressions exercées par le corps ainsi que leurs directions (penchement du corps à droite, penchement en avant, jambes croisées, etc.).



Figure 2.9 – Exemple de chaise utilisée pour la détection émotionnelle

L'analyse de la posture dans STEI représente une alternative prometteuse qui peut rivaliser avec la caméra et le microphone pour la détection émotionnelle. D'Mello (2009) recommande l'utilisation de l'analyse de la posture pour la détection des émotions qu'il est difficile de reconnaître par une simple caméra comme l'état d'ennui par exemple.

Certains travaux proposent d'utiliser des systèmes de détection et de suivi des mouvements des yeux pour détecter les états émotionnels (Conati et al., 2007; D'Mello et al., 2012b; Wang et al., 2006). D'autres ont recours à une souris équipée avec des détecteurs de pression pour évaluer le niveau de frustration de l'apprenant (Qi et al., 2001; Reynolds, 2001). Chaffar et Frasson (2004) proposent d'utiliser l'électromyographie (EMG) qui consiste à mesurer l'activité électrique des muscles afin de suivre la nature des réactions émotionnelles

(valence positive ou valence négative). L'électromyographie a été également proposée par les chercheurs comme une alternative à l'utilisation de la caméra pour la reconnaissance des émotions à partir des expressions faciales. L'idée est de déterminer directement les actions des muscles du visage (par exemple les zygomatiques, les sourcils, etc.) en y plaçant des capteurs (appelés électrodes) qui mesurent l'activité musculaire (Conati, 2002; Dimberg et al., 2000).

Le point commun entre ces techniques précédemment citées c'est qu'elles se basent sur des modèles capables d'inférer les émotions à partir des données collectées par des dispositifs tels que la caméra, le microphone, la chaise, etc. Une autre approche pour la détection des émotions à partir des mesures comportementales consiste à se baser principalement sur l'analyse des variables d'interactions enregistrées par le STEI. Ces variables peuvent englober des indicateurs tels que : le temps émis pour répondre, le temps passé sur les activités, le nombre de fois où l'apprenant sollicite l'aide fournie par le système, la séquence de bonnes réponses ou de mauvaises réponses, le nombre de clics et le mouvement de la souris, etc. Des modèles d'apprentissage machine se basant sur ces variables sont souvent utilisés pour déterminer l'état de l'apprenant. Par exemple, Baker et ses collègues (2004, 2006) utilisent un ensemble de variables d'interaction (fréquence de clics sur le bouton aide, temps passé sur les activités, etc.) pour vérifier si l'apprenant est en train d'utiliser le système sérieusement et n'est pas en train de le faire déjouer (« gaming the system »). Le modèle proposé permet de détecter si l'apprenant n'est pas intéressé à apprendre et s'il est tout simplement en train de passer rapidement au travers le curriculum sans vraiment prêter attention et faire l'effort nécessaire. Beck (2005) analyse le temps de réponse pour savoir si l'apprenant n'est pas réellement impliqué dans la résolution de l'exercice et s'amuse à essayer de deviner les réponses (« guessing »).

Bien que les variables d'interaction peuvent renseigner sur certains aspects de l'état émotionnel de l'apprenant, leur utilisation est surtout dédiée à l'évaluation et au suivi des performances des apprenants dans les différentes tâches exécutées.

### 2.7.3 Les mesures physiologiques

La majeure partie des travaux portant sur la définition et la modélisation des émotions s'accorde sur le fait que l'occurrence des émotions s'accompagne avec des changements *physiologiques* (Scherer, 2005). De plus, contrairement aux manifestations comportementales (telles que les expressions faciales, la gestuelle ou la voix), les individus n'ont pratiquement aucun contrôle sur les changements physiologiques qui se produisent spontanément avec le déclenchement d'une émotion (Cacioppo et al., 2000). Pour ces raisons, l'approche physiologique pour la détection émotionnelle représente une solution prometteuse en informatique affective, et a donné naissance à un nouveau domaine de recherche connu sous le nom de *l'informatique physiologique* (« physiological computing ») (Fairclough, 2009).

Un des objectifs de l'informatique physiologique est de créer des systèmes capables d'extraire des informations sur l'état interne de l'utilisateur en analysant des données physiologiques collectées à partir d'un ou plusieurs senseurs physiques (Allanson et al., 2004). Ces données incluent la fréquence respiratoire ou cardiaque, l'activité cérébrale, le diamètre de la pupille, la pression sanguine, etc.

Il existe principalement deux façons d'utiliser les données physiologiques pour la détection émotionnelle. La première façon consiste à extraire directement, à partir des senseurs physiologiques, un ou plusieurs indices physiologiques qui sont soit corrélés avec des dimensions émotionnelles (principalement valence et activation) (Lang, 1995), soit associés directement à l'apparition de certaines classes discrètes d'émotions (frustration, stress, anxiété, etc.) (Chanel et al., 2007; Zhai et al., 2006). La deuxième façon consiste à utiliser ces données physiologiques comme des variables d'entrée dans des algorithmes d'apprentissage machine pour la classification et la prédiction des états émotionnels (Conati, 2002; Hussain et al., 2011; Kapoor et al., 2007; Mcquiggan et al., 2007b; Jraidi, 2014).

Parmi les indices physiologiques les plus utilisés dans la détection émotionnelle, on peut citer la réponse galvanique de la peau (« Galvanic Skin Response », GSR) qui traduit l'activité électrodermale. Cet indice est généralement mesuré en positionnant sur deux doigts de la main deux capteurs dédiés qui mesurent la différence de potentiel entre les deux zones cutanées.

Afin de ne pas mobiliser toute la main de l'individu pour mesurer la réponse galvanique de la peau, de nouveaux capteurs existent sous la forme d'un bracelet ou d'un gant, et sont souvent utilisés dans les STEI (voir figure 2.10 [a] et [b]). La réponse galvanique de la peau est fortement corrélée avec l'activation émotionnelle (Lang, 1995) : plus l'intensité de l'émotion est forte, plus la réponse la conductance est élevée et vice versa.

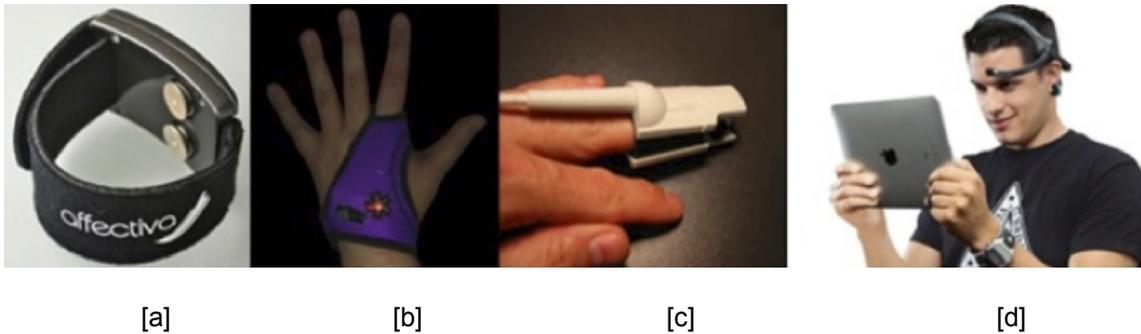


Figure 2.10 – Exemple de senseurs physiologiques. [a] et [b] représentent respectivement un bracelet et un gant utilisés pour mesurer la réponse galvanique de la peau, [c] est un capteur de pression du volume sanguin et [d] représente le système EEG de NeuroSky.

Un autre indice physiologique qui est directement relié avec une dimension émotionnelle est la fréquence cardiaque (« Heart Rate », HR). En effet, cet indice est corrélé avec la valence émotionnelle, et il est établi qu'une émotion positive provoque une phase d'accélération cardiaque plus importante qu'une émotion négative (Lang, 1995). La fréquence cardiaque peut être mesurée directement à partir d'un électrocardiogramme ou à partir des données d'un capteur pression du volume sanguin (« Blood Volume Pressure », BVP) (voir figure 2.10 [c]). Ces capteurs de fréquences cardiaques permettent également d'extraire d'autres indices physiologiques tels que la pression sanguine, l'intervalle inter-battement ou la variabilité de la fréquence cardiaque dont l'augmentation est souvent associée à la présence de stress et de frustration (Gunes et al., 2010; Mandryk et al., 2006).

Une combinaison de senseurs, qui mesurent la réponse galvanique de la peau et la fréquence cardiaque, permet de renseigner mutuellement sur la valence et l'activation émotionnelle, et donc de situer selon l'approche dimensionnelle l'état de l'individu. Ces deux indicateurs ont été également employés dans des modèles d'apprentissage machine pour détecter des émotions discrètes telles que la frustration des apprenants (Mcquiggan et al.,

2007b), le stress (Zhai et al., 2006) ou encore pour déterminer l'état d'immersion dans un environnement virtuel (Wiederhold et al., 2001).

La fréquence de respiration permet également de renseigner sur l'activation émotionnelle (Gunes et al., 2010). Cet indicateur est mesuré à partir d'un capteur généralement sous la forme d'une ceinture placée sur la cage thoracique permettant de calculer le nombre de cycles respiratoires (inspirations et expirations) par une minute. Une respiration profonde et lente correspond généralement à un état émotionnel avec faible activation comme l'état de relaxation par exemple. En revanche, une accélération de la fréquence cardiaque est synonyme d'une émotion de forte activation comme la colère ou la joie. Par exemple, la fréquence de respiration a été utilisée par Blanchard et ses collègues (2007) pour évaluer l'influence (positive ou négative) de stratégies motivationnelles sur l'état de l'apprenant.

L'activité cérébrale est également utilisée pour renseigner sur l'état émotionnel. En effet, cet indicateur est mesuré à l'aide d'une technique connue sous le nom de L'ÉlectroEncéphaloGraphie (EEG) qui consiste à positionner un ensemble de capteurs (électrodes) sur la surface du cuir chevelu (voir figure 2.10 [d]). L'utilisation de l'EEG pour la détection émotionnelle est souvent reliée à la détection d'émotions discrètes en utilisant des algorithmes d'apprentissage machine. Heraz et ses collègues (2009) ont proposé de prédire un ensemble d'états émotionnels de l'apprenant en se basant sur des algorithmes d'apprentissage machine, qui utilisent comme prédicteurs des bandes de fréquences extraites de l'activité cérébrale (Heraz et al., 2009; Chaouachi et al., 2009; Jraïdi et al., 2009).

L'utilisation des principes de l'informatique physiologiques est de plus en plus courante dans les IHM en général et les STEI en particulier, surtout avec le progrès technologique constant en matière de création de capteurs sophistiqués, précis, non intrusifs et de faible coût. Ces capteurs peuvent être intégrés facilement dans les systèmes et permettent d'améliorer considérablement leur capacité à détecter les émotions. Plusieurs STEI optent pour une approche multimodale qui combine un ensemble de capteurs physiologiques, des mesures comportementales et même des techniques d'auto-évaluation (Arroyo et al., 2009; Burleson et al., 2007; D'Mello et al., 2007; Jraïdi et al., 2013b). Arroyo et ses collègues (2009) ont même expérimenté la possibilité d'utiliser cette approche en dehors du contexte de laboratoire, en

installant un ensemble de senseurs dans une salle de classe réelle de 25 apprenants. Chaque apprenant interagissait avec un STEI appelé Wayang Outpost dédié à l'enseignement de la géométrie. Les émotions ont été détectées à partir de la réponse galvanique de la peau (avec l'utilisation d'un gant), la posture (avec l'utilisation d'un siège détecteur de posture), les expressions faciales (avec l'utilisation d'une caméra) et la pression exercée sur la souris (avec l'utilisation d'une souris avec des détecteurs de pression). Les résultats obtenus par cette étude prouvent qu'il est possible d'utiliser avec succès cette approche multimodale hors contexte laboratoire sans gêner ou perturber les apprenants. Les auteurs ont également montré que la combinaison de senseurs permet d'obtenir de meilleures performances pour la détection émotionnelle que le recours à un seul mode de détection en utilisant chacun de ces senseurs d'une manière unique (Arroyo et al., 2009).

## 2.8 Stratégies d'adaptation dans les STEI

Dans cette section, nous allons survoler les principales stratégies utilisées par les STEI pour s'adapter aux émotions des apprenants. En effet, cet axe de recherche constitue un aspect fondamental dans le fonctionnement des STEI dans la mesure où les interventions du système doivent être déterminées en fonction de l'état émotionnel de l'apprenant et selon une logique favorable à son apprentissage. Il s'agit de définir comment un système qui mesure les émotions des apprenants peut agir intelligemment en conséquence (Picard et al., 2004). L'action du système peut être, par exemple, orientée pour éviter ou corriger une situation émotionnelle négative pour l'apprentissage. Elle peut être également initiée en vue de générer ou maintenir un état émotionnel favorable à l'apprentissage.

Les stratégies d'adaptation dans les STEI définissent le moment et la nature des interventions établies par le système. On peut distinguer deux types de stratégies qui peuvent être adoptées par les STEI à savoir : les stratégies *orientées tâches* (ou *problèmes*) et les stratégies *orientées émotions* (Chaffar et al., 2009; Robison et al., 2009b).

Les stratégies d'intervention orientées tâches consistent à agir sur la tâche d'apprentissage que l'apprenant est en train d'exécuter en vue d'améliorer ou de corriger un état émotionnel généralement défavorable à l'apprentissage (Robison et al., 2009b). Il s'agit

donc d'agir indirectement sur les émotions en intervenant sur les éléments de la tâche ou l'activité d'apprentissage. Par exemple, le système peut fournir une assistance sous la forme d'une astuce ou d'explications, s'il détecte un état de frustration chez l'apprenant lors de la résolution d'un exercice.

Les stratégies d'intervention orientées émotions visent à agir directement sur les émotions de l'apprenant sans modifier ou remédier à la situation d'apprentissage qui est à l'origine de ces émotions. Le principe de cette stratégie est de rendre l'apprenant conscient de son état émotionnel, dans l'optique de l'engager dans un processus de réflexion sur comment ses émotions sont en train d'influencer (ou affecter) son apprentissage (Picard et al., 2004). Par exemple lors de la résolution d'un exercice, si un état de frustration est détecté chez l'apprenant, une intervention orientée émotion consisterait à lui fournir un message d'encouragement. Toujours dans les stratégies orientées émotions, certains chercheurs proposent d'utiliser des stratégies orientées émotions implicites en utilisant des techniques subliminales ( Jraidi et al., 2012; Jraidi et al., 2010a; Jraidi et al., 2010b, 2011).



Figure 2.11 – Interface du système de Prendinger et Ishizuka (2005)

Dans les STEI, ces deux types de stratégies d'intervention peuvent être utilisées d'une manière combinée. Toutefois, les chercheurs ont tendance à privilégier davantage les stratégies orientées émotions. Dans leur système nommé « The Empathic Compagnon », Prendinger et Ishizuka (2005) utilisent une stratégie d'intervention basée principalement sur *l'empathie* pour s'adapter aux émotions des apprenants (voir figure 2.11). Ce système, qui enseigne comment réussir un entretien d'embauche à travers un ensemble de questions et de mise en situation, mesure les émotions des apprenants en analysant leurs mesures physiologiques (fréquence cardiaque et réponse galvanique de la peau). Un caractère animé (compagnon virtuel) interagit avec l'apprenant en lui fournissant des répliques empathiques adaptées en fonctions de l'état émotionnel détecté (calme, joie ou frustration) et de la situation d'apprentissage. Par exemple, si l'apprenant est frustré suite à une question, le compagnon intervient afin de montrer à l'apprenant qu'il est conscient de son état et qu'il comprend que la question peut être un peu gênante pour lui. L'objectif du compagnon est donc de rassurer l'apprenant et l'aider à surmonter sa frustration. En revanche, le compagnon intervient avec des messages d'encouragements et de félicitations si l'apprenant est calme ou joyeux (Prendinger et al., 2005a).

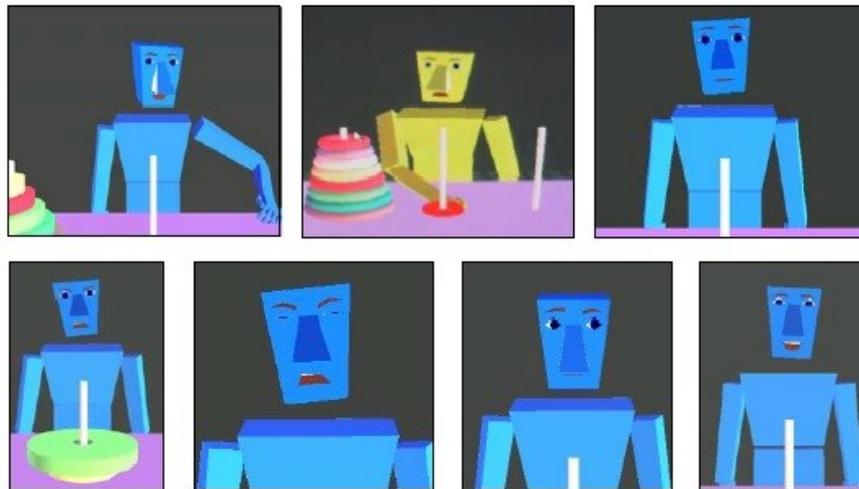


Figure 2.12 – Exemples de réactions émotionnelles exprimées par l'agent virtuel de Burleson (2006)

Le concept d'empathie ainsi que le recours à un caractère virtuel ont été également repris par Burleson (2006). L'idée était de créer un agent émotionnel (compagnon) qui

intervient en *miroitant* les émotions des apprenants (figure 2.12). Dans son système qui enseigne comment résoudre le problème des tours de Hanoi, Burleson (2006) utilise une approche multimodale pour la détection des apprenants avec : une caméra, un siège pour la détection de la posture, une souris à pression, un capteur physiologique pour la conductivité de la peau et une caméra pour les expressions faciales. La stratégie d'intervention empathique du système consiste à ce que l'agent virtuel exprime les mêmes états émotionnels de l'apprenant, en imitant ses expressions faciales et ses gestes. L'agent aura par exemple l'air ennuyé s'il détecte que l'apprenant est ennuyé, et il va sourire s'il observe un sourire chez l'apprenant. L'objectif principal de cette stratégie empathique est de renforcer la relation entre l'apprenant et la machine en renforçant l'aspect plus social et interactif. Pour Burelson (2006), le fait de miroiter des émotions négatives telles que la frustration en cas d'échec dans la résolution du problème par exemple peut aider l'apprenant à mieux contenir cet état émotionnel et donc à poursuivre ses tentatives de découverte de la solution.

Le système virtuel 3D Crystal Island pour l'enseignement de la biologie englobe également des stratégies d'interventions basées sur l'empathie pour répondre aux émotions de l'apprenant (Rowe et al., 2009). Ce système, qui utilise un modèle de transition affective construit à partir d'une série d'expérimentations, déclenche son intervention en fonction de l'émotion détectée, l'émotion que le système vise à produire chez l'apprenant et des variables de la situation d'apprentissage (McQuiggan et al., 2006; McQuiggan et al., 2008a). Par exemple, si l'apprenant ressent de l'ennui, le système va choisir en fonction de cet état émotionnel et de la situation d'apprentissage, l'intervention optimale qui permettra à l'apprenant de transiter de cet état d'ennui à un état émotionnel positif cible (tel que la joie ou le plaisir). Les stratégies d'interventions prévues par le système peuvent être empathiques à travers les personnages virtuels de l'environnement qui vont exprimer les mêmes émotions que l'apprenant. Elles peuvent être également sous la forme d'encouragements et de dialogues qui incitent l'apprenant à changer d'état émotionnel. Ce type d'interventions est qualifié par les auteurs d'empathie réactive. Le système prévoit également des interventions centrées sur la tâche en fournissant des indications et des rappels (Robison et al., 2009a; Robison et al., 2009b). La technique de détection émotionnelle utilisée dans Crystal Island se base sur une approche multimodale qui combine l'auto-évaluation, des variables comportementales

d'interactions (telles que les actions et les intentions de l'apprenant) et des variables physiologiques dans un modèle d'apprentissage machine (McQuiggan et al., 2007a; McQuiggan et al., 2008b).

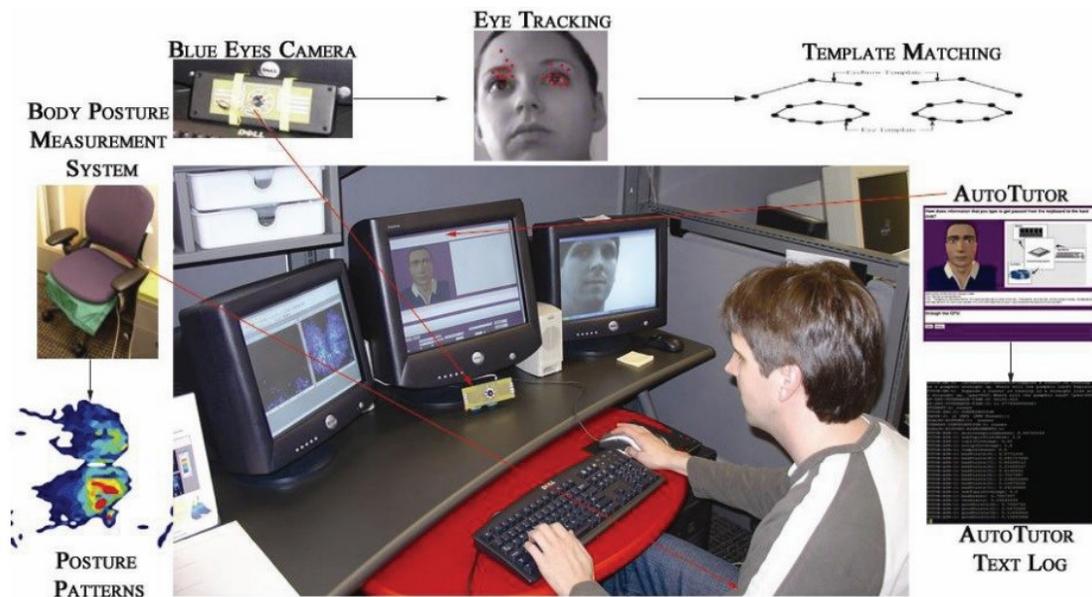


Figure 2.13 – Dispositifs de détection des émotions utilisés dans le système Autotutor (D'Mello et al., 2007).

Le STI Autotutor (voir figure 2.13) initialement développé sans considération émotionnelle et uniquement avec des stratégies d'intervention purement cognitives (Graesser et al., 1999) a lui aussi subi au fil des années plusieurs innovations afin d'intégrer les principes de l'informatique affective. Ce système qui utilise des techniques de traitement de langues naturelles pour comprendre, corriger et répondre aux actions des apprenants a commencé d'abord à intégrer un ensemble de techniques de mesure comportementales des émotions (détection des émotions à partir de l'analyse du dialogue, la posture et les expressions faciales) (D'Mello et al., 2007), et par la suite, physiologiques (rythme cardiaque, respiration, conductivité de la peau et activité musculaire) (Hussain et al., 2011). Les stratégies d'interventions employées par Autotutor se basent sur un ensemble de règles de production qui permettent d'inférer la meilleure action que le système va entreprendre (D'Mello et al., 2009). Ces règles utilisées tiennent compte de l'émotion négative détectée (ennui, frustration ou confusion), du niveau de confiance de la classification de cette émotion, de l'état

émotionnel précédent, du niveau de compétence de l'apprenant et de la qualité de la réponse donnée. L'intervention initiée par le système peut être sous la forme d'une assistance (p. ex. explication, rappel, résumé, astuce, indication sur l'erreur, etc.), d'une réponse émotionnelle de soutien (empathie ou encouragement) ou bien sous la forme d'une réponse émotionnelle qui tente de subtilement *secouer* (« shakeup Autotutor ») l'apprenant et de l'inciter à modifier lui-même son état. Par exemple, dans le cas où un état d'ennui est détecté, le système peut secouer l'apprenant en l'incitant vivement à finir l'activité au lieu de l'assister.

D'autres STEI comme le système Wayang Outpost (Arroyo et al., 2007) ou le système GURU (D'Mello et al., 2012b) utilisent des stratégies d'interventions afin de réparer des états négatifs spécifiques tels que l'ennui et le désengagement. Le principe de ces interventions consiste à réorienter l'apprenant vers l'activité et de l'informer sur les conséquences de son état sur son apprentissage. Chaffar et Frasson (2004) dans leur système EMILIE proposent d'utiliser des techniques d'induction émotionnelle à fin d'agir sur l'état émotionnel de l'apprenant. Ce système propose à l'apprenant d'imaginer des situations plaisantes à fin réduire l'impact des émotions négatives (Chaffar et al., 2004).

## **2.9 Problématiques de recherche**

L'intégration de la dimension émotionnelle a constitué un véritable changement dans la conception des STI. Comme nous l'avons montré dans les sections précédentes, plusieurs travaux de recherche ont été entrepris dans le but d'augmenter ces systèmes par des mécanismes permettant de détecter les émotions de l'apprenant, et d'adapter l'apprentissage selon son état émotionnel. Toutefois, malgré les progrès réalisés dans ce domaine, plusieurs insuffisances sont à noter.

En effet, la reconnaissance des émotions a toujours été une tâche très complexe en raison de la complexité même de leurs définitions, origines et manifestations. Les techniques communément utilisées pour la détection des émotions, en dépit de plusieurs améliorations, ne sont pas toujours très efficaces.

Les questionnaires d'auto-évaluations sont généralement sujets à des considérations subjectives de la part de l'apprenant. Par exemple, certains apprenants vont se soucier de la façon avec laquelle ils vont être perçus par le système plutôt que d'évaluer le plus objectivement possible leurs émotions (Picard et al., 2005).

La caméra, qui est souvent utilisée pour l'analyse des expressions faciales, n'arrive pas à déchiffrer correctement les émotions des apprenants dont les visages sont peu expressifs ou impassibles (Russell et al., 2003). Par ailleurs, l'analyse de la posture et de la gestuelle n'est pas toujours efficace pour reconnaître tous les types d'émotions (Gunes et al., 2010; Tao et al., 2005). Enfin, l'analyse de la voix est aussi insuffisante, dans la mesure où elle ne permet pas de relater de manière précise l'état émotionnel de l'individu (Jraidi et al., 2013b; Tao et al., 2005).

Les mesures physiologiques quant à elles, malgré la corrélation de certains indicateurs avec des dimensions émotionnelles (tel que le rythme cardiaque qui est corrélé avec la valence des émotions, et les réponses galvaniques de la peau qui sont associées à leur activation), ne disposent pas de modèles génériques qui permettent de reconnaître d'une façon précise tous les types d'émotions. En effet, les individus ont tendance à réagir différemment lorsqu'ils ressentent une émotion. Certains sont nettement plus réactifs que d'autres, ce qui fait que ces indices physiologiques peuvent avoir des niveaux de variation différents d'un individu à un autre (Arroyo-Palacios et al., 2008; Liao et al., 2006; Ward et al., 2003).

Ces limites font que plusieurs systèmes optent pour une approche multimodale combinant plusieurs sources d'informations (en recourant à plusieurs senseurs) pour évaluer les émotions de l'apprenant. Néanmoins, cette approche qui vise à améliorer la détection des émotions nécessite des mécanismes de synchronisation et de fusion des différentes sources de données émotionnelles (Gunes et al., 2010; Pantic et al., 2003; Tao et al., 2005). Ces mécanismes doivent être suffisamment sophistiqués, non seulement pour traiter, coordonner et concilier les données, mais également pour trouver les relations mutuelles qui existent entre les différentes sources d'informations (Tao et al., 2005). Pour Tao et Tan (2005), une approche multimodale sans mécanismes précis de fusion et de coordination peut dégrader considérablement la justesse de la détection. De plus, la complexité, entre autres

algorithmique, de ces mécanismes tend généralement à croître en fonction du nombre de moyens de détection utilisés, ce qui rend très difficile leur utilisation dans les environnements interactifs.

Une autre problématique liée à la modélisation émotionnelle même dans les STEI est l'absence de consensus clair quant aux émotions qu'il faut en compte lors des sessions d'interaction avec les apprenants (Baker et al., 2010; Craig et al., 2004; D'Mello et al., 2007; Graesser, 2009). La liste des émotions considérées dans les STEI varie d'un auteur à un autre. Certains chercheurs optent pour la modélisation d'un seul état émotionnel, tels que la frustration (Kapoor et al., 2007), la confusion (Grafsgaard et al., 2011), l'ennui (D'Mello et al., 2012b), l'immersion (« flow ») (Rodrigo et al., 2008), etc.

D'autres chercheurs favorisent l'idée d'identifier un ensemble d'émotions. Par exemple, D'Mello et ses collègues (2007) modélisent un ensemble d'états émotionnels, que les auteurs considèrent comme les plus fréquemment exprimées lors de l'apprentissage. Ces états émotionnels incluent notamment : l'ennui, la confusion, l'immersion, la frustration et la curiosité. Cependant, l'ensemble des états considérés ne cesse d'évoluer à travers les différentes études réalisées (D'Mello et al., 2009, 2012a; D'Mello et al., 2012b; D'Mello et al., 2008). Forbes-Riley et ses collègues (2008) s'intéressent à identifier la frustration et l'incertitude. Les auteurs ont motivé leur choix de modéliser ces deux états par la nature de leur système tutoriel, qui est essentiellement basé sur les dialogues entre l'apprenant et le système, et donc qui est capable de générer fréquemment ces émotions chez l'apprenant (Forbes-Riley et al., 2008). Arroyo et ses collègues (2009) modélisent les états d'intérêt, d'excitation, de frustration et de confiance. Pour les auteurs, le suivi de ces états émotionnels peut être bénéfique pour prédire les performances d'apprentissage dans un contexte hors laboratoire (une classe réelle). En revanche, plusieurs travaux optent plutôt pour une approche dimensionnelle pour la modélisation des émotions (Dragon et al., 2008; Hussain et al., 2011; Prendinger et al., 2005a; Zakharov et al., 2008). Cette approche ne restreint pas le nombre des émotions à considérer ; elle permet de situer, de façon approximative, l'état émotionnel de l'apprenant dans un espace à deux dimensions selon : la valence et l'activation. Par exemple, une valence positive et une activation positive indiquent que l'apprenant ressent une certaine

émotion parmi la joie, l'intérêt ou le plaisir, sans déterminer explicitement la nature exacte de l'émotion.

Ainsi, les recherches dans la modélisation émotionnelle dans les STEI ne disposent pas d'une approche standard pour modéliser les émotions, ni de théorie claire qui précise, selon le contexte d'apprentissage, quelle émotion mesurer. L'analyse de la littérature nous montre également que les émotions représentées peuvent varier selon plusieurs facteurs tels que le type de l'environnement du STEI (web, jeux sérieux, système narratif, etc.), la nature des activités (lecture, résolution de problèmes, questions-réponses, etc.) ou les techniques utilisées pour mesurer les émotions (auto-évaluation, expressions faciales, posture, senseurs physiologiques, etc.).

De plus, bien qu'il soit établi que les émotions jouent un rôle prépondérant dans le processus d'apprentissage, il ne suffit pas de les reconnaître pour pouvoir évaluer directement leur impact sur la qualité de l'apprentissage. En effet, la relation qui existe entre les émotions et l'apprentissage n'est pas forcément linéaire (Graesser et al., 2007). L'apparition de certaines émotions, qualifiées de négatives a priori, telles que la frustration, le stress ou la confusion, ne signale pas forcément des conditions émotionnelles défavorables à l'apprentissage. Par exemple, une récente analyse de 21 études sur les manifestations émotionnelles au cours de l'apprentissage par ordinateur (au total 1420 apprenants et 1058 heures d'interactions) a révélé que la confusion est parmi les états les plus fréquemment ressentis par les apprenants (D'Mello et al., 2014). Pour les auteurs, l'apprentissage de notions complexes provoque souvent de la confusion qui peut être dans certains cas révélatrice d'un certain effort d'analyse et de synthèse ainsi que d'une profonde réflexion de la part des apprenants. D'un autre côté, plusieurs études ont pu trouver que ce même état de confusion est symptomatique d'une impasse cognitive, qui peut bloquer le processus d'apprentissage (Craig et al., 2004; VanLehn et al., 2003). L'état de stress peut également avoir deux facettes : positive et négative. Le stress positif appelé également stress fonctionnel (« eustress ») est synonyme de motivation et de stimulation cognitive. Le stress négatif appelé détresse (« distress ») monopolise les fonctions cognitives de l'apprenant et l'empêche de se concentrer (Muse et al., 2003). Enfin, l'état de frustration, qui est souvent considéré comme un état négatif pour l'apprentissage, est

non seulement inévitable dans les Interactions Homme-Machine, mais peut aussi dans certaines situations évoquer un certain degré d'implication de la part de l'apprenant (Gee, 2004). Par exemple, un apprenant qui tient à résoudre et à comprendre un exercice peut sentir de la frustration en cas d'échec.

De ce fait, la relation complexe qui existe entre les émotions et l'apprentissage humain fait qu'une simple reconnaissance des émotions de l'apprenant n'est pas toujours suffisante pour les STEI. Ces systèmes doivent également être capables de déterminer comment ces émotions sont en train d'influencer les capacités cognitives de l'apprenant et son processus d'apprentissage. Cet aspect n'est pas évident, étant donné que la manière avec laquelle les individus ressentent les émotions et la façon avec laquelle les émotions les influencent peuvent dépendre de plusieurs facteurs. Ces facteurs peuvent être d'ordres individuels (tels que le genre, la personnalité, le tempérament ou le vécu d'un individu), contextuels (tels que les circonstances déclencheurs de l'émotion ou l'environnement d'apprentissage), motivationnels (tels que les objectifs fixés, les objectifs réalisés ou les moyens à disposition, etc.) ou culturels (Bachorowski et al., 1995; Keltner et al., 2001; Markus et al., 1991; Russell, 2003).

### **2.9.1 Vers l'intégration des indicateurs cérébraux d'états mentaux**

Afin de pallier ces différentes problématiques, nous proposons dans le cadre de cette thèse d'utiliser l'activité cérébrale et plus précisément l'ÉlectroEncéphaloGraphie (EEG) comme principal moyen de détection. En particulier, nous allons nous intéresser à l'identification d'indicateurs cérébraux *d'états mentaux*.

Un état mental peut être vu comme un état affectif complexe qui englobe certes des dimensions émotionnelles, mais surtout cognitives (el Kaliouby, 2005). Dans l'interaction entre humains, il est très fréquent pour un individu de donner un sens aux actions d'un autre en lui attribuant un état mental. Par exemple, en observant qu'un individu regarde fréquemment sa montre on peut en déduire qu'il est pressé (Baron-Cohen et al., 1997) ; des émotions (négatives ou positives) peuvent être associées à cet état (par exemple de la frustration ou au contraire de l'engouement).

Ainsi, l'identification d'un état mental, plutôt qu'une émotion, peut aider à mieux comprendre le comportement et les actions d'un individu et par conséquent à mieux répondre ou s'adapter à son état.

Dans cette thèse, nous proposons donc de modéliser l'état mental de l'apprenant afin de mieux analyser et suivre son comportement lors de son interaction avec le STI. Plus précisément, nous allons nous intéresser à la modélisation de l'état *d'engagement mental* et la *charge mentale de travail* à partir de l'activité cérébrale de l'apprenant. La modélisation de l'état d'engagement permettra d'évaluer quantitativement le *niveau d'attention et d'implication* mentale de l'apprenant dans les différentes activités d'apprentissage. La charge mentale de travail, appelée également charge cognitive « cognitive load », permettra quant à elle de mesurer la quantité *d'énergie mentale* dépensée par l'apprenant dans la réalisation des différentes tâches d'apprentissage.

Notre objectif est donc d'avoir deux indicateurs fiables pour quantifier respectivement le niveau d'investissement de l'apprenant dans la tâche, et l'effort fourni durant l'apprentissage.

Nous allons également nous intéresser dans cette thèse à l'analyse de ces états mentaux par rapport aux émotions. Autrement dit, nous allons vérifier si les variations de l'engagement et de la charge mentale de travail sont accompagnées par certaines manifestations émotionnelles que nous tenterons d'identifier.

### **2.9.2 Un système adaptatif selon l'état mental de l'apprenant**

En dépit des nombreux travaux de recherche menés sur la modélisation des émotions, très peu de systèmes ont été développés pour détecter *en temps réel* les émotions des apprenants. En effet, la majeure partie des algorithmes de reconnaissance proposés dans la littérature ont été développés et testés rétrospectivement (en « offline ») pour identifier et analyser, à posteriori, les émotions de l'apprenant. De plus, comme nous avons pu le noter dans la section 2.8 de ce chapitre, les systèmes tutoriels qui ont intégré des mesures émotionnelles pour adapter dynamiquement leur processus de formation en temps réel sont aussi relativement rares.

Dans cette thèse, nous allons donc tenter de démontrer la possibilité de *détecter* les indicateurs cérébraux d'engagement et de charge de travail de l'apprenant en temps réel, mais aussi *d'adapter*, en temps réel également, le processus d'apprentissage en fonction de ces indicateurs.

Par ailleurs, les stratégies actuelles d'adaptation orientées émotions, c.-à-d. qui visent à agir directement sur les émotions des apprenants à travers des messages d'encouragement, de félicitation ou d'empathie, ne sont pas toujours efficaces (Bailey et al., 2008; Monk et al., 2002). Par exemple, le recours à des agents virtuels, qui vont miroiter les émotions de l'apprenant, peut dans certains cas provoquer des réactions négatives chez ce dernier. (Woolf et al., 2009). En effet, bien que le caractère social d'un agent animé puisse divertir l'apprenant, le fait de mimer ses expressions faciales ou ses gestuelles, lorsque celui-ci n'arrive pas à comprendre un concept ou à résoudre un exercice, peut ne pas être pris pour de l'empathie et générer au contraire des émotions négatives comme de la frustration ou de la colère.

En outre, comme le démontrent certaines études (Robison et al., 2009b), les apprenants ont tendance à préférer les stratégies d'adaptation orientées tâches qui ne visent pas à répondre directement à leurs émotions, mais qui visent plutôt à adapter la tâche (ou l'activité) qu'ils sont en train d'exécuter en proposant une indication, un rappel ou une définition.

Enfin, plusieurs études empiriques, qui ont analysé l'impact de ces stratégies d'adaptation orientées sur les émotions, ne démontrent pas d'effet positif statistiquement significatif, ni sur les performances de l'apprenant ni sur ses réponses émotionnelles (Beale et al., 2009; Burleson et al., 2007; Kim, 2005; Mori et al., 2003; Woolf et al., 2009).

Dans cette thèse nous essayerons de démontrer qu'un système tutoriel adaptatif, selon les indicateurs cérébraux de l'apprenant, ou autrement dit qui tient compte de l'engagement et de la charge mentale de travail dans ses stratégies d'adaptation, peut améliorer significativement l'expérience d'apprentissage aussi bien au niveau des performances cognitives des apprenants, mais aussi au niveau de leur satisfaction par rapport au déroulement de leur apprentissage.

## **Chapitre 3**

# **Modélisation de l'engagement mental des apprenants**

### **3.1 Introduction**

Comme nous l'avons spécifié dans l'introduction de cette thèse, notre objectif global est d'améliorer le fonctionnement des STEI en intégrant des indicateurs cérébraux d'engagement et de charge mentale de travail. Dans cette optique, nous proposons dans une première étape d'intégrer, dans un contexte d'apprentissage, un index d'engagement mental extrait à partir de l'activité cérébrale de l'apprenant. Cet index a été développé dans le cadre des travaux de recherche menée dans la « National Aeronautics and Space Administration » (NASA) afin de mesurer le niveau d'attention des pilotes lors de simulations de vols (Pope et al., 1995).

Dans ce chapitre, nous proposons, à travers une première étude expérimentale, de mesurer cet index d'engagement mental dans un environnement d'apprentissage et d'analyser les contributions que pourrait apporter l'utilisation de cet indicateur dans le suivi du processus d'apprentissage.

Nous allons tout d'abord (1) vérifier si cet index d'engagement est valide et pertinent pour les STEI. Nous allons nous assurer que cet indicateur cérébral, qui a été utilisé

principalement dans le contexte d'aviation, est valable dans le cadre de l'apprentissage. Nous vérifierons entre autres si l'index d'engagement peut aider à mieux suivre et analyser les performances et le comportement des apprenants. Nous allons ensuite (2) vérifier si — en plus de l'information cognitive sur l'attention de l'apprenant — cet index d'engagement peut donner des informations sur *l'état affectif* de l'apprenant. En d'autres termes, nous allons analyser la relation qui pourrait exister entre les variations de l'index d'engagement mental et les états émotionnels des apprenants.

Ce chapitre est structuré de la façon suivante. Dans la deuxième section, nous présentons la notion d'engagement mental. Dans la troisième section, nous exposons la technique que nous allons utiliser pour extraire l'index d'engagement à partir de l'activité cérébrale. Dans la quatrième section, nous décrivons l'étude expérimentale réalisée pour analyser cet index dans un contexte d'apprentissage. Dans la cinquième section, nous présentons les principaux résultats trouvés. Nous clôturerons enfin ce chapitre par une discussion de ces résultats.

### **3.2 La notion d'engagement mental**

L'engagement mental (ou l'attention) est généralement défini comme un processus de détection et de stockage de stimuli qui sont pertinents par rapport aux objectifs d'un individu. En fonction de ces objectifs, certains stimuli seront traités par l'individu d'une manière plus prioritaire que d'autres (Reed, 1977). Pour Gagné (1985), l'engagement mental est la phase initiale dans un processus d'apprentissage (Gagné, 1985). Dans cette phase, la réception d'un stimulus provoque une activation de la fonction de perception sélective qui enregistrera dans la mémoire à court terme certaines caractéristiques du stimulus jugées comme pertinentes par l'apprenant. Par exemple, suite à la visualisation d'une image qui intéresse l'apprenant, certaines caractéristiques visuelles de cette image (telles que la forme, les couleurs, etc.) sont directement retenues dans la mémoire à court terme. Et ce n'est qu'à la suite de cette première phase d'attention, qu'un processus d'acquisition se déclenche pour encoder ces informations dans une forme plus sémantique qui servira pour les autres processus cognitifs (tels que la mémorisation à long terme, le raisonnement, etc.).

L'engagement mental peut être également défini comme la capacité d'un individu à se focaliser sur une tâche en particulier, sans être perturbé par d'autres facteurs externes (à cette tâche) (Stamenović et al., 2005). Si, au cours de la réalisation d'une tâche, un individu arrive à maintenir un haut niveau d'attention sur une période relativement longue, alors, on dit qu'il est *concentré*. Si son attention est orientée de façon à réagir immédiatement au moindre changement qui se produit aléatoirement, alors on dit qu'il est *vigilant*.

L'engagement mental est considéré également comme une composante de la motivation (Keller, 1987). En effet, cette notion peut également refléter un certain degré d'intérêt ou de motivation, si l'état d'engagement mental au cours de la réalisation d'une tâche est accompagné d'états affectifs positifs (tel que la satisfaction par exemple). Toutefois, il est important de ne pas automatiquement associer la notion de motivation à l'engagement. En effet, un apprenant peut être engagé dans une tâche sans être vraiment motivé (par exemple, un apprenant peut être contraint à faire une tâche qu'il n'apprécie pas particulièrement).

En apprentissage, le manque d'engagement (ou désengagement) est souvent attribué à la nature de la tâche que l'apprenant est en train d'effectuer. Les tâches trop faciles, non stimulantes, monotones, ou à l'inverse difficiles, contenant beaucoup d'informations, ou qui dépassent les compétences de l'apprenant, ont tendance à engendrer un abaissement du niveau d'engagement mental (Chanel et al., 2008 ; Pekrun, 2006).

### **3.2.1 Modélisation de l'engagement dans les STEI**

Dans les STEI, comparées à des émotions comme la frustration, le stress ou la confusion, les tentatives de modélisation de l'engagement sont beaucoup moins fréquentes, malgré l'importance fondamentale de cette notion dans l'apprentissage. La plupart des approches utilisées se basent généralement une approche comportementale qui surveille la présence (ou l'absence) d'engagement à partir d'indices contenus dans les expressions faciales, la posture et les mouvements du corps ou encore les variables d'interaction (Asteriadis et al., 2009; Banda et al., 2011; Beck, 2005; D'Mello et al., 2007; Sanghvi et al., 2011).

L'analyse des expressions faciales pour la détection de l'engagement procède de la même manière que l'approche utilisée pour reconnaître les émotions, c'est-à-dire en se basant sur l'identification des unités d'action relatives à l'expression d'engagement. Par exemple, un apprenant engagé aura tendance par exemple à serrer les yeux et à lever les sourcils (Kapoor et al., 2001).

L'analyse de la posture et des mouvements du corps pour la détection de l'engagement se fait généralement à l'aide de caméras ou de sièges équipés avec des détecteurs de postures et de mouvements. Par exemple, Asteriadis et ses collègues (2009) ont développé une approche qui prédit le niveau d'attention des apprenants lors de l'apprentissage de la lecture en analysant la direction du regard, les mouvements de la tête et la position des mains (Asteriadis et al., 2009). L'emplacement des mains et la direction du regard ont été également utilisés pour savoir si l'apprenant est engagé ou pas dans l'activité (Kapoor et al., 2001); par exemple si l'apprenant pose ses mains ailleurs que sur la souris ou le clavier, et regarde ailleurs, alors il est considéré comme désengagé de la tâche. Sanghvi et ses collègues (2011) utilisent également des algorithmes de vision par ordinateur pour extraire des informations sur la posture de l'apprenant (le degré d'inclinaison du dos, l'angle de courbure de dos, la quantité de mouvement et le degré de contraction du haut du corps), et entraîner un modèle d'apprentissage machine capable de déterminer si l'apprenant est engagé ou non lors de l'interaction avec un système d'apprentissage d'un jeu d'échecs (Sanghvi et al., 2011). D'Mello et ses collègues (2007) ont également utilisé l'analyse de la posture pour déterminer si l'apprenant est engagé dans sa résolution d'exercices de physique, en utilisant un siège de détection de posture (mesurant le niveau de pression au niveau du dossier et de l'assise) (D'Mello et al., 2007).

La détection de l'état d'engagement de l'apprenant à partir des variables d'interaction se fait en analysant des indicateurs de performances dans la tâche comme : le temps émis pour lire l'énoncé ou pour répondre, la fréquence de la demande d'aide, le taux de bonnes réponses, le temps émis dans la lecture de la solution en cas d'erreur, le taux de mauvaises réponses, le nombre de fois où l'apprenant change son choix avant de répondre, etc. (Beal et al., 2006; Beck, 2005).

Les différentes techniques utilisées pour modéliser l'engagement des apprenants dans les STEI présentent les mêmes limites de la modélisation émotionnelle exposées dans la section 2.9 du chapitre précédent. En effet, les indices externes et visibles du comportement (tels que les expressions faciales ou la posture) ne sont pas toujours efficaces pour évaluer l'état de l'individu (Fairclough, 2009). Par exemple, un apprenant qui ne regarde pas l'écran, qui ne pose pas ses mains sur le clavier (ou la souris) ou qui ne présente pas d'expressions faciales relatives à un état d'engagement n'est pas forcément mentalement désengagé. L'utilisation des variables d'interaction pour juger du niveau d'engagement est également imprécise et souvent très dépendante de la nature de la tâche que l'apprenant est en train d'effectuer, et donc difficilement généralisable (D'Mello et al., 2005; Picard, 1997).

Dans cette thèse, nous nous basons sur une approche physiologique qui extrait l'engagement mental des apprenants à partir de leurs activités cérébrales. Cette approche présente l'avantage d'être plus objective et plus précise que l'approche comportementale (Fairclough, 2009). En effet, l'index d'engagement mental que nous proposons d'intégrer est directement déterminé à partir de changements dans l'activité cérébrale qui — comme toutes les réactions physiologiques — sont spontanés et incontrôlables (Allanson et al., 2004; Fairclough, 2009 ; Wagner et al., 2005). De plus, contrairement à l'analyse de la posture ou des variables d'interactions, l'index que nous proposons permet de donner une indication quantitative et continue dans le temps du niveau d'engagement mental de l'apprenant ; ceci permettra donc un meilleur suivi du comportement des apprenants tout au long de la session d'apprentissage.

### **3.3 Calcul de l'index d'engagement mental**

Dans cette partie, nous introduisons la technique d'extraction de l'engagement mental à partir de l'activité cérébrale. Nous commençons tout d'abord par présenter le procédé de *l'électroencéphalographie* utilisé pour enregistrer l'activité cérébrale.

### 3.3.1 L'électroencéphalographie

Le cerveau humain est une structure complexe composée d'environ  $10^{11}$  cellules nerveuses appelées *neurones*, reliées entre elles par  $10^{14}$  *connexions synaptiques* faisant intervenir près de  $10^{18}$  *neurotransmetteurs* chaque seconde. L'acheminement et le traitement des informations par le cerveau se fait par des charges électriques appelées *influx nerveux*.

L'ÉlectroEncéphaloGraphie (EEG) représente une des plus anciennes modalités de mesure et de suivi de l'activité cérébrale : ce procédé a été employé pour mesurer l'activité cérébrale des humains dès l'année 1920 par le physiologiste allemand Hans Berger. L'ÉlectroEncéphaloGraphie repose sur l'utilisation d'un ensemble d'électrodes apposées sur le scalp humain qui permettent de mesurer la valeur du potentiel électrique du cerveau.

La popularité grandissante de l'utilisation de cette technique dans divers domaines d'application est justifiée par le fait que c'est un procédé simple, bien maîtrisé et qui produit des données facilement interprétables avec une grande résolution temporelle. La figure 3.1 montre un exemple d'un signal EEG mesuré à partir d'une électrode pendant une seconde.

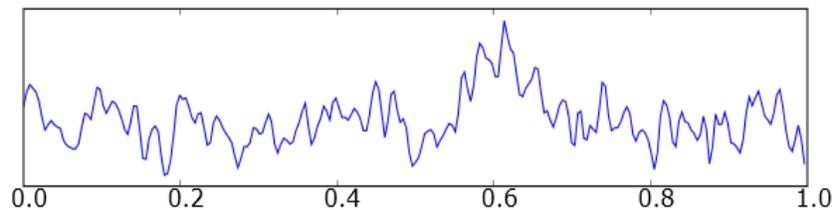


Figure 3.1 – Exemple de signal EEG sur une durée d'une seconde

L'interprétation des signaux EEG peut se faire selon deux méthodes : les **potentiels évoqués** et les **ondes cérébrales**.

1. La méthode des **potentiels évoqués** est généralement utilisée pour analyser l'impact d'un stimulus extérieur (visuel, auditif ou tactile) sur la morphologie (c.-à-d. la forme) du signal EEG brut. Cette analyse se fait sur un laps de temps d'environ une demi-seconde suivant la présentation du stimulus. En fonction de la forme que présente le signal EEG, une interprétation peut être faite sur la réponse du cerveau par rapport à ce

stimulus. Par exemple, il est établi en neuroscience qu'une augmentation du potentiel environ 300 à 500 ms après la présentation d'un stimulus (par exemple l'apparition d'une lettre sur un écran d'ordinateur) est un signe de réaction d'attention particulière vis-à-vis de ce stimulus. On parle alors de la présence de la composante *P300* (la lettre *P* fait référence à l'amplitude positive du signal et le *300* à son délai d'apparition appelé également latence). La méthode des potentiels évoqués est surtout utilisée dans le développement des Interfaces Cerveau-Machine (ICM) telles que « p300 speller » introduites par Donchin (Donchin et al., 2000). Cette interface qui se base sur la détection des composantes *P300* permet à l'utilisateur d'écrire des lettres à l'écran (une lettre à la fois) uniquement en fixant une matrice 6x6 de caractères. Les lignes et les colonnes de cette matrice s'illuminent aléatoirement provoquant une composante *P300* chez l'individu à chaque fois que la ligne ou la colonne contenant la lettre qu'il veut écrire s'illumine. Des algorithmes de détection de cette composante *P300* permettent alors à l'interface de reconnaître cette lettre.

2. La méthode des **ondes cérébrales** se base sur une décomposition du signal EEG en un ensemble de plages de fréquences (ou ondes cérébrales). L'interprétation du signal EEG se fait en fonction de la présence (ou de la prédominance) de certaines fréquences par rapport aux autres. Le tableau 3.1 présente des exemples de fréquences du signal EEG, avec les tendances mentales qui leur sont associées (Demos, 2005). Cette liste de plages de fréquences n'est pas exhaustive, et est généralement utilisée dans les applications de suivi médical ou de biofeedback.

La méthode des ondes cérébrales inclut également d'autres types de traitement sur les plages de fréquences tels que :

- La *synchronisation* : comparaison de la différence d'amplitude d'une même plage de fréquences entre deux régions. Ceci permet de détecter la présence de certains traitements cognitifs (comme le raisonnement logique, le raisonnement intuitif, etc.) (Davidson, 1988; Ehrlichman et al., 1980).

- Les *ratios* : analyse des proportions entre deux ou plusieurs plages de fréquences différentes (Lubar, 1991). Nous allons nous baser sur cette technique d'analyse pour le calcul de l'index d'engagement mental.

Tableau 3.1 – Exemples de fréquences du signal EEG (Demos, 2005)

Nom de la plage	Plage	Exemple d'activité mentale associée
Delta ( $\Delta$ )	1-4 Hz	Sommeil profond, repos et relaxation
Theta ( $\theta$ )	4-8 Hz	Intuition, concentration et créativité
Alpha ( $\alpha$ )	8-13 Hz	Repos et relaxation
Beta ( $\beta$ )	13-22 Hz	Traitement des informations et attention soutenue
High Beta (High $\beta$ )	28-32 Hz	Anxiété et hyperactivité
Gamma ( $\gamma$ )	38-42 Hz	Traitement de tâches complexes

### 3.3.2 Extraction de l'index d'engagement mental à partir des signaux EEG

La modélisation de l'engagement mental à partir des signaux physiologiques, et plus particulièrement de l'activité cérébrale, a été amplement traitée dans le domaine de l'aviation (Freeman et al., 2000; Prinzl et al., 1995). L'objectif était de créer des systèmes capables d'automatiser l'exécution de certaines tâches du pilote sans que ce dernier perde le contrôle de la situation. En d'autres termes, bien que certaines tâches soient exécutées automatiquement par le système (tel que le suivi de la trajectoire, la surveillance de différents indicateurs de bord, etc.), le pilote doit toujours maintenir un bon niveau d'attention et ne pas perdre le contrôle par rapport à ces tâches automatisées. On parle alors d'un niveau de *conscience de la situation* « situation awareness » (Prinzl Iii et al., 2002) qui permettrait de réduire le risque d'erreurs ou d'accidents (par exemple lors de la reprise du pilotage manuel). L'engagement mental du pilote était donc considéré comme un critère fondamental dans les décisions d'adaptation du niveau d'automatisation.

Dans cette optique, Pope et ses collègues (Pope et al., 1995) à la « National Aeronautics and Space Administration » (NASA) ont développé un index d'engagement mental basé sur les ondes cérébrales et inspiré d'un ensemble de résultats de recherches neuroscientifiques sur l'attention, la vigilance et l'éveil mental (Davidson, 1988; Lubar, 1991; Lubar et al., 1995; Offenloch et al., 1990). Le calcul de cet index d'engagement se fait en établissant un ratio entre les trois plages de fréquences  $\theta$  (4-8 Hz),  $\alpha$  (8-13 Hz) et  $\beta$  (13-22 Hz) de la manière suivante :

$$Eng\_Index = \frac{\beta}{\theta + \alpha}$$

Les plages de fréquences  $\theta$ ,  $\alpha$  et  $\beta$  utilisées pour l'établissement de ce ratio sont des *plages combinées*, c'est-à-dire qu'elles proviennent de la somme des plages de chaque région cérébrale mesurée.

Afin de réduire les fluctuations de cet index, les auteurs recommandent de le lisser en utilisant une moyenne sur une fenêtre mobile d'une période de temps de 20 ou de 40 secondes. Ainsi, la valeur d'un index à un instant  $t$  correspond à la moyenne totale des ratios précédents calculée sur cette période. Par exemple, si un ratio est calculé chaque seconde, l'index d'engagement à un instant  $t$ , avec une fenêtre de  $a$  secondes, sera lissé de la manière suivante :

$$Eng_t = \frac{1}{a} * \sum_{i=t-a}^t \frac{\beta_i}{\theta_i + \alpha_i}$$

Avec  $\theta_i$ ,  $\alpha_i$  et  $\beta_i$  les plages de fréquences combinées des différentes régions à l'instant  $t$ .

L'utilisation de cet index d'engagement comme critère d'adaptation dans les systèmes de simulations de vols a permis d'améliorer considérablement les performances des pilotes et de réduire leurs taux d'erreurs. Pour cette raison, nous proposons dans cette thèse de l'intégrer dans les environnements d'apprentissage afin de suivre l'engagement mental des apprenants.

### 3.3.3 Analyse de l'index d'engagement

L'index d'engagement ainsi calculé est généralement discrétisé en deux états possibles (*positif* ou *négatif*) dans le but d'avoir une idée globale sur le comportement de l'engagement mental (Freeman et al., 2000) :

- L'état positif de l'index peut être interprété, selon la tâche et le contexte, comme un indicateur d'attention, de vigilance ou de concentration.
- L'état négatif de l'index indique un désengagement mental. En fonction de la tâche et du contexte, on peut parler de relâchement mental, de manque d'implication ou encore de distraction.

Deux méthodes de discrétisation des valeurs de cet index sont fréquemment employées (Freeman et al., 2000; Pope et al., 1995) :

- 1. Comparaison par rapport à la valeur de référence « Baseline » :** appelée également *méthode absolue*, cette technique est très utilisée dans les systèmes à base de senseurs physiologiques (Levenson, 1988). Son principe consiste à identifier une valeur de référence permettant d'évaluer les données physiologiques en provenance d'un senseur. Pour l'index d'engagement, une période de 5 minutes est préconisée (Freeman et al., 2000). La référence (ou seuil) est établie en moyennant les valeurs de l'index calculées sur cette période de 5 minutes. Lors de l'établissement de cette référence, l'individu doit être dans l'état le plus neutre possible. De cette manière, si la moyenne des index d'engagement d'un individu durant une activité est inférieure à cette référence, l'état d'engagement est considéré comme *négatif*. En revanche, si cette valeur est supérieure à la référence, l'état d'engagement est considéré comme *positif*.
- 2. Analyse de la pente des données :** Cette technique consiste à déterminer une droite qui s'ajuste le plus possible à l'ensemble des valeurs de l'index d'engagement dans le temps. Il s'agit donc de faire une régression linéaire avec la méthode des moindres carrés qui minimise la fonction  $F$  déterminée par :

$$F = \sum_{i=1}^n (y_i - (ax_i + b))^2 ,$$

$$\frac{\partial F}{\partial a} = 0 \text{ et } \frac{\partial F}{\partial b} = 0$$

Avec  $y_i$  la valeur de l'index d'engagement calculé à l'instant  $i$ . La minimisation de cette fonction  $F$  permet de générer la pente  $a$  ainsi que l'ordonnée à l'origine  $b$  de cette droite de tendance. L'engagement mental durant une activité est considéré comme *positif*, si la valeur de cette pente est positive ; et *négatif* dans le cas contraire. La figure 3.2 illustre un état d'engagement mental positif sur une période d'une minute.

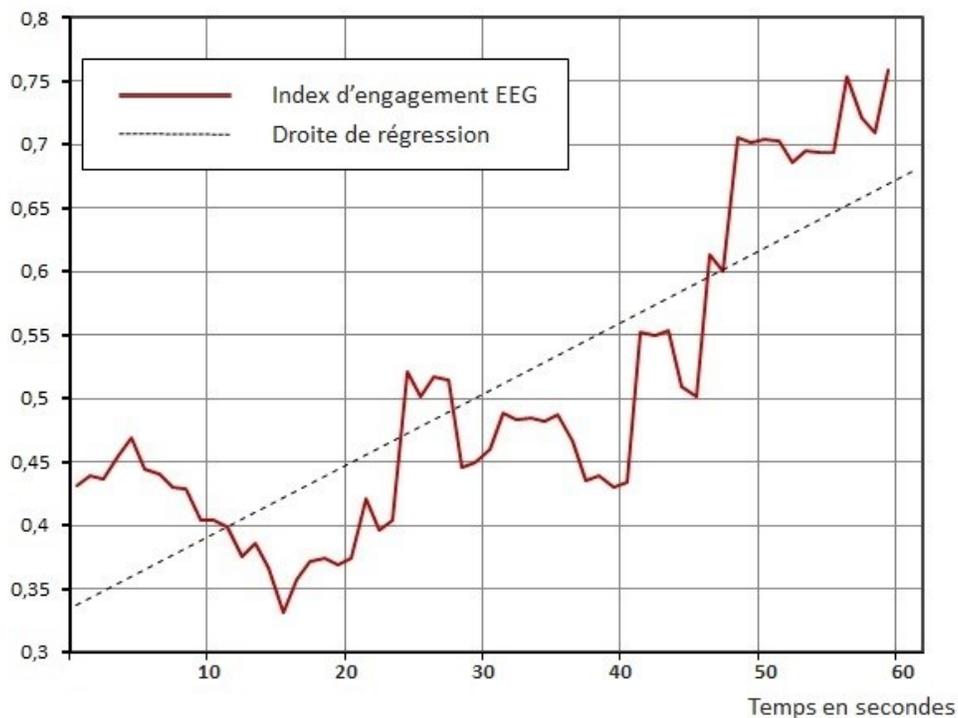


Figure 3.2 – Engagement mental positif sur une période d'une minute

Les deux méthodes de discrétisation sont considérées comme équivalentes par les développeurs de cet index (Freeman et al., 1999; Pope et al., 1995; Prinzel et al., 1995). Toutefois, la première méthode de discrétisation est souvent utilisée pour analyser

rétrospectivement (en « offline ») l'évolution de l'index d'engagement lors d'une session de mesure, tandis que la deuxième méthode est généralement utilisée lors de l'analyse en temps réel (Freeman et al., 1999 ; Freeman et al., 2000 ; Prinzel et al., 2000a ; Prinzel et al., 1995). Dans ce chapitre, nous allons nous baser sur la première méthode de discrétisation.

Après avoir introduit la technique d'extraction de l'index d'engagement mental à partir des signaux EEG, nous présentons dans la prochaine section la première expérimentation que nous avons réalisée en vue d'analyser la validité de cet index dans un contexte d'apprentissage.

### **3.4 Expérience 1 : Étude de l'index d'engagement mental dans un contexte d'apprentissage**

L'objectif de cette étude est de mesurer l'index d'engagement mental des apprenants, et d'analyser la contribution que peut apporter l'utilisation de cet index dans un contexte d'apprentissage.

Dans un premier temps, nous allons vérifier si cet index cérébral, qui a été utilisé uniquement dans un contexte de pilotage, dans des environnements hautement contrôlés et où les tâches à exécuter consistent à surveiller certains indicateurs de vol principalement, peut être valide et pertinent dans un contexte d'apprentissage, faisant intervenir des environnements beaucoup moins contrôlés, et où les apprenants sont amenés à exécuter des tâches beaucoup plus complexes, telles que le raisonnement, le calcul ou la mémorisation.

Dans un deuxième temps, nous allons analyser les relations qui peuvent exister entre cet index d'engagement et les réactions émotionnelles de l'apprenant. En effet, comme nous l'avons évoqué dans le chapitre précédent, les émotions jouent un rôle fondamental dans l'apprentissage humain. L'objectif de cette analyse sera donc de mettre en évidence l'apport que peut avoir l'intégration de cet index dans les STEI sur le plan émotionnel.

Nous formulons ainsi les deux hypothèses de recherche suivantes que nous tenterons de vérifier à travers cette première étude expérimentale :

- (1) Les apprenants qui parviennent à maintenir un bon niveau d'engagement mental, selon cet index, auront tendance à obtenir de meilleures performances d'apprentissage.
- (2) L'évolution de l'index d'engagement permet de donner des indications sur l'état affectif de l'apprenant.

La vérification de la première hypothèse nous permettra de confirmer la validité de l'index d'engagement et de souligner l'apport de l'utilisation de cette mesure cérébrale dans un contexte d'apprentissage. La validation de la deuxième hypothèse nous permettra d'affirmer que l'intégration de cet index dans les STEI peut contribuer à l'analyse des réactions émotionnelles des apprenants.

Nous décrivons dans ce qui suit notre méthodologie et dispositif expérimental, et exposons les principaux résultats trouvés.

### **3.4.1 Méthodologie et dispositif expérimental**

Afin de valider nos deux hypothèses de recherche, nous avons réalisé une expérience dans laquelle un certain nombre d'apprenants devaient interagir avec un environnement d'apprentissage de type jeu-concours. Un protocole expérimental a été mis en place afin de nous permettre d'enregistrer et d'analyser rétrospectivement leurs données physiologiques.

Le dispositif expérimental utilisé dans cette expérience était constitué des éléments suivants (voir figure 3.3) :

- Un bonnet EEG ajustable contenant un ensemble d'électrodes permettant de mesurer l'activité cérébrale. Les données en provenance de ces senseurs ont été utilisées pour extraire l'index d'engagement mental.
- Des amplificateurs (un pour région mesurée) dont l'objectif est d'amplifier le potentiel mesuré par les électrodes.
- Deux senseurs physiologiques pour mesurer respectivement la pression du volume sanguin/artériel (BVP « Blood Volume Pressure ») et la conductivité de la peau (SC « Skin Conductance ») utilisées pour la détection des états émotionnels des apprenants.

- Un encodeur qui permet de récupérer les données mesurées par les différents capteurs et les fournir à la machine.

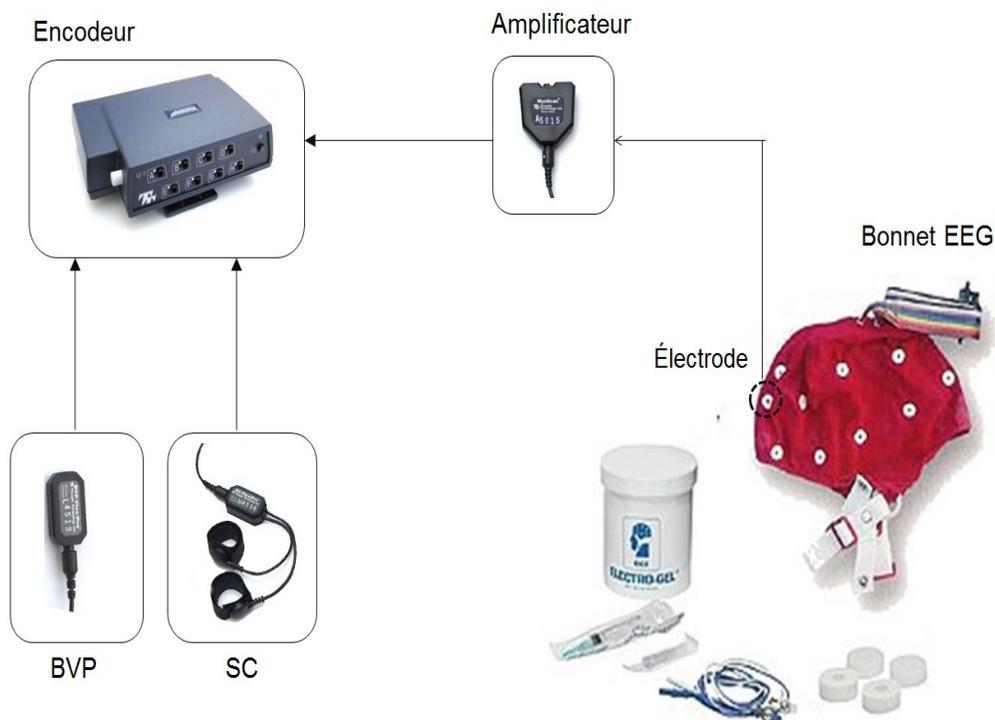


Figure 3.3 – Dispositif expérimental de la première expérience

Dans cette expérience, nous avons également utilisé deux flux vidéo afin d'enregistrer toutes les sessions de mesures. Ceci nous a permis de revoir et analyser à posteriori les différentes sessions, et donc de conserver uniquement les parties de l'expérience dont nous avons besoin pour notre analyse (ceci nous a permis par exemple d'éliminer les parties de nos expérimentations où l'on expliquait les principes de l'expérience aux apprenants ou encore les moments d'interruption). Les différentes sources de données physiologiques ont été synchronisées minutieusement dans le temps, afin de faire correspondre à chaque valeur de l'index d'engagement mental, un état émotionnel extrait à partir des deux autres capteurs physiologiques (BVP et SC). De cette manière, l'analyse conjointe de l'évolution de l'index d'état mental avec les états émotionnels était possible.

Par ailleurs, nous avons également enregistré deux paramètres d'interaction qui mesurent les performances des apprenants lors de leur interaction avec l'environnement. Ces

deux paramètres sont l'exactitude de la réponse fournie et le temps émis pour répondre aux questions posées (RT pour « Response Time »).

## **Participants**

Trente-cinq apprenants ont pris part à l'expérience. Leur participation a été compensée avec 10 dollars canadiens. Tous les participants ont signé un formulaire de consentement écrit. Les apprenants étaient tous francophones, étudiants de l'Université de Montréal et ne présentaient aucune difficulté à comprendre les questions fournies par l'environnement d'apprentissage utilisé dans cette expérience.

## **Déroulement de l'expérience**

Avant de commencer l'expérience, les apprenants étaient équipés avec les différents senseurs physiologiques (EEG, BVP et SC). Puis, durant une période d'environ 5 minutes, les valeurs de référence (« Baseline »), pour les différents senseurs, étaient enregistrées pour chaque participant. Ces valeurs permettaient de fournir un état neutre pour l'analyse de l'index d'engagement mental à partir des signaux EEG, ainsi que les émotions à partir des senseurs BVP et SC. Durant cette période, les apprenants devaient se mettre dans l'état le plus neutre possible en fixant une croix dans l'écran ; les apprenants ont été incités à ne pas être trop engagés, ni trop détendus non plus.

## **L'environnement d'apprentissage**

L'environnement d'apprentissage consistait en un jeu-concours ou « Quiz » dans lequel les apprenants devaient répondre à trois séries successives de dix questions (vrai/faux). Ce type de tâches est très utilisé dans l'analyse des réactions émotionnelles des apprenants surtout en présence de capteurs physiologiques (Kim, 2007; Kim et al., 2006; Prendinger et al., 2005b; Shen et al., 2009).

En effet, ce genre d'activité sollicite généralement une bonne implication de la part des apprenants dans la mesure où la nature même du jeu-concours est plaisante et suscite le sens de la compétition et du défi chez les apprenants. De plus, en fonction du succès ou des échecs

dans les réponses de l'apprenant, ce type d'environnement est également propice à l'apparition de certaines émotions telle que la joie ou la frustration, la confusion ou l'ennui.

La plupart des questions utilisées dans notre environnement étaient relativement simples et ne nécessitaient aucun prérequis ou compétence particuliers. Cependant, un bon niveau d'attention et de vigilance de la part des apprenants était nécessaire pour éviter de faire des erreurs faciles. Les questions étaient choisies de façon à ce que si le niveau d'attention baissait chez un apprenant, celui-ci pourrait alors commettre très facilement une erreur. Certaines questions évidentes, en apparence, étaient spécifiquement conçues pour induire délibérément les apprenants en erreur. Les apprenants étaient parfois tellement sûrs de l'exactitude de leur réponse que la découverte du contraire générait chez eux de l'étonnement voire même de la frustration et de la contrariété.

Le temps alloué pour répondre à chaque question était limité à 20 secondes. Cette contrainte temporelle a été ajoutée dans le but de provoquer un certain degré de stress ou d'excitation chez les apprenants durant cette phase de question-réponse.

Avant le début du test, les apprenants ont été informés qu'une réponse correcte leur rapporterait un point, alors qu'une réponse incorrecte ou une absence de réponse valaient zéro point. Ils étaient aussi informés que l'objectif était d'obtenir le plus de points possible dans les limites de temps imposées.

La première série de questions se rapportait à des connaissances générales en géographie. Un exemple de ces questions était : « *Est-ce que Rio De Janeiro est la capitale du Brésil ?* ». La deuxième série de questions portait sur la vérification orthographique. Les participants devaient déterminer si un mot présenté à l'écran était correctement orthographié ou pas. Les mots étaient présentés un par un à l'écran, et certains pouvaient contenir des erreurs d'orthographe communément commises, que l'apprenant peut ne pas détecter s'il n'est pas suffisamment attentif. Un exemple de ces mots mal orthographiés : les mots «Dilemne» ou «Supérieur». Enfin, la troisième série de questions nécessitait un certain niveau de raisonnement, les participants devaient répondre par vrai ou faux à des énoncés de logique, par exemple : « *si  $X < Y - 2$ , alors  $X < Y$*  ».

Après chaque réponse fournie, le système interagissait avec les apprenants pour leur indiquer si leurs réponses étaient correctes ou pas. Pour cela, des messages textuels comportant des connotations émotionnelles particulières étaient affichés. En cas de réponse correcte, le message était encourageant, par exemple : « *Excellente réponse ! Tu as l'air d'être bien concentré* ». Dans le cas contraire, un conseil ou un message d'empathie était affiché, par exemple : « *Mauvaise réponse ! Tu dois te concentrer davantage* » ou « *Je suis sûr que tu avais la bonne réponse* ». L'objectif derrière la mise en place de ces messages était d'accentuer les manifestations émotionnelles lors de l'interaction. Cette stratégie à base de messages a été employée dans plusieurs études sur les réactions émotionnelles physiologiques des apprenants lors de jeu-concours (p.ex. (Ochs et al., 2004; Prendinger et al., 2005b)).

### 3.4.2 Enregistrement des signaux EEG

Durant l'acquisition des données, les apprenants étaient équipés d'un bonnet EEG (de la marque « Electro-Cap ») ; les signaux ont été recueillis à partir de six régions actives. Les quatre premières régions sont situées sur le scalp aux emplacements *P3*, *C3*, *Pz* et *Fz*, tels que définis par la norme internationale 10-20 (Jasper, 1958), et référencés par rapport à *Cz*. Les deux dernières régions correspondent aux emplacements *A1* et *A2*, soient respectivement le lobe de l'oreille gauche et le lobe de l'oreille droite.

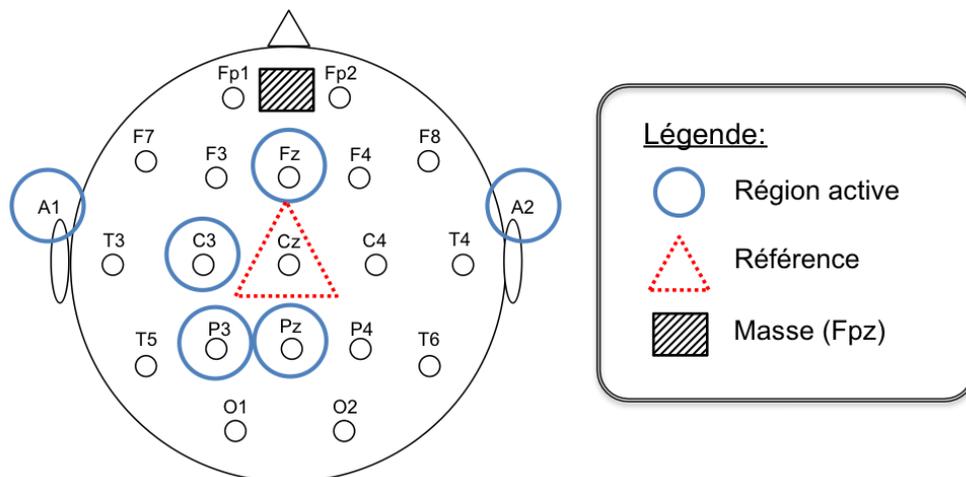


Figure 3.4 – Emplacements des canaux EEG

Cette configuration particulière — appelé également montage — est désignée techniquement par « *the referential linked ears montage* », illustrée dans la figure 3.4. En résumé, cette configuration permet de recentrer le signal par rapport aux deux hémisphères. Ce type de montage permet d'obtenir mathématiquement un signal EEG beaucoup plus précis et plus net en corrigeant le signal provenant de chaque emplacement du scalp par rapport à celui du milieu du cerveau. Par exemple, le signal *C3* corrigé sera donné de la façon suivante :

$$C3_{\text{corrigé}} = C3 - \frac{A1 + A2}{2}$$

L'impédance des électrodes a été maintenue en dessous de 5 Kilo Ohms durant toute la durée de l'expérimentation, en appliquant un gel non collant sur les emplacements mesurés. Ce seuil d'impédance est un indicateur de la qualité de contact entre l'électrode et le crâne. Une impédance réduite permet de diminuer le bruit et donc d'avoir une bonne qualité de signal (Reilly, 1999). Le taux d'échantillonnage pour l'enregistrement du signal EEG était de 256 Hz.

De plus, comme le signal électrique du cerveau est très faible (de l'ordre de microvolts :  $10^{-6}$  volts) et qu'il est très souvent bruité, le signal EEG a été amplifié et filtré. Les sources du bruit sont notamment l'électricité statique ou les champs électromagnétiques produits par les appareils environnants. Un filtre coupe-bande de 60 Hz a été appliqué pour éliminer ces interférences durant la phase d'acquisition des données. Il est important de préciser qu'un filtre coupe-bande de 50 Hz doit être utilisé en Europe, où la distribution de l'électricité (110 volts) diffère de celle de l'Amérique du Nord (120 volts).

Outre les bruits externes, le signal EEG peut être fortement contaminé par des artefacts qui proviennent des mouvements du corps ou des clignements très fréquents des yeux. Par conséquent, un filtre passe-haut de 48 Hz et un filtre passe-bas de 1 Hz ont été appliqués pour éliminer ces artefacts.

**Extraction de l'index d'engagement à partir des signaux EEG.** L'index d'engagement a été dérivé à partir des signaux EEG en appliquant une Transformée de Fourier Rapide (TFR). Les bandes de fréquences  $\theta$  (4-8 Hz),  $\alpha$  (8-13 Hz) et  $\beta$  (13-22 Hz) en provenance de chaque

région ont été extraites pour calculer un index d'engagement chaque seconde en utilisant la procédure décrite dans la section 3.3.2 de ce chapitre.

### 3.4.3 Détection des émotions à partir des senseurs physiologiques

La méthode utilisée pour la détection des états émotionnels des apprenants se base sur la transformation des données en provenance des senseurs physiologiques BVP et SC selon les deux indicateurs suivants :

- La fréquence cardiaque « Heart Rate » (HR) : mesure les variations issues des contractions des muscles du cœur. Cette mesure est calculée à partir de l'intervalle de temps entre les variations des données en provenance de la pression du volume sanguin. Plus précisément, il s'agit de mesurer l'intervalle, calculé en millisecondes, qui existe entre chaque deux battements successifs « interbeat interval » (IBI). Comme illustré par la figure 3.5, il s'agit de repérer les pics successifs dans les données en provenance du senseur BVP. La fréquence cardiaque est donnée par l'équation suivante :

$$HR = \frac{60000}{IBI}$$

Le numérateur désigne une minute exprimée en millisecondes. L'unité de la fréquence cardiaque est le battement par minute ou «beat per minute » (BPM).

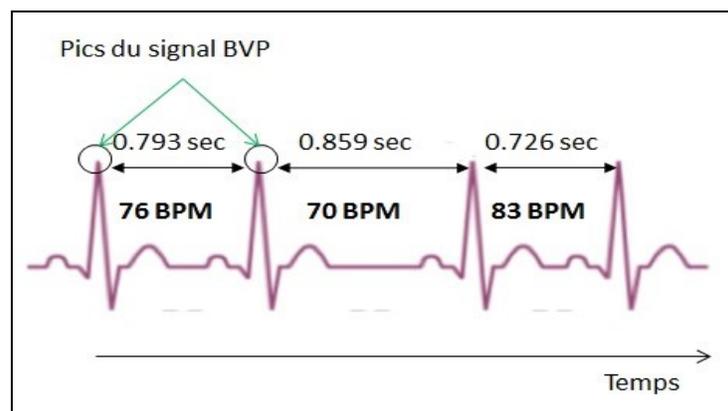


Figure 3.5 – Exemple d'un signal BVP

- La réponse (résistance) galvanique de la peau « Galvanic Skin Response (Resistance) » (GSR) : est une mesure du réflexe psycho-galvanique (ou réponse électrodermale) extraite à partir du senseur de conductivité de la peau (SC). Cet indicateur traduit l'activité des glandes sudoripares particulièrement présentes sur les mains et les pieds, responsables de la sécrétion de la sueur (Gil, 2009).

Les signaux physiologiques ont été enregistrés avec un taux d'échantillonnage de 1024 Hz.

Ces deux mesures sont très largement utilisées dans la détection et le suivi des états émotionnels (Healey, 2000; Jraidi et al., 2010c; Kim, 2007; Prendinger et al., 2005b; Shen et al., 2009). Le principe de cette approche consiste à utiliser ces indicateurs pour caractériser l'état émotionnel en terme de valence et d'activation, et à les projeter dans l'espace bidimensionnel du modèle d'émotions.

En effet, comme nous l'avons spécifié dans le chapitre 2, le rythme cardiaque est utilisé comme un indicateur de la valence de l'émotion. Cet indicateur fait apparaître trois phases lors d'une manifestation émotionnelle chez un individu, à savoir : (1) une décélération de l'activité (2) suivie d'une phase d'accélération (3) et enfin une seconde phase de décélération. Plusieurs études ont montré qu'une émotion négative implique une première décélération de l'activité cardiaque plus importante qu'une émotion positive, tandis qu'une émotion positive provoque une accélération plus importante qu'une émotion négative (Lang, 1995; Picard, 1997). Ainsi, plus la valence d'une émotion est positive, plus la valeur du HR augmente par rapport à une valeur de référence au repos. Le HR permet alors de situer l'état émotionnel par rapport à l'axe horizontal de la valence du modèle bidimensionnel des émotions.

Le GSR est utilisé pour déterminer l'activation ou l'intensité de l'émotion. Lorsque le système nerveux humain est stimulé émotionnellement, il active des glandes de sudation de la main et des pieds provoquant la sécrétion d'une substance saline. Ceci réduit la résistance électrique de la peau, et par conséquent, augmente la réponse galvanique mesurée. À titre d'exemple, plusieurs personnes ont les mains moites lorsqu'elles sont stressées ou tendues. Ainsi plus la mesure du GSR est forte par rapport à une valeur de référence, plus l'intensité

émotionnelle est grande, et vice versa. L'indice de résistance GSR permet de situer l'état émotionnel par rapport à l'axe vertical du modèle bidimensionnel des émotions. La Figure 3.6 illustre le principe général de fonctionnement de cette approche.

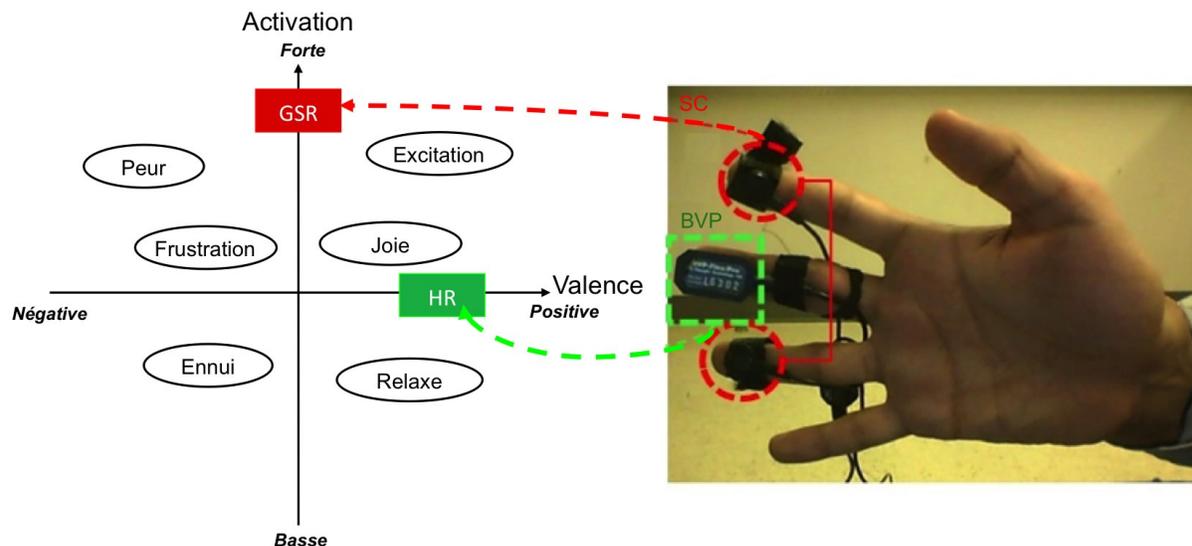


Figure 3.6 – Transposition des signaux physiologiques BVP et SC pour la détection des émotions.

Tout comme la plupart des travaux qui se basent sur l'approche dimensionnelle dans la modélisation des émotions, nous avons opté pour une division de cet espace en quatre sous-espaces émotionnels regroupant des classes d'émotions (Prendinger et al., 2005a; Strain et al., 2011; Yoo et al., 2005). Plus précisément, l'espace bidimensionnel des émotions est divisé selon la valence et l'activation en quatre régions appelées quadrants émotionnels (voir figure 3.7). Le quadrant émotionnel de l'apprenant est déterminé en normalisant et en comparant les variations des signaux HR et GSR par rapport aux valeurs de référence. Ainsi, selon les valeurs des indicateurs HR et GSR, les apprenants peuvent passer d'un quadrant à l'autre, selon leur état émotionnel. Les quadrants considérés dans nos analyses sont :

- Le quadrant Q1 : inclut des émotions positives à fortes intensités telles que la joie, la surprise et l'excitation.
- Le quadrant Q2 : inclut des émotions positives de faibles intensités. L'apprenant est dans un état calme et détendu.

- Le quadrant Q3 : contient des émotions négatives de fortes intensités. Celui-ci inclut des émotions comme la confusion, le stress, la frustration ou la colère.
- Le quadrant Q4 : regroupe des émotions négatives de basses intensités comme l'ennui par exemple.

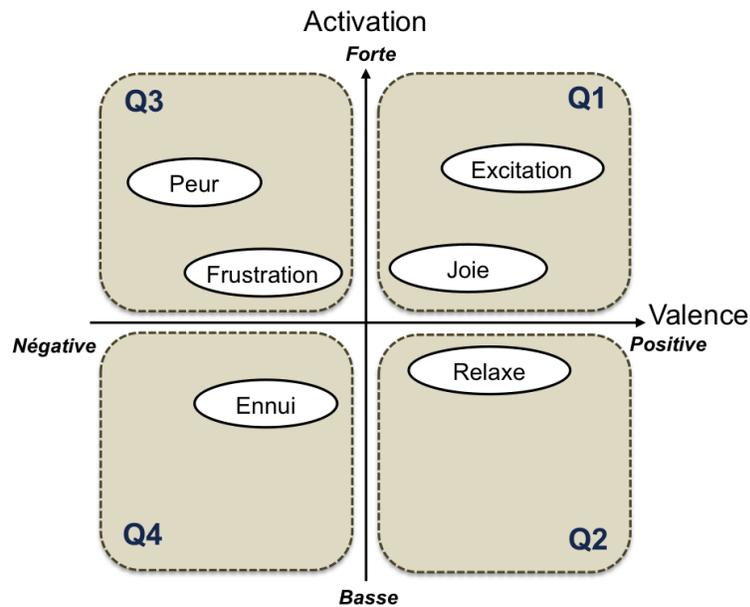


Figure 3.7 – Espace bidimensionnel valence/activation des émotions avec les 4 quadrants

### 3.5 Résultats expérimentaux

Durant nos expérimentations, chaque indice d'engagement mental calculé a été associé à l'un des quatre quadrants de l'espace bidimensionnel des émotions décrit précédemment, et ce pour tous les apprenants. Une analyse corrélacionnelle entre l'indice d'engagement et les performances des apprenants a également été réalisée. Les résultats obtenus sont présentés dans les quatre sous-sections suivantes.

### 3.5.1 Analyse du comportement de l'index d'engagement par rapport aux performances

Notre premier objectif a été de vérifier si l'index d'engagement peut donner une indication sur les performances des apprenants. Pour cela, nous avons analysé la relation entre le niveau de l'engagement mental et les performances dans la tâche mesurée par le score total obtenu. Deux groupes d'apprenants ont été considérés :

- (G1) Les apprenants dont le niveau d'engagement moyen était inférieur par rapport à leur valeur de référence (Baseline) sur toute la durée de l'expérience. Ces apprenants présentaient donc, selon l'interprétation de cet index, un niveau d'engagement globalement faible.
- (G2) Les apprenants, dont le niveau d'engagement était supérieur à leur valeur de référence sur toute la durée de l'expérience. Le niveau d'engagement pour ces apprenants était considéré comme globalement élevé.

Les résultats d'une analyse de variance (ANOVA) à un facteur ont montré que les performances dans le groupe G2 étaient significativement plus élevées que dans le groupe G1 :  $F(1, 33) = 19.782$ ,  $p < 0.001$ . Le tableau 3.2 montre la moyenne et l'écart type des performances réalisées dans les deux groupes.

Tableau 3.2 – Comparaison des performances entre les groupes G1 et G2

	<i>M (moyenne)</i>	<i>SD (écart type)</i>
G1 : mentalement moins engagés	16.00	3.120
G2 : mentalement engagés	19.78	2.025

Ce premier résultat a montré que l'index d'engagement peut donner une *indication qualitative* sur les *performances* des apprenants, surtout dans ce genre de tâche où la vigilance et l'attention jouent un rôle clé pour l'obtention de bonnes performances. En effet, maintenir un bon niveau d'engagement augmente la chance de discerner et d'éviter les différents pièges présents dans les questions. Les apprenants qui ont su maintenir un bon niveau d'engagement

mental tout au long de la session selon cet index ont pu réaliser un score significativement plus élevé.

### **3.5.2 Analyse de l'index d'engagement avec les indicateurs physiologiques émotionnels, le temps de réponse et les performances**

Étant donné qu'il existe un rapport entre la performance globale et le niveau d'engagement des apprenants, nous nous sommes intéressés à examiner la relation qui pourrait exister entre l'évolution de l'index EEG d'engagement, le comportement dans la tâche (en termes de temps émis pour répondre aux questions et exactitude des réponses) et les indicateurs physiologiques émotionnels.

Une régression linéaire a été réalisée avec comme variable dépendante la valeur moyenne de l'index d'engagement sur chaque question. Les quatre facteurs prédictifs introduits dans le modèle de régression sont : (1) la moyenne de la fréquence cardiaque (HR), (2) la moyenne de la réponse galvanique de la peau (GSR), (3) le temps de réponse (RT) à la question courante, et (4) le résultat réalisé par l'apprenant dans la question courante (codé par +1 s'il s'agit d'une bonne réponse et -1 dans le cas contraire).

Le modèle global était statistiquement significatif :  $F(4,947) = 121.45$ ,  $p < 0.01$  ;  $R^2 = 0.289$ . L'analyse conditionnelle de l'effet principal a montré un effet significatif du HR (beta = -0.001,  $p < 0.05$ ), du GSR (beta = .42,  $p < 0.05$ ) ainsi que du RT (beta = 0.21,  $p < 0.05$ ). En revanche, il n'y avait pas d'effet significatif de la réponse de l'apprenant (beta = 0.03,  $p = 0.46$ ).

Ce résultat démontre que même si, sur l'ensemble de la session, les apprenants qui étaient les mieux engagés mentalement avaient de meilleurs résultats, cet effet n'était pas observé pour chaque question prise à part. En effet, ce résultat — bien qu'il est contre-intuitif au premier abord — confirme plusieurs tendances qu'on a remarquées lors de l'analyse de ces différentes sessions.

En effet, le comportement de l'engagement mental des apprenants pouvait varier énormément selon le déroulement et l'enchaînement des réponses. Pour certains apprenants,

une série de mauvaises réponses consécutives réduisait considérablement leur niveau d'engagement mental. Une sorte de relâchement voire de lassitude apparaît et les apprenants semblaient de plus en plus désengagés par rapport à la tâche. Paradoxalement, ce même effet de décroissance de l'engagement mental a été observé, mais dans le cas d'une série consécutive de bonnes réponses. Après un enchaînement de bonnes réponses, les apprenants avaient tendance à se détendre, à être plus surs d'eux-mêmes et donc à faire moins attention dans leurs réponses, moins relire les questions, etc.

Ces constatations confirment la nature complexe de la relation qui existe entre l'engagement et les performances. Les systèmes doivent être capables de reconnaître les différentes tendances dans l'engagement mental des apprenants, et intervenir intelligemment soit pour maintenir un bon niveau d'engagement, ou pour le redresser (l'augmenter). La gestion de l'état mental d'engagement à travers cet index EEG semble ainsi une solution avantageuse pour les environnements d'apprentissage.

Par ailleurs, bien que la valeur de son coefficient bêta ne soit pas très grande, la relation entre la variable RT et l'index d'engagement était prévisible vu la nature de la tâche exécutée par les apprenants. En effet, en répondant aux questions, plus les apprenants prenaient leurs temps pour répondre (lecture et analyse de la question, réflexion, etc.), plus leurs index d'engagement étaient importants. Ce résultat démontre une certaine fiabilité de cet index dans la mesure où son évolution par rapport au temps de réponse était cohérente avec l'esprit de la conception de ce test. Un apprenant qui n'est pas attentif et qui se précipite dans sa réponse peut ne pas déceler les pièges existants dans les questions.

Le résultat de cette régression a également montré l'existence d'une relation entre les indicateurs physiologiques émotionnels (à travers le HR et le GSR) et l'engagement mental. Lors d'une réponse, l'engagement mental d'un apprenant peut alors s'accompagner de certaines manifestations émotionnelles exprimées par ces indicateurs émotionnels.

### 3.5.3 Corrélations de l'indice d'engagement avec les signaux GSR et HR

Afin de mieux comprendre la nature de la relation qui existe entre les indicateurs physiologiques émotionnels et l'index mental d'engagement, une corrélation bivariée a été calculée entre l'indice d'engagement et la moyenne des signaux HR et GSR pour tous les apprenants.

Les résultats ont montré l'existence d'une relation significative entre l'indice d'engagement et le signal GSR ( $r = 0.68$ ,  $p = 0.032$ ), et une relation non significative entre l'index d'engagement et le signal HR ( $r = -0.042$ ,  $p = 0.197$ ).

L'index d'engagement semble donc avoir une relation avec la dimension d'activation du modèle émotionnel. Plus l'engagement mental est élevé, plus l'intensité des émotions ressenties par les apprenants augmente. En revanche, il n'y avait pas de relation significative entre l'engagement et la dimension de valence du modèle.

En raisonnant en terme de quadrants émotionnels, on peut faire les constatations suivantes :

- Le passage du quadrant émotionnel Q1 au quadrant Q3 (p.ex. l'apprenant passe d'un état joyeux à un état confus) ou de Q2 à Q4 (p.ex. l'apprenant passe d'un état détendu à un état d'ennui), causé par un changement dans la valence (calculée à partir des signaux HR), a un effet insignifiant sur l'indice d'engagement.
- Le passage du quadrant émotionnel Q1 au quadrant Q2 (p.ex. l'apprenant passe d'un état joyeux à relaxe) ou de Q3 à Q4 (p.ex. l'apprenant passe d'un état confus à un état d'ennui), causé par un changement dans l'activation (calculée à partir des signaux GSR), est accompagné d'un changement significatif dans la tendance d'évolution de l'index d'engagement.

La figure 3.8 montre un exemple de l'augmentation visible de l'indice d'engagement d'un apprenant suite à une bonne réponse. Nous pouvons clairement observer qu'à la suite d'une bonne réponse, l'état émotionnel de l'apprenant a tout de suite changé du quadrant Q2

vers Q1 (par exemple de stressé à content) résultant en une augmentation de l'engagement mental.

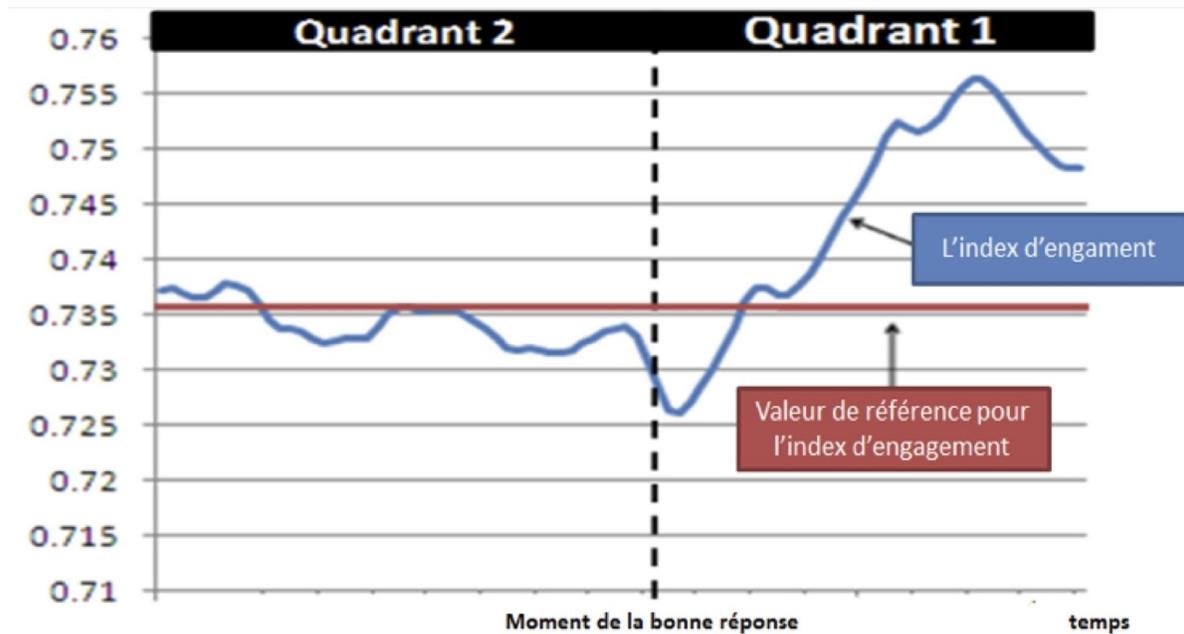


Figure 3.8 – Changement de l'indice EEG d'engagement mental à travers les quadrants émotionnels Q1 et Q2

### 3.5.4 Relation entre l'état émotionnel et l'index d'engagement

À travers le résultat précédent, nous avons pu observer que l'index d'engagement peut donner une indication sur les dimensions émotionnelles. Pour mieux analyser la relation entre ces deux variables, une ANOVA à un facteur a été effectuée avec comme variables indépendantes les quadrants émotionnels et comme variable dépendante l'index d'engagement. Les résultats ont montré qu'il y a un effet principal significatif de l'état émotionnel sur la valeur de l'indice :  $F(3, 6064) = 115.749, p < 0.05$  pour tous les participants.

Plus précisément, l'analyse de ce résultat a révélé que les valeurs moyennes de l'indice d'engagement étaient significativement plus élevées lorsque l'état émotionnel de l'apprenant était dans Q1 (valence positive et activation élevée :  $M = 0.769, SD = 0.085$ ) par rapport aux autres quadrants (voir figure 3.9). Ainsi, les émotions positives qui surviennent dans le

quadrant Q1 (comme la joie ou l'excitation) se sont accompagnées avec le plus haut niveau d'engagement comparé aux autres quadrants.

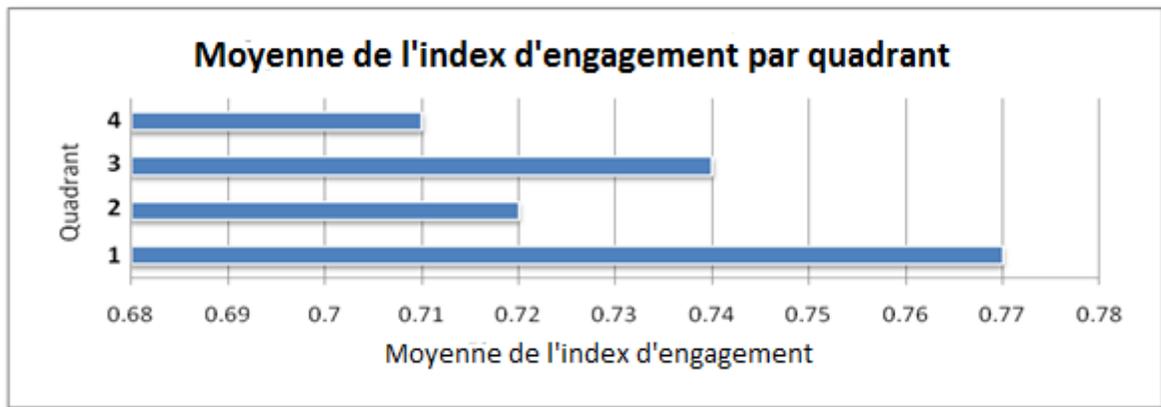


Figure 3.9 – Moyenne de l'indice d'engagement selon les quatre quadrants

Par ailleurs, la deuxième plus grande moyenne de l'indice d'engagement a été trouvée dans Q3 (valence négative et activation élevée :  $M = 0.743$ ,  $SD = 0.074$ ), soulignant ainsi que les émotions associées à cet état (p. ex. : la confusion ou la frustration) pouvaient également être synonymes d'un niveau élevé d'engagement chez les apprenants.

Pour mieux comprendre ce résultat, nous donnons l'exemple suivant qui était très récurrent au cours de cette expérience : l'apprenant répond à une question en étant confiant que sa réponse est correcte, mais le système indique qu'au contraire, la réponse était erronée, suscitant chez l'apprenant une émotion négative à forte intensité comme la confusion ou la frustration (quadrant Q3). Suite à cette erreur, le niveau d'engagement de l'apprenant peut spontanément augmenter, tout en restant dans le même état émotionnel, dans le but de ne pas se tromper une deuxième fois.

Enfin, les deux plus faibles moyennes de l'indice d'engagement ont été recensées dans les quadrants Q2 (valence positive et activation négative :  $M = 0.72$ ,  $SD = 0.086$ ) et Q4 (valence négative et activation négative :  $M = 0.712$ ,  $SD = 0.105$ ). Ce résultat montre que les émotions avec un faible niveau d'activation (ennui, relaxe, etc.) se manifestent avec un faible index d'engagement mental.

### **3.6 Discussion**

Nous avons choisi de faire notre première évaluation de cet index EEG d'engagement à travers une tâche d'apprentissage (jeu-concours) relativement facile et plaisante, mais qui nécessite un bon niveau d'attention et de vigilance pour pouvoir réaliser de bonnes performances.

L'analyse des résultats a montré que conformément à ce qui a été prévu lors de la conception du jeu-concours, les apprenants qui avaient maintenu un bon niveau d'engagement selon cet index (valeur moyenne supérieure à la valeur de référence) avaient obtenu les meilleures performances sur l'ensemble des questions. De même, l'index d'engagement a montré une variation linéaire avec le temps de réponse des apprenants. Ce comportement est également cohérent avec la conception du jeu-concours dans la mesure où dans ce genre d'activités, plus les apprenants prenaient leurs temps pour répondre, plus ils étaient susceptibles de faire attention aux éléments de la question, de réfléchir, de vérifier leurs réponses, etc.

Ces résultats viennent donc confirmer cet index à fournir des indications sur la manière avec laquelle les apprenants abordent les tâches demandées en matière d'engagement, d'attention et de vigilance. Ainsi l'utilisation de cet index dans un STEI permettra d'augmenter leurs capacités à suivre le processus d'apprentissage et à s'adapter en conséquence.

D'un autre côté, les analyses de cet index avec les quadrants émotionnels déduits à partir des indicateurs physiologiques (HR et GSR) ont montré que l'index d'engagement pouvait également renseigner sur l'état émotionnel de l'apprenant. En particulier, les valeurs moyennes de cet index, qui étaient associées aux différents quadrants émotionnels, étaient compatibles avec plusieurs résultats expérimentaux et théoriques précédents, qui se sont intéressés à étudier la relation entre l'engagement et les émotions. En effet, nous avons trouvé que les apprenants qui avaient globalement le meilleur niveau d'engagement dans l'activité, selon cet index, exprimaient des émotions positives à forte intensité (Q1). Ceci est en accord avec plusieurs résultats qui soutiennent la corrélation entre l'apparition des émotions positives (comme la joie) et l'augmentation de l'attention (Fredrickson, 2001; Fredrickson et al., 2005).

De plus, selon ce même index, les émotions négatives à forte intensité dans Q3 se sont également accompagnées d'un niveau d'engagement élevé. Ce résultat est aussi compatible avec plusieurs études qui insistent sur le rôle fondamental de telles émotions comme la confusion ou le stress (émotions présentes dans Q3) dans le processus d'apprentissage. En effet, ces émotions sont considérées, par plusieurs auteurs, comme des facteurs stimulateurs essentiels dans la construction des connaissances (Baker et al., 2010; Craig et al., 2004 ; D'Mello et al., 2014 ; Muse et al., 2003).

Le faible niveau d'engagement trouvé dans le quadrant Q2 et surtout Q4, qui enregistre le plus faible niveau d'engagement mental moyen chez les apprenants, est compatible avec la théorie de l'immersion mentale de Csikszentmihalyi (1990, 1997), selon laquelle il existe un état optimal d'interaction (appelé « flow »), de grande concentration, dans lequel l'apprenant est complètement absorbé par une tâche, pour laquelle il éprouve un très grand intérêt (Csikszentmihalyi, 1990; Csikszentmihalyi, 1997). Cet auteur affirme que la présence d'émotions comme l'ennui (émotion présente dans Q4) est un signe d'un manque d'immersion et d'engagement dans la tâche. Cette conclusion a été également reprise par plusieurs études expérimentales qui ont souligné l'effet négatif de l'ennui sur le niveau d'engagement des apprenants dans les tâches d'apprentissage (Baker et al., 2010; D'Mello et al., 2011). Selon D'Mello et Graesser (2011), l'état d'ennui est plus persistant dans le temps que les autres états ; conséquemment, il est plus difficile de réengager l'apprenant dans une tâche s'il ressent de l'ennui.

Ainsi, l'analyse de l'index d'engagement a montré sa capacité à fournir des indications pertinentes sur les manifestations émotionnelles qui accompagnent l'interaction des apprenants avec le système. Ceci renforce notre idée sur l'importance de l'intégration de cet index dans les STEI pour le suivi et la gestion de l'engagement mental des apprenants.

### **3.7 Conclusion**

Dans ce chapitre, nous nous sommes intéressés au premier indicateur cérébral que nous voulons intégrer dans les STEI à savoir l'index d'engagement mental. Cet index, qui est extrait à partir des signaux EEG des apprenants, a été développé dans la NASA pour évaluer le

niveau d'attention des pilotes dans le cadre de simulations de vols. Nous avons proposé, à travers une étude expérimentale, d'évaluer cet indicateur cérébral dans une situation d'apprentissage.

Les résultats de cette étude ont permis de confirmer la validité de cet index d'engagement et sa capacité à fournir des indications sur le comportement des apprenants au cours de leurs interactions avec un environnement d'apprentissage (de type jeu-concours). En effet, nous avons trouvé que cet index était relié aux performances des apprenants en termes de résultats réalisés et de temps de réponse. Notre étude a aussi démontré que l'index d'engagement permet de donner également des indications sur les états émotionnels des apprenants. En effet, nous avons trouvé que cet indicateur était relié aux réponses émotionnelles de l'apprenant mesurées en terme de rythme cardiaque et de réponse galvanique de la peau.

Dans le chapitre suivant, nous allons nous intéresser à un autre indicateur mental extrait à partir de l'activité cérébrale, et que nous proposons d'intégrer dans les STEI, à savoir : la charge mentale de travail.



## Chapitre 4

# Modélisation de la charge mentale de travail des apprenants

### 4.1 Introduction

Ce chapitre s'inscrit dans la continuité du premier objectif de cette thèse qui vise à augmenter les STEI par des indicateurs cérébraux dans le but d'améliorer leurs capacités à analyser et interagir en fonction de l'état de l'apprenant. En effet, dans le chapitre précédent, nous nous sommes intéressés à notre premier indicateur cérébral : l'index d'engagement ; nous avons pu démontrer que cet index fournit des indications pertinentes quant au déroulement du processus d'apprentissage ainsi qu'à l'état émotionnel de l'apprenant. Dans ce chapitre, nous nous intéressons au deuxième indicateur cérébral que nous proposons d'intégrer dans les STEI à savoir l'indicateur de *charge mentale de travail* (CMT).

Bien qu'elle ait un impact direct sur les capacités cognitives (Meshkati et al., 2011), et l'apprentissage en particulier (Paas et al., 2003), la notion de charge mentale de travail (appelée également charge cognitive « cognitive load ») n'a pas été suffisamment abordée dans les STEI. De plus, contrairement à l'engagement mental pour lequel nous nous sommes basés sur un index cérébral utilisé dans les neurosciences, il n'existe pas dans la littérature de

modèle préétabli pour la charge mentale de travail, que nous pourrions intégrer directement dans les STEI.

Ainsi, dans ce chapitre nous présentons notre approche pour modéliser la CMT des apprenants à partir de l'activité cérébrale. Nous nous intéresserons également à analyser conjointement l'indicateur cérébral de CMT avec l'index d'engagement et les émotions, afin d'évaluer l'impact de ces deux index dans le suivi de l'expérience d'apprentissage. En effet, l'objectif ultime de cette thèse sera de proposer un système adaptatif qui permettra de détecter en temps réel l'état mental de l'apprenant et d'ajuster ses stratégies d'interventions en fonction de cet état.

Avant cela, nous présentons dans ce chapitre la deuxième étude expérimentale que nous avons réalisée pour modéliser la CMT dans un contexte d'apprentissage et évaluer la relation entre cet index, l'index d'engagement et les émotions. Le but de cette étude est tout d'abord (1) d'entraîner un modèle de CMT, de le valider et d'évaluer son impact sur le déroulement de la session d'apprentissage. Nous allons ensuite (2) analyser, à travers cette étude, l'engagement mental de l'apprenant. Ceci nous permettra d'évaluer le comportement de cet index dans un contexte d'apprentissage différent de celui que nous avons étudié dans le chapitre précédent. En effet, l'environnement d'apprentissage que nous allons utiliser pour cette deuxième étude est beaucoup plus complexe que l'environnement que nous avons utilisé pour la première expérience : les apprenants vont être amenés à suivre un cours portant sur la trigonométrie, suivi d'une session de résolution de problèmes. Ceci nous permettra ainsi de valider la pertinence de cet index. Nous allons enfin (3) examiner la relation entre les deux index cérébraux (engagement et CMT) et les émotions de l'apprenant, afin de vérifier que ces index sont associés aux émotions.

Ce chapitre est organisé de la façon suivante. Dans la deuxième section, nous présentons la notion de charge mentale de travail. Dans la troisième section, nous décrivons notre protocole expérimental. La quatrième section sera consacrée à notre approche pour modéliser la charge mentale de travail, et la cinquième section présentera les résultats de notre étude.

## 4.2 La charge Mentale de travail

La charge mentale de travail est définie d'une manière générale comme étant la *quantité d'effort mental* mise en œuvre par un individu pour faire face aux exigences d'une *tâche* (Barrouillet, 1996 ; Tricot et al., 1996). Par effort mental, on sous-entend le degré de mobilisation des *ressources cognitives* telles que le traitement des informations, la prise de décision, la mémorisation à court et moyen terme, etc. (Parasuraman et al., 2002; Vidulich et al., 2012).

Dans le contexte d'apprentissage, la charge mentale de travail représente la quantité d'effort mental investi par un apprenant lors de l'acquisition de nouvelles connaissances ou compétences, ou encore au cours de situations nécessitant l'application de connaissances ou de compétences (Paas et al., 2003). La charge mentale de travail est conjointement dépendante des deux caractéristiques suivantes :

- La tâche d'apprentissage : en effet, selon la théorie de la charge cognitive (« cognitive load theory ») qui s'est développée avec les travaux de Sweller (1988) portant sur la relation entre la charge cognitive et l'apprentissage, la nature et la conception de la tâche ainsi que ses exigences en termes de performance ont une grande incidence sur le niveau de charge mentale de travail des apprenants (Sweller, 1994, 2003; Sweller et al., 2011). Par exemple, une activité d'apprentissage mal conçue (contenant des informations superflues et mal structurées, des énoncés non clairs, etc.) augmente considérablement le niveau de charge mentale de travail des apprenants.
- Les apprenants : le niveau de charge mentale de travail des apprenants peut dépendre de plusieurs facteurs internes qui varient dans le temps et selon les individus. Parmi les facteurs propres à l'apprenant, on peut citer : ses ressources cognitives (c'est-à-dire les ressources disponibles à un instant donné), les capacités cognitives (c'est-à-dire les ressources maximales qu'il peut mobiliser), son niveau de compétence (appelé également puissance cognitive), sa motivation, etc. (Cain, 2007; Hancock et al., 1986; Lee et al., 2006; Meshkati et al., 2011; Welford, 1978).

Malgré sa complexité, la notion de charge mentale de travail a toujours suscité l'intérêt des scientifiques en particulier dans le domaine de l'apprentissage (Lee et al., 2006). L'une des raisons principales de cet intérêt est que les capacités cognitives et plus particulièrement la mémoire de travail (c'est-à-dire la mémoire à court et à moyen terme) disposent d'une *capacité limitée*. Ainsi, l'apprenant est souvent confronté à des situations d'apprentissage dans lesquelles la quantité d'informations à traiter dépasse ses capacités cognitives. On parle alors d'un état de *surcharge cognitive* (« mental overload »), appelé également saturation cognitive (Kirsh, 2000; Mayer et al., 2003). Cet état affecte énormément le processus d'apprentissage et peut carrément le paralyser (Eppler et al., 2004; Hung, 2012). En effet, lorsqu'un apprenant commence à être surchargé, ses performances se dégradent rapidement et ses erreurs s'accroissent. Il commence également à éprouver des difficultés à comprendre la matière, à bien raisonner et à organiser ses connaissances. Il devient également moins créatif, plus anxieux et stressé et fait de plus en plus de mauvais choix.

L'état de *sous-charge cognitive* (« mental underload ») est également considéré comme un état négatif pour le processus d'apprentissage (Paas et al., 2004; Paas et al., 2003). En effet, la sous-charge cognitive désigne un faible niveau d'effort mental fourni par l'apprenant comparé à ce qui est demandé ou prévu par la tâche. De la même manière que la surcharge cognitive, ce manque d'investissement dégrade les performances des apprenants et augmente le nombre de leurs erreurs et de fausses conceptions (Paas et al., 2004; Young et al., 2002). Par exemple, un apprenant qui ne fait pas l'effort nécessaire pour lire, comprendre et assimiler tous les éléments d'un problème risque de mal le résoudre.

Ainsi, la modélisation de la charge mentale de travail dans les environnements d'apprentissage permettra de mieux adapter leurs interactions et d'éviter l'occurrence de la surcharge ou de la sous-charge cognitive. Le déroulement du processus d'apprentissage doit être adapté intelligemment par les systèmes de façon à mettre l'apprenant dans une zone de confort cognitif située entre la surcharge et la sous-charge (Hoc et al., 2003; Paas et al., 2004).

### 4.2.1 Modélisation de la charge mentale de travail

La modélisation de la charge mentale de travail dans l'Interaction Homme-Machine utilise principalement trois catégories de mesures (Cain, 2007) :

1. Les questionnaires d'auto-évaluation : appelés également mesures subjectives, ces questionnaires psychologiques permettent de déterminer la charge mentale de travail telle qu'elle est perçue par l'apprenant. Parmi les questionnaires les plus utilisés, on peut citer la technique d'évaluation subjective de la charge mentale de travail (« Subjective Workload Assessment Technique ») (SWAT) (Reid et al., 1988) ou encore l'instrument NASA TLX (Hart et al., 1988a) que nous présenterons plus tard dans ce chapitre.
2. Les mesures de performances : appelées également mesures objectives, elles sont extraites à partir de la manière avec laquelle les individus exécutent la tâche. Ces mesures incluent notamment la performance (ou l'efficacité), le taux d'erreur, le temps d'exécution, le taux de réussite par rapport au taux d'échec, etc.
3. Les mesures physiologiques : parmi les indicateurs physiologiques utilisés pour mesurer la charge mentale de travail, on peut citer le mouvement des yeux, le rythme cardiaque, la réponse galvanique de la peau et l'activité cérébrale.

Nonobstant l'utilisation de ces indicateurs dans la modélisation de la charge mentale de travail dans divers contextes, plusieurs limites sont à signaler (Cain, 2007; Castor et al., 2003; Just et al., 2003). En effet, les questionnaires d'auto-évaluation présentent l'inconvénient de devoir interrompre l'apprenant ce qui peut le gêner et même augmenter sa charge mentale de travail (Speier et al., 1999). Les mesures de performances sont souvent dépendantes de la nature de la tâche et manquent de précision (Cain, 2007). Certains indicateurs physiologiques, malgré leurs capacités à fournir une estimation continue et spontanée, ne sont pas souvent utilisables dans tous les types de tâches (Castor et al., 2003; Sirevaag et al., 2000). Par exemple, l'analyse des certains paramètres des mouvements des yeux (tels que la dilatation du diamètre de la pupille, la durée de fixation, la vitesse des mouvements de l'œil) peuvent renseigner sur l'augmentation de la charge mentale de travail. Toutefois, les variations de ces paramètres sont généralement très légères et difficilement détectables dans un contexte où

l'individu peut avoir un certain degré de liberté comme celui de l'apprentissage. De même, la fréquence cardiaque (plus précisément la plage de fréquence cardiaque allant de 0.07 à 0.14 Hz) et les réponses galvaniques de la peau, souvent utilisées pour modéliser la charge mentale de travail, manquent de fiabilité et de précision (Castor et al., 2003; Mulder et al., 2000). En effet, les variations de ces indicateurs sont fréquemment biaisées par des manifestations émotionnelles qui ne sont pas toujours le résultat d'un certain effort mental produit par l'individu (Mulder et al., 2000).

Contrairement à ces indicateurs, l'utilisation des données de l'activité cérébrale dans la modélisation de la charge mentale de travail est le moyen le plus précis et le plus directement relié à cette notion (Anderson et al., 2011; Berka et al., 2005; Just et al., 2003; Stipacek et al., 2003). En effet, l'activité cérébrale a été fréquemment employée par les scientifiques et les concepteurs des systèmes physio-cognitifs<sup>1</sup> pour extraire des indicateurs relatifs au niveau de charge mentale de travail, et ce malgré l'absence d'un index standard directement extrait des signaux EEG (tel que l'index d'engagement mental que nous avons utilisé dans le chapitre précédent) (Berka et al., 2004; Berka et al., 2007; Berka et al., 2005; Hankins et al., 1998; Russell et al., 2005; Stevens et al., 2006; Stevens et al., 2007a; Stevens et al., 2007b; Wilson, 2002). Pour cela, ces travaux ont recours généralement à des algorithmes d'apprentissage machine de classification ou de régression (linéaire et non linéaire), qui extraient à partir des caractéristiques du signal EEG (principalement les plages de fréquences) un indicateur de charge mentale de travail (Berka et al., 2007; Galán et al., 2012; Gevins et al., 2003; Heger et al., 2010; Kohlmorgen et al., 2007; Wilson, 2002). Par exemple, Wilson (2002) a utilisé les réseaux de neurones afin d'entraîner un modèle capable de classer la charge mentale de travail en trois classes (fort, moyen et faible) à partir des valeurs des plages de fréquences  $\theta$  (4-8 Hz),  $\alpha$  (8-13 Hz) et  $\beta$  (13-22 Hz). Les réseaux de neurones ont été également employés par Gevin et Smith (2003) pour extraire un index permettant de suivre la charge mentale de travail des utilisateurs lors de la réalisation de différentes activités (tels que la correction de texte, la

---

<sup>1</sup> Les systèmes qui utilisent les principes de l'informatique physiologique (c'est-à-dire le recours à des senseurs physiologiques pour analyser et suivre l'état cognitif des utilisateurs) sont communément qualifiés de systèmes physio-cognitifs (Berka et al., 2007).

recherche rapide d'informations sur le Web et la mémorisation). D'autres algorithmes tels que l'analyse discriminante linéaire ou les machines à vecteur de support ont été également employés pour modéliser la charge mentale de travail à partir des données EEG (Berka et al., 2007; Galán et al., 2012; Heger et al., 2010; Kohlmorgen et al., 2007).

Dans cette thèse nous allons également utiliser une approche basée sur des techniques d'apprentissage machine, et plus particulièrement la *régression par processus gaussiens*, pour modéliser la charge mentale de travail à partir des données EEG. Notre approche sera validée dans un contexte d'apprentissage.

### **4.3 Expérience 2 : Modélisation de la charge mentale de travail dans un contexte d'apprentissage**

L'objectif de cette étude est de modéliser la charge mentale de travail à partir de l'activité cérébrale et de la valider dans un environnement éducatif. En effet, comme mentionné précédemment, on ne trouve pas dans la littérature un modèle préétabli, que nous pourrions directement utiliser pour notre approche ; contrairement à l'indice d'engagement cérébral, qui peut être extrait directement à partir des plages de fréquence EEG  $\theta$ ,  $\alpha$  et  $\beta$  (voir chapitre 2 section 3.3.2).

La CMT requiert l'entraînement d'un modèle à partir de caractéristiques extraites des signaux EEG. Cette phase d'entraînement recourt à des tâches cognitives faisant intervenir des fonctions mentales spécifiques assez simples (telles que la mémorisation, le calcul mental ou la logique). De plus lors de la réalisation de ces tâches, un protocole expérimental strict doit être imposé pour limiter les actions d'interactions de l'apprenant à des manœuvres bien précises, par exemple : entrer (au clavier) une séquence de chiffres mémorisés, le résultat d'un calcul mental ou d'un exercice de logique. Les données EEG rassemblées lors de cette phase seront utilisées comme des entrées (inputs) pour construire notre modèle prédictif de CMT. Pour ce faire, nous allons combiner différentes techniques d'extraction d'attributs, de réduction de dimensionnalité, de normalisation et de régression que nous détaillerons plus loin

dans ce chapitre. Le résultat (output) de notre modèle sera sous la forme d'un index EEG de charge mentale de travail (EEG\_CMT).

En plus des mesures physiologiques EEG, nous allons également mesurer dans cette étude des mesures subjectives (NASA\_TLX (Hart et al., 1988a)) et objectives (performance et temps réponse) de la charge mentale de travail. Ces mesures seront utilisées pour l'entraînement et l'évaluation de notre modèle prédictif. Plus spécifiquement, ce modèle sera entraîné pour prédire une estimation de la charge mentale de travail : l'index EEG\_CMT, à partir des évaluations subjectives attribuées par les apprenants. L'évaluation de ce modèle se fera en utilisant à la fois les mesures subjectives et objectives.

Ainsi notre méthodologie consiste, dans un premier temps, à recueillir un ensemble de données EEG : *phase d'entraînement cérébral*. Durant cette phase, l'apprenant sera amené à exécuter des tâches purement cognitives. En particulier, trois différentes tâches ont été implémentées pour cette étude : (1) Empan Numérique (EN) (2) Empan Numérique Inversé (ENI) et (3) Test de Logique et de Calcul Mental (TLCM). Les deux premières tâches incluent des capacités de mémorisation et la troisième des capacités de calcul, de raisonnement et d'inférence. Chacune de ces tâches a été conçue avec un niveau de difficulté progressivement croissant. L'objectif de cette variation croissante de difficulté est de solliciter graduellement un niveau de charge mentale de travail de plus en plus élevé chez les apprenants. L'index EEG\_CMT dérivé par notre modèle sera d'abord testé durant cette phase d'entraînement cérébral.

Notre méthodologie inclut, dans un deuxième temps, une *phase d'apprentissage* durant laquelle les apprenants seront amenés à utiliser un environnement éducatif portant sur la trigonométrie. Cet environnement, spécifiquement implémenté pour cette étude, comporte trois étapes successives : (1) un pré-test pour situer le niveau de connaissances des apprenants par rapport à la trigonométrie, (2) un cours sous la forme d'une leçon sur certaines notions de base en trigonométrie et (3) une phase de résolution de problèmes pour évaluer le niveau de connaissances des apprenants après la leçon. Durant cette session d'apprentissage, nous allons tout d'abord valider notre modèle de charge mentale de travail en analysant le comportement de l'index EEG\_CMT dérivé, dans un contexte concret d'apprentissage. Nous vérifierons,

entre autres, si cet index permet de refléter l'évolution de leur expérience d'apprentissage. En effet, contrairement à la phase d'entraînement cérébrale précédente qui impliquait des tâches strictes, contrôlées et purement cognitives, cette deuxième session d'interaction va faire intervenir une combinaison de différentes fonctions mentales plus complexes telles que la prise de décision, le raisonnement, la déduction, la généralisation, l'inférence et la construction de schémas mentaux. De plus, l'apprenant va disposer lors de son interaction, d'une certaine liberté d'action. Par exemple durant la phase (3) de résolution des problèmes, l'apprenant peut consulter le cours, utiliser une calculatrice, demander une astuce, etc. Les questions que l'on se pose alors sont les suivantes : est-ce qu'un modèle entraîné dans des conditions « laboratoire » strictes, avec des tâches assez simples et précisément conçues pour solliciter différents niveaux de charges mentales, est valide dans un environnement d'apprentissage moins contrôlé et plus complexe ? Est-ce que ce modèle reflète l'évolution de l'activité d'apprentissage ?

Nous allons par ailleurs analyser durant cette phase d'apprentissage l'index d'engagement mental que nous avons introduit dans le chapitre précédent. Ceci nous permettra d'évaluer la pertinence de cet index dans un nouveau contexte d'apprentissage. Nous nous intéresserons également à analyser conjointement nos index cérébraux (engagement et CMT) par rapport aux émotions, afin de vérifier si la combinaison de ces deux index permet de donner des indications pertinentes quant aux émotions de l'apprenant.

### **4.3.1 Protocole expérimental**

Le dispositif expérimental est constitué d'un bonnet EEG mesurant huit canaux cérébraux ainsi que de deux flux vidéo. Ces derniers ont été utilisés pour enregistrer les sessions d'expérimentations avec les apprenants, que nous avons visionnées et analysées pour extraire et synchroniser les données EEG avec les différentes phases de l'expérience.

17 participants en tout ont été recrutés pour cette étude. Tous les participants ont été informés du processus ainsi que des objectifs expérimentaux, et ont signé un formulaire de consentement. Une compensation de 20 dollars a été attribuée pour chaque participant.

À leur arrivée, les participants étaient équipés avec un bonnet EEG et familiarisés avec le dispositif et l'environnement de nos expérimentations. Ils étaient ensuite amenés à compléter les deux phases de notre étude, à savoir : la phase d'entraînement cérébral et la phase d'apprentissage

#### **4.3.1.1 Phase d'entraînement cérébral**

La phase d'entraînement cérébral se compose de trois tâches successives : (1) Empan Numérique (EN) (2) Empan Numérique Inversé (ENI) et (3) Test de Logique et de Calcul Mental (TLCM).

Comme décrit précédemment, l'objectif de cette phase est d'induire différents niveaux de charge mentale de travail chez les apprenants. Les trois tâches cognitives que nous avons considérées pour cette phase sont très fréquemment utilisées pour étudier et modéliser la charge mentale de travail (Berka et al., 2007). En effet, ces tâches permettent de solliciter, directement, sans effort physique supplémentaire et sur un court laps de temps, les capacités mentales d'un individu.

Chacune des trois tâches cognitives que nous avons considérées (EN, ENI et TLCM) dispose de différents niveaux de difficulté, avec un niveau progressivement croissant. Cet accroissement graduel de difficulté a pour effet d'agir sur le niveau d'effort mental requis pour l'exécution de chaque tâche, et donc de constamment augmenter le niveau de charge mentale de travail.

Avant de commencer chaque tâche, les participants étaient amenés à effectuer un ensemble d'exercices d'échauffement pour se familiariser avec le rythme imposé par la tâche. Par ailleurs, tous les participants ont effectué les tâches de cette phase dans le même ordre.

**Empan Numérique (EN).** Ce test sollicite les capacités d'attention et de mémorisation. Dans cette tâche, une série de chiffres sont successivement présentés à l'écran. Les participants doivent mémoriser toute la séquence, puis entrer les chiffres dans le bon ordre (ordre de leur apparition à l'écran). Cette tâche comprend 6 niveaux de difficulté graduellement croissante, avec chacun un certain nombre de chiffres dans la séquence à retenir. L'augmentation de la

difficulté se fait en augmentant la longueur de cette séquence. Le premier niveau comprenait une série de 20 séquences de 3 chiffres, le niveau deux : 12 séquences de 4 chiffres, le niveau trois : 8 séquences de 5 chiffres, le niveau quatre : 6 séquences de 6 chiffres, le niveau cinq : 4 séquences de 7 chiffres et le niveau six : 4 séquences de 8 chiffres.

**Empan Numérique Inversé (ENI).** Ce test se base sur le même principe que la tâche précédente (EN). Les participants doivent mémoriser une séquence de chiffres présentés à l'écran, sauf qu'ils doivent pour ce test entrer les chiffres dans l'ordre inverse de leur apparition. Cinq niveaux de difficulté ont été considérés en augmentant le nombre de chiffres de la séquence présentée. Le premier niveau consiste en une série de 12 séquences de 3 chiffres, le deuxième niveau : 12 séquences de 4 chiffres, le troisième niveau : 8 séquences de 5 chiffres, le quatrième niveau : 6 séquences de 6 chiffres et le cinquième niveau : 4 séquences de 8 chiffres. Aucune limite de temps n'a été imposée aux participants pour exécuter les tâches ENI et EN.

**Test de Logique et de Calcul mental (TLCM).** Deux sortes d'exercices ont été utilisés pour ce test : (1) additionner ou soustraire mentalement deux nombres différents ayant le même nombre de chiffres, et (2) déduire un élément manquant dans une série de nombres affichés à l'écran, en inférant la règle logique qui les unissait. Un exemple d'une telle série est : «  $38 - 2 - 19 - 9 - 3 - 3 - 40 - 4 - ?$  », où « ? » est l'élément manquant à trouver, et qui serait dans ce cas : «  $10$  ». Autrement dit, la règle logique que l'on doit déduire serait que pour chaque groupe de trois nombres, le dernier nombre est le résultat de la division du premier par le second.

Le test (TLCM) comprend trois niveaux de difficulté dont chacun consistait en une série de 5 questions. La manipulation du degré de difficulté de chaque niveau se fait en augmentant les nombres de chiffres des nombres à calculer pour les questions de calcul mental, et en augmentant la complexité de la règle logique à trouver pour les questions de logique. Les participants disposaient d'un délai maximal fixe de 30 secondes pour répondre aux questions de logique, et de 15 secondes pour les exercices de calcul mental.

Après avoir terminé la phase d'entraînement cérébral, les participants prenaient une pause de dix minutes environ. Puis, ils étaient amenés à compléter la deuxième phase de nos expérimentations, à savoir : la phase d'apprentissage.

Determine the measure of angle  $\widehat{DOC}$  in the trapezoid below

81      86      90       Skip

You can save the results

Intermediate result 1

Intermediate result 2

Intermediate result 3

Calculator      Help      Next

*Consider using the tangent (opposite over the adjacent).  
Remember that the sum of the angles of a triangle is 180*

**103**  
Time remaining

Figure 4.1 – Interface de l'environnement de trigonométrie utilisé dans la deuxième expérience<sup>2</sup>

#### 4.3.1.2 Phase d'apprentissage

Cette phase se compose de trois étapes successives : (1) un pré-test, (2) une leçon et (3) une session de résolution de problèmes.

<sup>2</sup> Cet environnement a également été utilisé dans le cadre des travaux de recherche de Jraidi (2014) pour la modélisation des émotions de l'apprenant.

**Pré-test.** Le pré-test est composé de 10 questions élémentaires de type vrai/faux, couvrant certains aspects fondamentaux de la trigonométrie. Un exemple de question posée est : « *Est-ce que la tangente d'un angle est égale à la longueur du côté opposé divisée par la longueur du côté adjacent ?* ». Les participants devaient répondre aux questions du pré-test sans aucune interruption, aide ou limite de temps. L'objectif de cette phase préliminaire était de situer le niveau de connaissances des participants par rapport à la trigonométrie.

**Leçon de trigonométrie.** Durant cette étape, les participants devaient utiliser un environnement éducatif pour apprendre (ou réviser) certaines notions en trigonométrie ; des propriétés et relations trigonométriques fondamentales sont présentées. L'environnement fournit des définitions et des concepts de base *ainsi* que leurs démonstrations mathématiques illustrées par des schémas explicatifs. En outre, pour chaque concept présenté, l'environnement fournit un ensemble d'exemples afin d'aider les apprenants à mieux le maîtriser.<sup>3</sup>

**Session de résolution de problèmes.** Durant cette étape, les apprenants devaient résoudre une série de problèmes trigonométriques relatifs aux concepts exposés dans la leçon suivie précédemment. Ces problèmes nécessitent une capacité d'appliquer, de généraliser et de raisonner sur ces concepts. Aucun autre prérequis, mis à part les éléments de la leçon, n'était nécessaire pour pouvoir résoudre les problèmes. Toutefois, un bon niveau de raisonnement et de concentration était requis pour trouver la solution. Six problèmes en tout, avec un niveau de difficulté graduellement croissant, ont été sélectionnés et présentés dans le même ordre pour tous les participants. Chaque problème était sous forme d'une question à choix multiples, illustrée par une figure géométrique. Un délai était imposé pour répondre à chaque problème. Ce délai variait d'un problème à l'autre, en fonction de son niveau de difficulté. En particulier, à chaque problème est associé un certain nombre d'étapes intermédiaires pour pouvoir atteindre la solution finale ; le niveau de difficulté a été augmenté en décuplant le nombre de ces étapes intermédiaires. Par exemple, pour calculer le sinus d'un angle, les apprenants

---

<sup>3</sup> À la fin de l'expérience, tous les participants ont indiqué qu'ils étaient satisfaits de la qualité de l'environnement ainsi que de l'approche pédagogique utilisée pour présenter le matériel didactique.

devaient d'abord calculer le cosinus dans une première étape. Ensuite, ils devaient l'élever au carré, et le soustraire de  $1$  dans une deuxième étape. Enfin, ils devaient dans une troisième étape calculer sa racine carrée. Tous les problèmes étaient indépendants en termes de concepts trigonométriques auxquels ils se rapportaient, à l'exception des problèmes 4 et 6 qui portaient sur la même règle géométrique, soit : « *la somme des angles d'un triangle est égale à 180 degrés* ».

L'environnement de résolution de problèmes fournissait aux apprenants un certain nombre d'indices pour chaque problème, ainsi qu'une calculatrice pour effectuer les différents calculs nécessaires. La figure 4.1 illustre une interface de notre environnement.

### **4.3.2 Enregistrement des signaux physiologiques EEG**

Pour enregistrer les données physiologiques EEG, nous avons utilisé les mêmes configurations de notre première étude décrite dans le chapitre précédent à savoir :

- Enregistrement des signaux EEG bruts à l'aide d'un bonnet électroencéphalographique.
- Mesure des signaux à partir des emplacements  $P3$ ,  $C3$ ,  $Pz$  et  $Fz$ , avec  $Cz$  comme référence et  $Fpz$  comme masse, en utilisant la norme internationale 10-20 de positionnement des électrodes sur le scalp (Jasper, 1958).
- Utilisation d'un montage référentiel par rapport aux lobes des deux oreilles avec les régions  $A1$  et  $A2$  comme régions actives.
- Maintien de l'impédance en dessous de 5 Kilo Ohms tout au long de l'expérience. Pour ce faire, une petite quantité de gel non collant était introduite à l'intérieur des électrodes positionnées au-dessus des emplacements mesurés.
- Application d'un filtre coupe-bande de 60 Hz lors de l'acquisition des données pour éliminer le bruit et les interférences causés par les dispositifs électroniques ambiants.
- Utilisation d'un taux d'échantillonnage de 256 Hz.
- Traitement du bruit après l'enregistrement des signaux de deux différentes manières : (1) en examinant manuellement les signaux et éliminant les parties contaminées par les

clignements des yeux et les mouvements musculaires ; (2) à travers l'exécution d'une procédure automatique qui élimine les parties du signal qui dépassent de plus de 25 % un seuil préalablement fixé. Si un bruit est détecté à la seconde  $t$  dans une portion du signal à un emplacement particulier, alors cette partie contaminée est également éliminée des signaux des autres emplacements.

### **4.3.3 Mesures subjectives et objectives de la charge mentale de travail**

Afin de pouvoir entraîner et évaluer notre modèle, nous avons considéré deux types de mesure de la charge mentale de travail : (1) une *mesure subjective* à travers l'utilisation de l'instrument « NASA-Task Load Index » (NASA-TLX) (Hart et al., 1988b), que nous allons décrire dans la sous-section suivante, et (2) deux *mesures objectives* : la performance et le temps de réponse des apprenants. En effet comme mentionné précédemment, la charge mentale de travail a un impact direct sur ces mesures (Cain, 2007), et peut ainsi être évaluée de manière objective à travers ces variables.

La mesure de performance a été calculée en termes de proportions de bonnes réponses lors de l'exécution des différentes tâches, à savoir : proportions des séquences correctement mémorisées pour les tâches EN et ENI, et proportions des réponses correctes dans le test TLCM, le pré-test et la session de résolution de problèmes. Le temps de réponse correspond au temps émis par l'apprenant pour l'exécution de chaque tâche.

#### **Mesure subjective de la CMT : NASA\_TLX**

Sous la forme d'un questionnaire d'auto-évaluation (voir Annexe A), cet instrument est utilisé pour évaluer le niveau de charge mentale de travail d'un individu lors de l'exécution d'une tâche. Il s'agit donc de recueillir une estimation subjective (personnelle) d'un individu par rapport aux efforts qu'il a produits lors de l'exécution de cette tâche. Comparé à d'autres mesures subjectives telles que le « Subjective Workload Assessment Technique » (SWAT) (Reid et al., 1988) ou encore le « Workload Profile » (WP) (Tsang et al., 1996), l'instrument NASA\_TLX a montré une plus grande validité ainsi qu'une meilleure sensibilité pour discriminer entre de fines variations de CMT (Rubio et al., 2004).

L'évaluation de la charge mentale de travail avec le NASA\_TLX se fait selon six principales dimensions : l'exigence mentale, l'exigence physique, l'exigence temporelle, la performance, l'effort et le niveau de frustration. Chaque dimension est présentée sous la forme d'une échelle, et le participant doit attribuer un score entre zéro et cent à chacune d'entre elles. Chaque échelle est divisée en 20 graduations (5 unités par intervalle de graduation).

Dans notre étude, les participants procédaient à cette évaluation immédiatement après avoir terminé un niveau de chaque tâche, autant dans la phase d'entraînement cérébral que dans la phase d'apprentissage.

Par ailleurs, étant donné que chaque tâche peut solliciter certaines dimensions de façon plus importante par rapport à d'autres, l'instrument NASA\_TLX prévoit une procédure supplémentaire de pondération de ces échelles. Il s'agit d'une séquence de comparaison par paires (« pair-wised comparison ») des six dimensions du test. À chaque comparaison, l'apprenant doit décider laquelle des deux dimensions présentées est la plus importante dans la tâche. Ainsi, un total de 15 comparaisons est établi ; chaque dimension dispose d'une pondération allant de 0 (si la dimension n'a été choisie dans aucune comparaison) à 5 (si la dimension a été choisie dans toutes les comparaisons).

Cette procédure de pondération est effectuée à la fin de chaque activité de la phase d'entraînement cérébral (EM, ENI et TLMC) ainsi que de la phase d'apprentissage (pré-test, leçon et résolution de problèmes). Les apprenants étaient amenés à comparer les six dimensions de l'instrument afin de pondérer chaque score qu'il a donné.

L'index NASA\_TLX pondéré est calculé en utilisant la formule suivante :

$$\text{NASA\_TLX} = \frac{\sum_{i=1}^6 X_i p_i}{15}$$

Avec  $p_i$  le poids obtenu par la dimension  $i$  à travers la procédure de pondération et  $X_i$  le score de cette même dimension  $i$ , attribué par l'apprenant lors de la phase d'évaluation.

Ainsi bien que l'instrument NASA\_TLX mesure cinq différentes dimensions de charge mentale (exigence mentale, exigence physique, exigence temporelle, performance, effort et

niveau de frustration), ces dimensions ne peuvent être interprétées de manière indépendante ; seul l'index NASA\_TLX pondéré est utilisé comme une estimation subjective de la charge mentale de travail (Hart et al., 1988a). Nous avons donc utilisé ce score pour notre étude.

#### 4.3.4 Autres variables mesurées

En plus des signaux EEG et des mesures objectives et subjectives de CMT décrits précédemment, nous avons mesuré lors de la phase d'apprentissage les émotions des participants en recourant au modèle bidimensionnel valence/activation utilisé lors de notre première étude.

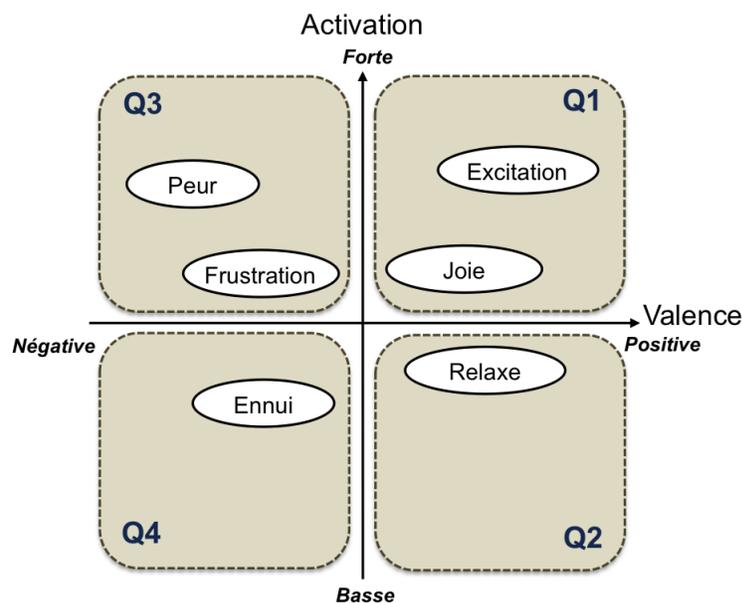


Figure 4.2 – Modèle bidimensionnel valence/activation utilisé dans la deuxième étude

Plus précisément, les participants devaient à la fin du pré-test, de la leçon et de chaque problème, spécifier le quadrant émotionnel Q1, Q2, Q3 ou Q4 (voir figure 4.2) qui correspond le plus à leur état lors de la réalisation de chaque étape. Le quadrant Q1 englobe des émotions de valence positive à forte activation (ou intensité) telles que la joie, la surprise, l'excitation. Le quadrant Q2 comprend des émotions positives de faibles intensités telles que l'état calme ou relaxe. Le quadrant Q3 inclut des émotions négatives avec une forte intensité telles que la confusion, le stress, la frustration ou la colère. Enfin, le quadrant Q4 regroupe des émotions négatives de faible intensité telles que l'ennui.

Par ailleurs, avant d'entamer le pré-test, les apprenants étaient amenés à indiquer leur niveau de compétence (faible, modéré ou expert) en trigonométrie.

## 4.4 Construction du modèle prédictif de CMT

Nous présentons dans cette section les différentes étapes de notre approche pour construire, pour chaque apprenant, un modèle de prédiction quantitatif de la charge mentale de travail à partir des signaux EEG. Le résultat de ce modèle sera sous la forme d'un index EEG de charge mentale de travail (EEG\_CMT). La figure 4.3 illustre les principales étapes de notre approche.

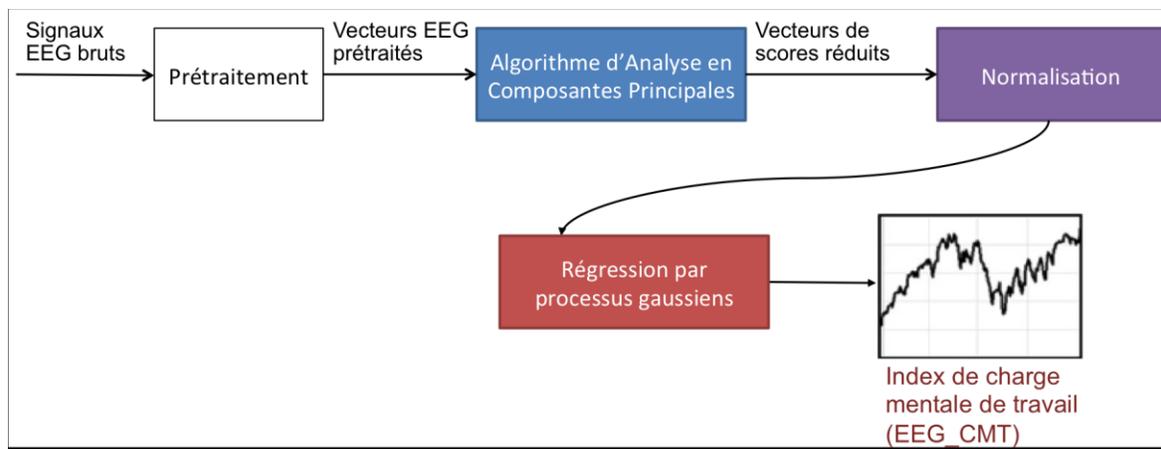


Figure 4.3 – Approche utilisée pour la construction du modèle de CMT

Les données EEG collectées durant la réalisation des activités de la phase d'entraînement cérébral serviront à l'entraînement de ce modèle. Ces données vont tout d'abord être prétraitées, réduites et normalisées.

### 4.4.1 Prétraitement des données EEG et extraction des attributs

Cette première étape consiste à découper les données EEG de chaque région en des segments d'une longueur d'une seconde avec un recouvrement d'une demi-seconde entre les segments successifs (voir figure 4.4).

Chaque segment est multiplié par une fenêtre de Hamming pour réduire l'effet de bord et une Transformée de Fourier Rapide (TFR) est appliquée pour extraire un spectre de

fréquences. Cette procédure, fréquemment utilisée dans le prétraitement des données EEG, nous a permis de transformer chaque seconde du signal physiologique en un vecteur de bandes de fréquences d'une largeur égale à 1 Hz. Pour chaque région, seules les bandes de fréquences allant de 4 Hz à 47 Hz sont conservées, et les vecteurs résultants sont concaténés pour former une seule entrée. À la fin de cette étape, chaque seconde de données est transformée en une entrée contenant 176 attributs (4 régions x 44 bandes).

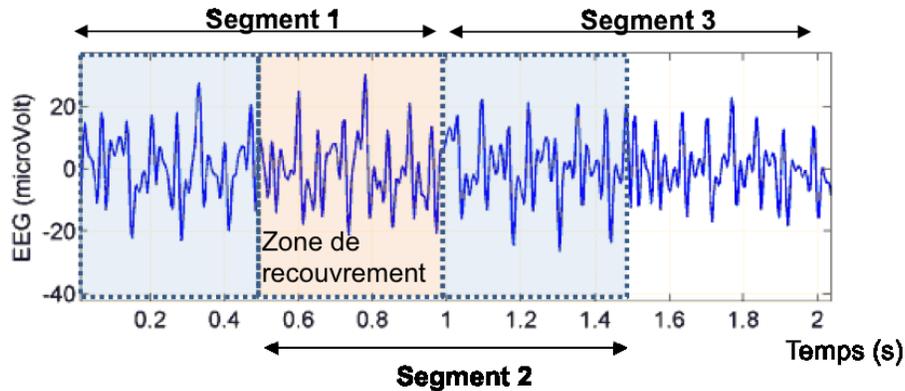


Figure 4.4 – Procédé de recouvrement pour la segmentation des données EEG

L'idée principale sera de faire correspondre à chaque seconde des signaux EEG prétraités, le score NASA\_TLX attribué par l'apprenant à travers l'évaluation subjective de l'activité. Par exemple pour une activité qui a duré 15 secondes, 15 vecteurs de données EEG seront conservés ayant chacun, comme valeur cible, le score NASA\_TLX attribué par l'apprenant.

#### 4.4.2 Réduction de la dimensionnalité

Afin de réduire la dimensionnalité des données EEG (176 attributs par vecteur EEG), nous proposons d'utiliser l'algorithme d'Analyse en Composantes Principales (ACP). Cette technique est communément employée dans les travaux de recherche qui s'intéressent à mesurer les signaux EEG, étant donnée la forte corrélation entre les données EEG en provenance des différentes régions du scalp (Heger et al., 2010; Jung et al., 1994; Liang et al., 2006; Valdés et al., 1992). L'algorithme permet de créer à partir d'un ensemble de données multivariées, un nouvel ensemble de données dont le nombre de variables (appelées composantes principales) est plus réduit que l'ensemble de départ. À chaque vecteur de

données de taille  $N$  ( $V_1, V_2, \dots, V_N$ ) dans l'ensemble de départ, correspond un vecteur de scores de taille  $k$  ( $C_1, C_2, \dots, C_k$ ) dont chaque élément  $C_i$  est le résultat d'une transformation mathématique des valeurs des variables dans le vecteur initial :

$$C_i = \beta_{1i}V_1 + \beta_{2i}V_2 + \dots + \beta_{Ni}V_N$$

Avec  $(\beta_{1i}, \beta_{2i}, \dots, \beta_{Ni})$  l'ensemble de coefficients de transformation dérivée par l'algorithme d'ACP pour la composante  $i$  ( $1 \leq i \leq k \leq N$ ).

Chacune des composantes résultantes de l'algorithme d'ACP permet d'expliquer un certain taux de variance dans l'ensemble de données de départ indépendamment des autres composantes. Plus précisément, la première composante explique le plus grand taux de variance ; la deuxième explique le second plus grand taux dans la variance laissée inexpliquée par la première ; la troisième explique le troisième plus grand taux de variance laissé inexpliqué par la première et la deuxième, et ainsi de suite pour le reste des composantes. Ainsi, le nombre de composantes sélectionnées détermine la variance totale expliquée par cette transformation, c.-à-d. la somme des variances expliquée par chaque composante.

Ainsi, les données EEG de chaque participant ont été transformées, en utilisant l'algorithme d'ACP, en un nouvel ensemble contenant 25 composantes (un pourcentage de 78.63 % de réduction). Le taux de variance expliquée pour chaque participant était en moyenne de l'ordre de  $M = 90.42$  %,  $SD = 3.30$  %, allant de 85.98 % (le taux de variance le plus faible) à 94.71 % (le taux de variance le plus fort).

**Normalisation des données.** Avant de procéder à l'entraînement de chaque modèle, les données transformées par l'algorithme d'ACP (appelée scores) sont normalisées en soustrayant, pour chaque composante principale, la moyenne des scores des données d'entraînement et en divisant par leur déviation standard.

#### 4.4.3 Régression par processus gaussiens

Une fois les données prétraitées, nous disposons d'un ensemble de vecteurs d'entrées  $\mathbf{X} = \{x_i\}_{i=1}^n$ , où chaque entrée  $x_i$  représente un vecteur de scores normalisés issu de

l'algorithme d'ACP et correspondant à une seconde d'une activité exécutée par l'apprenant. À chacune de ces entrées correspond une valeur cible  $y_i$  qui représente le score NASA\_TLX attribué par l'apprenant.

Ainsi notre démarche consiste, dans un premier temps, à utiliser un ensemble de données d'entraînement  $\mathcal{D} = \{x_i, y_i\}_{i=1}^n$  résultant des données collectées durant l'exécution des activités d'entraînement cérébral pour développer un modèle individuel permettant de donner une estimation quantitative de la charge mentale de travail de chaque apprenant (EEG\_CMT). Le modèle entraîné sera ensuite utilisé pour dériver la charge mentale de travail lors de la phase d'apprentissage.

Comme nous l'avons spécifié précédemment, une grande variété de techniques d'apprentissage machine ont été proposées pour modéliser la charge mentale de travail à partir des caractéristiques du signal EEG. Pour notre approche, nous proposons de construire notre modèle en utilisant les processus gaussiens. En effet, les processus gaussiens sont très fréquemment utilisés pour modéliser les données continues qui fluctuent au cours du temps (Davis, 2006) comme c'est le cas pour la charge mentale de travail. Les processus gaussiens ont également l'avantage par rapport aux autres techniques de fournir une estimation de probabilité sur la valeur prédite.

L'apprentissage machine par processus gaussiens est une technique bayésienne non paramétrique supervisée permettant de créer une fonction non-linéaire de régression (Rasmussen, 2006). Pour notre étude, nous avons entraîné une fonction à noyau gaussien et un bruit gaussien  $\sigma$  égal à 0.1 (voir l'Annexe B pour plus de détails sur les processus gaussiens). Cette fonction de régression sera testée pour prédire la valeur de la charge mentale de travail à partir des données EEG durant l'exécution des activités d'apprentissage.

## 4.5 Résultats et discussion

Les résultats de nos expérimentations sont présentés dans les sous-sections suivantes. La première partie est consacrée à la modélisation de la charge mentale de travail des apprenants.

Nous y analyserons d'abord nos données d'entraînement. Puis, nous nous y intéresserons à l'évaluation et la validation de notre modèle durant la phase d'apprentissage. Dans la deuxième partie, nous étudierons l'index d'engagement, et dans la troisième nous analyserons la relation entre les deux index cérébraux (engagement et charge mentale) et les émotions.

#### **4.5.1 Modélisation de la charge mentale de travail**

Avant de présenter les résultats de notre modèle prédictif de la charge mentale de travail, nous nous intéresserons à la validation des mesures subjectives, qui ont été utilisées comme les valeurs cibles de notre modèle.

##### **4.5.1.1 Analyse de la phase d'entraînement cérébral**

Afin de valider les estimations subjectives de la charge mentale de travail mesurées à l'aide de l'index NASA\_TLX, durant la phase d'entraînement cérébral, des analyses corrélationnelles entre cet index, le niveau de difficulté des différentes tâches, et les mesures objectives de CMT (performance et temps de réponse) ont été réalisées.

Une ANOVA à mesures répétées ( $N = 17$ ) a été réalisée pour étudier la variation des mesures subjectives de la charge mentale de travail à travers les différents niveaux de difficulté de chacune des trois tâches cognitives : Empan Numérique (EN), Empan Numérique Inversé (ENI) et Test de Logique et de Calcul Mental (TLCM). Les degrés de liberté ont été corrigés en utilisant les estimations de sphéricité de Greenhouse-Geisser pour EN et ENI ( $\epsilon = 0.35$  et  $0.54$ , respectivement) car l'hypothèse de sphéricité a été violée ( $\chi^2 = 60.11$  et  $54.40$ , respectivement,  $p < 0.05$ ), tandis que le test de Mauchly était non significatif pour le TLCM.

Les résultats ont révélé des variations significatives du score NASA\_TLX selon le niveau de difficulté dans les trois tâches cognitives :  $F(1.73, 27.65) = 25.65$ ,  $p < 0.001$  pour EN,  $F(2.18, 34.89) = 18.25$ ,  $p < 0.001$  pour ENI, et  $F(2, 32) = 43.51$ ,  $p < 0.01$  pour TLCM. Une augmentation linéaire significative en fonction des exigences requises par chaque niveau de difficulté a été trouvée. Les différentes moyennes des scores NASA\_TLX sont illustrées dans la figure 4.5.

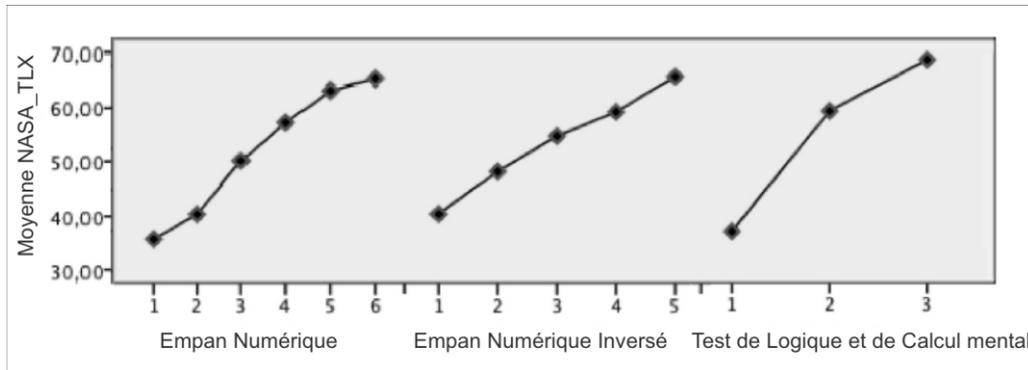


Figure 4.5 – Moyennes des estimations subjectives de la CMT (NASA\_TLX) à travers les différents niveaux de difficulté des tâches d’entraînement cérébral : Empan Numérique (6 niveaux de difficulté), Empan Numérique Inversé (5 niveaux de difficulté) et Test de Logique et de Calcul Mental (3 niveaux de difficulté)

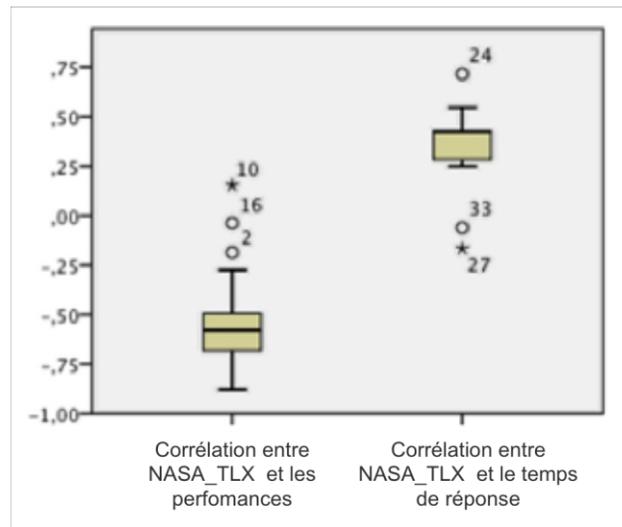


Figure 4.6 – Corrélations bivariées entre le score NASA\_TLX et les mesures objectives (performances et temps de réponse)

Des analyses de corrélations entre le score NASA\_TLX, les performances d’une part et le temps de réponse d’une autre part ont été effectuées. Ces corrélations ont été calculées pour chaque sujet à travers les 14 niveaux des trois tâches cognitives EN, ENI et TLCM. La corrélation médiane entre le score NASA\_TLX et les performances était de -0,58 avec un minimum de 0,15 et un maximum de 0,89. La corrélation médiane entre le NASA\_TLX et le temps de réponse était de 0,42 avec une variation de -0,16 à 0,72. La figure 4.6 illustre les résultats de ces corrélations.

Afin d'approfondir l'analyse entre ces deux mesures objectives et le score NASA\_TLX, des corrélations ont été calculées pour chacun des trois types d'activités cognitives à travers les 17 participants. Les résultats ont démontré que la performance diminue de façon linéaire avec l'augmentation de la charge mentale de travail dans EN ( $r = -0.59$ ,  $p < 0.001$ ) et TLCM ( $r = -0.50$ ,  $p < 0.001$ ), alors que la relation n'était pas linéairement significative dans ENI ( $r = -0.19$ ,  $p = \text{n.s.}$ ). Le temps de réponse augmente linéairement en fonction de la charge mentale de travail dans EN ( $r = 0.30$ ,  $p < 0.05$ ) et TLCM ( $r = 0.35$ ,  $p < 0.05$ ). Ces variables n'étaient pas linéairement corrélées dans ENI ( $r = 0.014$ ,  $p = \text{n.s.}$ )

Pour résumer ces résultats, le score NASA\_TLX de la charge mentale de travail a montré une relation linéaire avec les mesures objectives dans l'activité Empan Numérique et le Test de Logique et de Calcul Mental de la phase d'entraînement cérébral. Par ailleurs, la variation (ou l'évolution) de cette mesure semble refléter la conception expérimentale prévue pour notre étude, dont le but était d'induire des valeurs croissantes de charge mentale de travail en fonction du niveau de difficulté qui était de plus en plus élevé dans chaque tâche. Ces éléments démontrent ainsi que les évaluations subjectives de l'instrument NASA\_TLX constituent un indicateur fiable de la charge mentale de travail pour l'entraînement de nos modèles prédictifs EEG.

#### **4.5.2 Validation du modèle proposé durant la phase d'apprentissage**

Un modèle de prédiction de la charge mentale de travail a été entraîné à partir des données EEG issues de la phase d'entraînement cérébral, et ce pour chaque apprenant. Ce modèle a été testé durant la phase d'apprentissage pendant laquelle les apprenants étaient amenés à utiliser un environnement éducatif de trigonométrie.

Ainsi, pour chaque apprenant, les données EEG étaient enregistrées durant les différentes activités de la phase d'apprentissage (pré-test, leçon et problèmes 1 à 6 de la session de résolution de problèmes). Ces données ont ensuite été prétraitées, réduites et normalisées ; puis, introduites, comme entrée (inputs), dans son modèle 'individuel' de prédiction. À chaque seconde de données EEG, le modèle infère une valeur de charge mentale

de travail. Un index EEG\_CMT (global) pour chaque activité est établi en moyennant les valeurs de prédiction faite par le modèle sur toute la durée de l'activité.

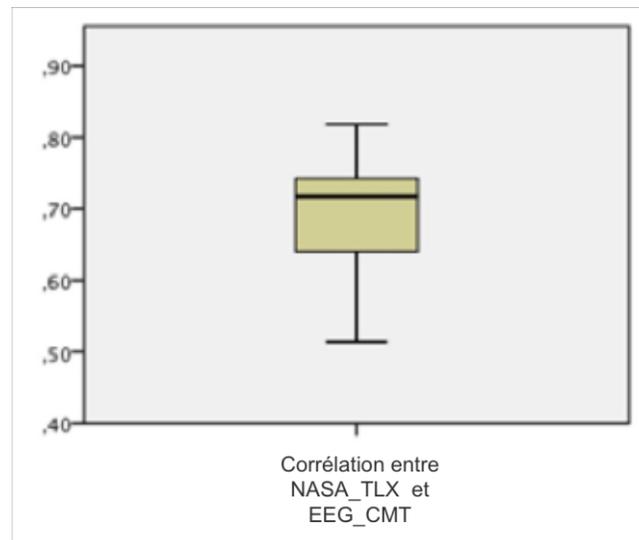


Figure 4.7 – Corrélations bivariées entre le score NASA\_TLX et l'index EEG\_CMT de charge mentale de travail dérivé par notre modèle

La boîte à moustaches de la figure 4.7 illustre les résultats des corrélations bivariées entre l'index prédit par le modèle EEG (EEG\_CMT) et les mesures subjectives des scores NASA\_TLX. Les corrélations ont été calculées pour tous les participants durant le pré-test, la session d'apprentissage et chacun des six problèmes résolus. Ces corrélations ont montré des relations significatives dans les huit activités. La corrélation médiane était de 0.72 avec un intervalle de variation allant de 0.52 ( $p < 0.001$ ) pour le problème 2, à 0.82 ( $p < 0.05$ ) pour le problème 1. Ces résultats prouvent que les prédictions de charge mentale de travail données par notre modèle correspondent bien aux évaluations subjectives des participants, démontrant ainsi la validité de notre modèle.

Dans la prochaine étape, nous allons nous intéresser à l'analyse de la progression du niveau de la charge mentale de travail — telle que prédite par notre modèle — durant la phase d'apprentissage. Le but est de vérifier si notre index cérébral reflète bien le déroulement de l'expérience de l'apprenant lors de l'exécution des différentes activités de l'environnement de trigonométrie.

Une ANOVA à mesures répétées a été réalisée pour examiner l'évolution de l'index EEG de charge mentale de travail (EEG\_CMT) à travers les différentes activités d'apprentissage à savoir : le pré-test, la leçon et chacun des six problèmes (prob.1 à prob. 6). Cette ANOVA a révélé qu'il y avait des variations significatives dans la valeur de l'index EEG\_CMT entre les différentes activités  $F(3.23, 51.61 = 2.76, p < 0.05)$ . Les degrés de liberté ont été corrigés à partir des estimations de sphéricité de Greenhouse-Geisser (epsilon = 0.46). La figure 4.8 illustre ces variations.

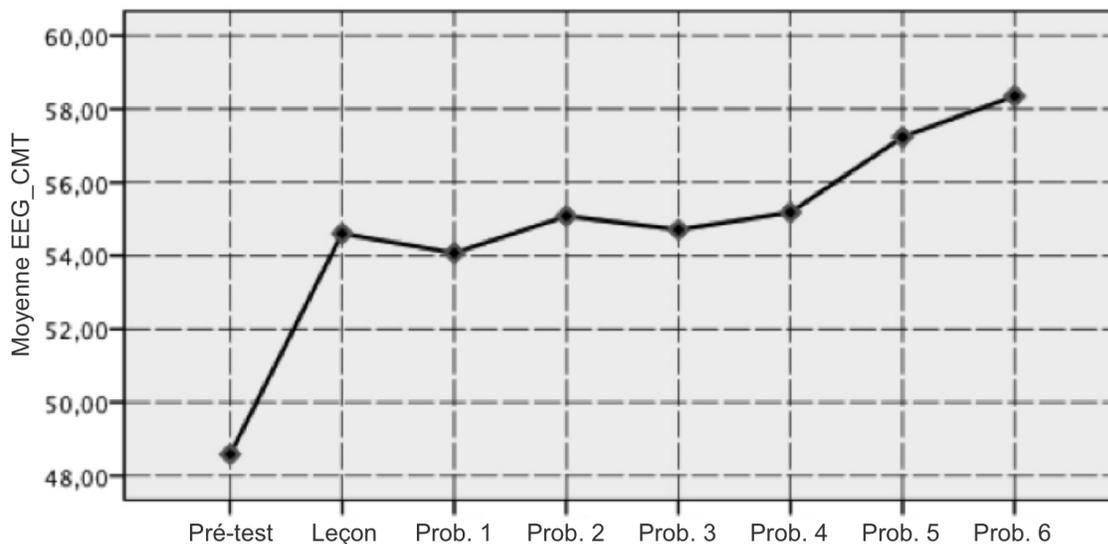


Figure 4.8 – Moyenne des mesures EEG\_CMT sur chaque activité de l'environnement d'apprentissage

Les résultats post-hoc ont montré que les mesures de l'index EEG\_CMT ont considérablement augmenté au cours de la leçon par rapport au pré-test ( $p < 0.05$ ). Cette augmentation reflète bien l'effort produit par les apprenants pour comprendre et mémoriser les concepts présentés dans cette leçon par rapport à la séance de pré-test, où les apprenants devaient répondre à des questions sans aucune obligation de bonne performances. En effet, lors du pré-test, il n'y avait pas de pression particulière sur les apprenants, qui étaient informés que le but du pré-test était de simplement situer leur niveau de connaissances en trigonométrie avant d'entamer la leçon.

Des augmentations significatives ont également été repérées entre le pré-test et les problèmes 4, 5 et 6 ( $p < 0.05$ ), démontrant que l'index EEG de charge mentale augmente de manière significative entre le début et la fin de l'interaction d'apprentissage. Une augmentation significative a également été constatée au cours du problème 5 par rapport aux problèmes 1, 3 et 4 ( $p < 0.05$ ). La variation de cet index montre ainsi un comportement cohérent avec la conception générale de la phase de résolution de problèmes, qui se base sur une augmentation graduelle du niveau de difficulté entre ses différents problèmes.

Ces résultats permettent donc d'établir les deux conclusions suivantes. Premièrement, l'approche proposée pour modéliser la charge mentale de travail à partir des caractéristiques du signal EEG a été validée dans un contexte d'apprentissage. En effet, l'index cérébral EEG\_CMT, que prédit notre modèle, présente de fortes corrélations avec les évaluations subjectives des apprenants. Celles-ci ont été mesurées avec l'instrument NASA\_TLX, que nous avons également validé lors de la phase d'entraînement cérébral, par rapport au niveau de difficulté des tâches cognitives. Deuxièmement, l'index cérébral proposé reflète bien le déroulement de la session d'apprentissage. Celui-ci présente en effet des variations qui sont conformes avec la conception de notre environnement d'apprentissage. La valeur de l'indice augmente significativement au fur et à mesure que les apprenants acquièrent des compétences en trigonométrie : une tendance notamment observée lorsque l'on compare les valeurs de l'index entre le pré-test, la leçon et la session résolution des problèmes.

### **Analyse de l'index EEG\_CMT par rapport aux mesures objectives**

Afin d'analyser l'indicateur cérébral de CMT par rapport aux mesures objectives (performance et temps de réponse) durant la phase d'apprentissage, deux types de corrélations bivariées ont été calculées.

Les premières corrélations ont été réalisées entre l'index *EEG\_CMT* et le *temps de réponse* sur tous les participants dans les six problèmes ( $N = 17 \times 6$ ), montrant une relation significative, mais assez faible entre les deux variables ( $r = 0.21$ ,  $p < 0.05$ ). Par ailleurs, une analyse corrélacionnelle de chaque problème séparément a été calculée. Celle-ci a révélé une corrélation significative uniquement pour le problème 5 ( $r = 0.56$ ,  $p < 0,05$ ). En effet,

contrairement aux tâches purement cognitives de la phase d'entraînement cérébral, qui impliquait des conditions d'exécution strictes, contrôlées, et des limites de temps imposées, l'environnement de trigonométrie de la phase d'apprentissage était moins contrôlé et moins contraignant pour les apprenants. Dans de telles conditions, un temps de réponse plus long n'implique pas nécessairement une charge mentale de travail supérieure, contrairement à la phase d'entraînement cérébral.

Les deuxièmes corrélations bivariées ont concerné l'index *EEG\_CMT* et les *performances* des apprenants dans le pré-test et chaque problème. Les résultats ont montré une relation linéaire significative entre ces deux variables dans le pré-test ( $r = -1.88$ ,  $p < 0.05$ ). Aucune corrélation significative n'a été trouvée dans les problèmes. Ces résultats confirment également la complexité de la relation entre la charge mentale de travail et la performance durant les tâches d'apprentissage : une augmentation linéaire de la charge mentale de travail n'impliquera pas forcément une diminution de la performance, et vice versa (comme nous l'avons remarqué durant les tâches cognitives d'entraînement cérébral Empan Numérique et Test de Logique et de Calcul Mental).

### **Analyse de l'index EEG\_CMT par rapport aux caractéristiques des apprenants**

Après avoir validé notre index cérébral par rapport aux évaluations subjectives de CMT des apprenants, examiné son évolution au cours de la session d'apprentissage et étudié son impact sur les performances et le temps de réponse, nous analysons dans cette partie le comportement de cet indicateur mental par rapport au *niveau de compétence* des apprenants. Nous rappelons que cette variable a été mesurée avant que les apprenants commencent leur session d'apprentissage avec le système : on leur a demandé de spécifier, avant le pré-test, leur niveau de connaissance en trigonométrie (faible, modéré ou expert).

Une ANOVA a été réalisée afin d'étudier l'impact de cette variable sur la charge mentale de travail. Nous avons trouvé un effet significatif du niveau de compétence sur la moyenne de l'index EEG\_CMT lors de la résolution des problèmes ( $F(2, 14) = 11.93$ ,  $p < 0.05$ ). Nous avons trouvé que les participants, ayant un niveau de compétence modéré, ont eu les valeurs de charge mentale de travail les plus élevées, durant la résolution des problèmes

( $M = 68.97$ ;  $SD = 14.16$ ), par rapport aux participants ayant un faible niveau de compétence ( $M = 57.25$ ;  $SD = 4.06$ ), ainsi qu'aux participants ayant un niveau expert, qui avaient les valeurs de charge de mentale les plus faibles ( $M = 41.35$ ,  $SD = 7.52$ ).

Ainsi, d'un côté, les apprenants ayant des compétences modérées ont eu tendance à produire plus d'effort mental (en terme de CMT prédite par notre modèle), par rapport à des apprenants ayant plus de compétences. Ceci pourrait être en effet interprété de la façon suivante : ces apprenants, plus experts en trigonométrie, sont plus à l'aise avec les notions trigonométriques présentées (qu'ils connaissent au préalable). Ils ont alors eu tendance à produire beaucoup moins d'effort lors de la résolution des problèmes.

D'un autre côté, les apprenants ayant le plus faible niveau de compétence ont eu tendance à produire moins d'effort mental et à être moins investis dans la réalisation des activités par rapport à des apprenants ayant plus de connaissances (niveau modéré)

La dernière investigation de cette partie porte sur l'impact global de la charge mentale de travail, du temps de réponse et du niveau de compétence sur les performances des apprenants dans la résolution des problèmes. Une analyse de régression multiple a été menée pour mesurer l'influence de chacun de ces paramètres. Le modèle global était significatif ( $F(3, 98) = 8.48$ ,  $p < 0.01$ ,  $R = 0.41$ ). Les analyses des effets conditionnels principaux ont révélé un effet positif de la charge mentale de travail ( $\beta = 0.24$ ,  $p < .05$ ) et du niveau de compétence ( $\beta = .42$ ,  $p < 0.001$ ), et un effet négatif du temps de réponse ( $\beta = -0.27$ ,  $p < 0.05$ ). Ceci suggère ainsi qu'une combinaison de ces différentes variables peut mieux prédire la performance d'apprentissage. En effet, comme nous l'avons constaté précédemment en étudiant la corrélation directe entre l'index EEG\_CMT et les performances, il n'y a pas d'association linéaire entre ces deux variables. Ce résultat confirme alors que d'autres variables telles que le temps de réponse ou le niveau d'expertise de la tâche peuvent contribuer à la prédiction de la performance des apprenants.

#### **4.5.2 Analyse de l'engagement mental**

Pour chaque participant, un index EEG d'engagement mental a été calculé pour chaque activité de la phase d'apprentissage, à savoir : le pré-test, la leçon et les problèmes 1 à 6. Pour

ce faire, les bandes de fréquences  $\theta$  (4-8 Hz),  $\alpha$  (8-13 Hz) et  $\beta$  (13-22 Hz) ont été extraites à partir des signaux EEG en provenance des différentes régions du scalp mesurées ( $P3$ ,  $C3$ ,  $Pz$  et  $Fz$ ), et l'index d'engagement a été dérivé en utilisant la procédure décrite dans le chapitre précédent (section 3.3).

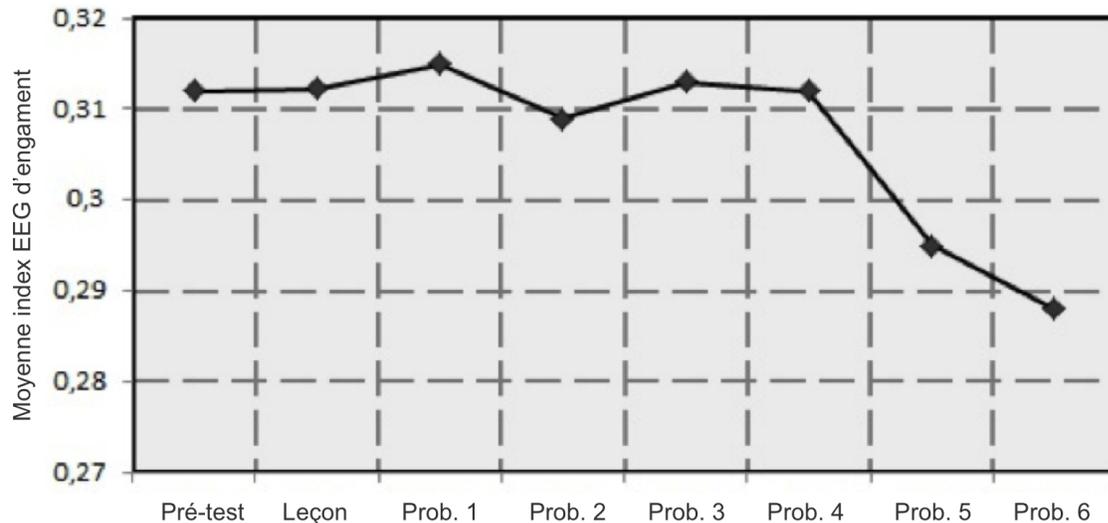


Figure 4.9 – Évolution de l'index EEG d'engagement mental à travers les activités d'apprentissage

Afin d'évaluer la pertinence de cet index dans ce nouvel environnement d'apprentissage, nous nous sommes intéressés, tout comme pour la charge mentale de travail, à analyser le comportement de cet index par rapport au déroulement des différentes activités proposées. À cette fin, une ANOVA à mesures répétées a été réalisée pour suivre l'évolution de l'index d'engagement au fil des différentes activités. Cette ANOVA a révélé qu'il n'y avait pas de changements significatifs dans l'index d'engagement des apprenants entre les activités d'apprentissage  $F(3.66, 58.70 = 0.69, p = n.s)$ . Les degrés de liberté ont été corrigés à partir des estimations de sphéricité de Greenhouse-Geisser ( $\epsilon = 0.52$ ).

En revanche, les résultats post-hoc ont montré qu'il existe une baisse significative dans les mesures de l'index d'engagement entre la leçon et le dernier problème de la session d'interaction avec le système ( $p < 0.05$ ). Ainsi, contrairement à la charge mentale de travail qui avait tendance à augmenter au fur et à mesure que les apprenants interagissaient avec le système, l'index d'engagement a montré une tendance à la baisse (voir figure 4.9).

La baisse globale de la valeur de l'index d'engagement peut notamment s'expliquer par la conception même de notre étude expérimentale, qui est destinée à solliciter un niveau de charge mentale de plus en plus croissant chez les apprenants. Ceci pourrait alors, à la longue, perturber leurs niveaux d'attention et de vigilance. En effet, le passage d'un problème à un autre par exemple peut ne pas avoir d'effet immédiat sur l'engagement, dans la mesure où la nature stimulante des problèmes en soi peut susciter l'intérêt des apprenants. Par contre au bout d'un certain nombre de résolutions, et avec l'augmentation de la charge mentale de travail, l'engagement mental peut baisser de manière assez importante.

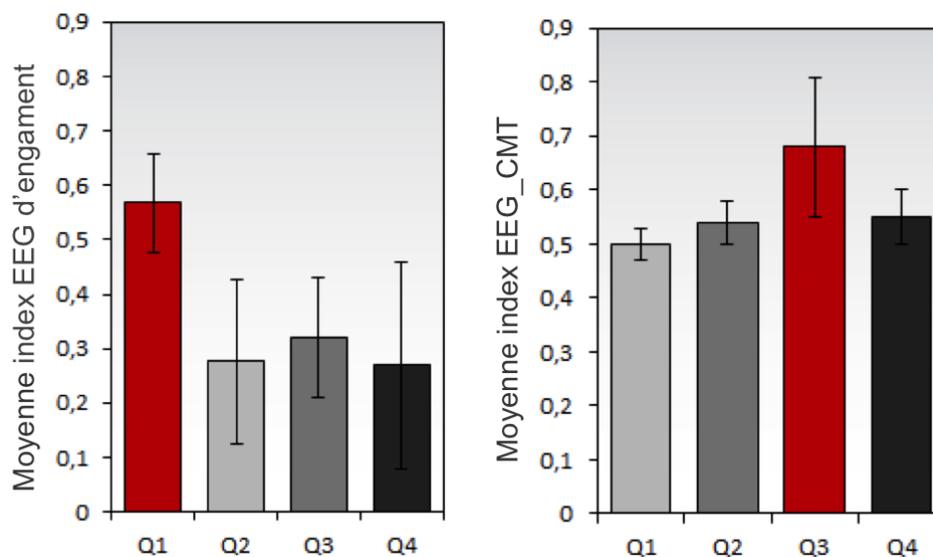


Figure 4.10 – Moyennes des index cérébraux (à gauche l'index d'engagement et à droite la charge mentale de travail) à travers les quatre quadrants d'émotions.

### 4.5.3 Relation entre les index cérébraux et les émotions

Afin d'analyser la relation entre les deux indicateurs cérébraux (charge de travail et engagement) et les émotions durant la phase d'apprentissage, des tests d'ANOVA ont été réalisés. Les émotions des apprenants ont été mesurées en termes de quadrants émotionnels : Q1, Q2, Q3 et Q4 du modèle bidimensionnel valence/activation (voir section 4.3.4).

Ces analyses de variance ont révélé qu'il existe un effet significatif de l'état émotionnel de l'apprenant, selon le type de quadrants, sur la valeur de l'index d'engagement :  $F(3, 132) = 3.32$ ,  $p < 0.05$ , ainsi que la valeur de l'index de charge mentale de travail :

$F(3, 132 = 4.52, p < 0.05)$ . Ces effets significatifs ( $p < 0.05$ ) ont également été observés pour chaque activité prise à part (c.-à-d. en effectuant des tests ANOVA sur les données du pré-test, de la leçon et de chaque problème séparément). La figure 4.10 illustre les moyennes et écarts types des index cérébraux à travers les quatre quadrants d'émotions.

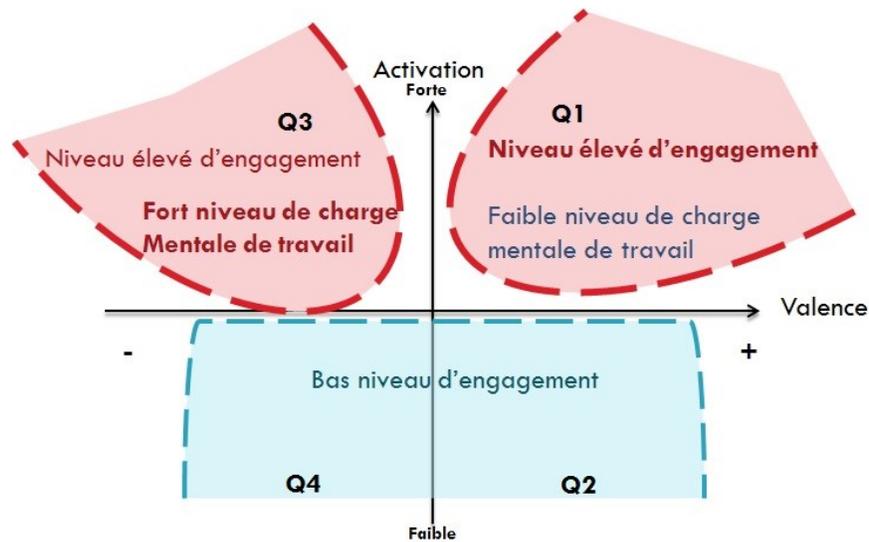


Figure 4.11 – Correspondance entre le modèle bidimensionnel valence/activation des émotions et les indicateurs cérébraux d'engagement et de charge mentale de travail.

Les analyses post-hoc nous ont permis d'arriver aux trois constatations suivantes (voir figure 4.11) :

1. Le quadrant Q1 (valence positive et activation élevée) présente en moyenne les valeurs d'engagement mental les plus élevées ( $M = 0.568, SD = 0.09$ ) et les valeurs de charge mentale de travail les plus basses ( $M = 0.50, SD = 0.03$ ).
2. Les quadrants Q2 et Q4 des émotions de faible activation présentent les moyennes d'engagement les plus faibles soit ( $M = 0.26, SD = 0.15$ ) pour Q2 (valence positive) et ( $M = 0.27, SD = 0.19$ ) pour Q4 (valence négative).
3. Le quadrant Q3 (valence négative et activation élevée) est associé aux valeurs de charge mentale de travail les plus élevées ( $M = 0.68, SD = 0.13$ ).

Ces résultats démontrent ainsi que ces indicateurs d'états mentaux sont bien associés aux émotions de l'apprenant. Le suivi de ces indicateurs pourra donc fournir des indications

pertinentes quant à l'expérience émotionnelle de l'apprenant durant son interaction avec l'environnement d'apprentissage, mais aussi cognitive comme nous avons pu le constater en analysant l'évolution de ces index au fil des différentes activités d'apprentissage. Dans le chapitre suivant, nous nous pencherons sur le dernier objectif de cette thèse, qui sera de proposer un système tutoriel adaptatif selon l'état mental de l'apprenant.

## 4.6 Conclusion

Dans ce chapitre, nous avons présenté notre approche pour modéliser la charge mentale des apprenants. Cette approche s'inscrit dans le cadre du premier objectif de cette thèse, qui vise à augmenter les Systèmes Tutoriels Intelligents par des indicateurs d'états mentaux. Pour ce faire, une deuxième étude expérimentale a été réalisée. Cette étude nous a permis d'atteindre les trois buts principaux que nous avons fixés à savoir :

- (1) Vérifier la possibilité de modéliser, à partir des caractéristiques du signal EEG, la charge mentale de travail dans un contexte d'apprentissage. Ceci inclut l'entraînement d'un modèle, sa validation dans un environnement éducatif, et l'évaluation de son impact sur le déroulement de la session d'interaction avec l'environnement.
- (2) Analyser le comportement de l'index d'engagement mental, présenté dans le chapitre précédent, dans un nouveau contexte d'apprentissage.
- (3) Examiner la relation entre les deux index cérébraux (charge mentale de travail et engagement) et les états émotionnels de l'apprenant.

Les questions que l'on se pose à la fin de ce chapitre sont les suivantes. Peut-on détecter *en temps réel* ces indicateurs cérébraux dans un Système Tutoriel Intelligent ? Et si oui, comment un tel système pourrait-il prendre en compte ces indicateurs pour *adapter* ces interventions de manière à ajuster son enseignement et améliorer l'expérience d'interaction de l'apprenant ?

Nous aborderons ces questions dans le prochain chapitre qui sera consacré à la présentation de notre système MENTOR (pour MENTal tuTOR) : un système tutoriel qui s'adapte en fonction des indicateurs mentaux de l'apprenant.



## Chapitre 5

# Le système MENTOR

### 5.1 Introduction

Dans les chapitres 3 et 4 de cette thèse, nous avons présenté deux études expérimentales qui ont permis de valider la possibilité de modéliser respectivement l'engagement mental et la charge mentale de travail dans un contexte éducatif à partir de l'activité cérébrale des apprenants. Dans ces deux chapitres, la détection et l'analyse de ces indicateurs d'états mentaux ont été réalisées *rétrospectivement* (c.-à-d. en « offline »). Ceci nous a permis d'examiner le comportement de ces indicateurs cérébraux (à postériori) et de souligner leur importance dans le suivi de l'expérience de l'apprenant aussi bien au niveau cognitif qu'au niveau émotionnel. L'intégration de ces indicateurs semble ainsi constituer une avenue très prometteuse pour les STEI.

Notre prochaine étape est désormais de montrer la possibilité de détecter ces indicateurs d'engagement et de charge de travail en temps réel, mais aussi d'adapter, en temps réel également, le processus d'apprentissage en fonction de ces indicateurs. En effet comme nous avons pu le constater à travers notre revue de la littérature, bien que plusieurs recherches se sont intéressées à l'utilisation des senseurs physiologiques pour étudier les réactions

émotionnelles des apprenants, les systèmes qui intègrent ces senseurs pour adapter dynamiquement leur processus de formation en temps réel sont relativement rares.

Ainsi, nous présentons dans ce chapitre notre approche pour intégrer ces indicateurs cérébraux dans un système tutoriel : le système MENTOR (« MENTal tuTOR »). Ce système permet de mesurer les signaux physiologiques EEG de l'apprenant et d'évaluer, son état mental en fonction de ses index d'engagement et de charge mentale de travail, qu'il pourra calculer et suivre activement tout au long de la session d'interaction. L'objectif de MENTOR est donc de tenir compte de ces indicateurs cérébraux pour adapter le déroulement de la session d'apprentissage de la façon qui conviendrait le plus à l'état mental de l'apprenant.

Ce chapitre est structuré en deux parties. La première partie sera consacrée à la présentation du Système MENTOR. Nous y détaillerons les différentes fonctionnalités et modules de traitement de notre système. La deuxième partie sera quant à elle dédiée à l'évaluation de notre système à travers une troisième étude expérimentale que nous avons réalisée dans le cadre de cette thèse. Le but de cette étude a été de mettre en évidence l'impact que peut avoir l'intégration de ces index cérébraux sur l'apprentissage. En particulier, nous nous sommes intéressés à ces trois questions de recherche. (1) Est-ce que l'intégration des indicateurs mentaux dans la politique d'adaptation d'un STI peut aider l'apprenant à mieux assimiler le contenu tutoriel et à réaliser de meilleures performances d'apprentissage ? (2) Est-ce qu'un système, qui est 'conscient' de l'état mental d'un apprenant, peut améliorer l'expérience de ce dernier vis-à-vis du déroulement de son apprentissage (en terme de satisfaction en particulier et de réactions émotionnelles en général) ? Et enfin, (3) est-ce qu'un tel système arrive à déceler et à corriger les difficultés mentales de l'apprenant ?

## **Partie I — Présentation du système MENTOR**

Le système MENTOR «MENTal tuTOR » que nous proposons dans le cadre de cette thèse est un système tutoriel qui intègre des indicateurs cérébraux dans sa logique d'adaptation avec l'apprenant. Son objectif est d'ajuster l'enseignement prodigué en adaptant ses interventions pédagogiques selon l'état mental de l'apprenant en vue d'enrichir la dynamique d'interaction avec l'apprenant et améliorer l'expérience d'apprentissage.

Plus précisément, notre système utilise les données physiologiques en provenance de senseurs électroencéphalographiques (EEG) pour calculer en temps réel les deux index cérébraux d'engagement et de charge mentale de travail. Ces index sont ensuite utilisés pour analyser l'état mental de l'apprenant, et adapter l'environnement d'apprentissage.

MENTOR dispose de trois différents modes de fonctionnement à savoir : le mode *Entrainement*, le mode *Affichage* et le mode *Apprentissage*.

- Le mode *Entrainement* (« Training ») : ce mode de fonctionnement est exclusivement dédié à la création d'un *modèle prédictif de la charge mentale de travail* pour un apprenant. Nous rappelons que contrairement à l'index d'engagement mental qui est directement calculé à partir des bandes de fréquences alpha, beta et thêta extraites du signal EEG, l'index de charge mentale de travail que nous proposons nécessite une phase d'entraînement (c.-à-d. une phase d'apprentissage machine). À l'issue de cette phase, un modèle de charge mentale de travail propre à chaque apprenant est créé. Ce modèle est utilisé par les deux autres modes de fonctionnement de MENTOR, à savoir le mode *Affichage* et le mode *Apprentissage*, pour dériver en temps réel l'indice de charge mentale de travail à partir des données EEG.
- Le mode *Affichage* (« Monitoring ») : ce mode de fonctionnement passif (c.-à-d. non adaptatif) présente un *tableau de bord* de l'état mental de l'apprenant. Le système enregistre et affiche sur l'écran les valeurs des index d'engagement et de charge mentale de travail calculées en temps réel lors d'une session d'interaction.
- Le mode *Apprentissage* (« Learning ») : ce mode offre un environnement d'apprentissage adaptatif. Le système surveille activement les index mentaux de l'apprenant lors d'une activité d'apprentissage, analyse l'évolution de son état mental, et décide d'initier des interventions pour adapter l'environnement d'apprentissage en fonction de l'état mental de l'apprenant.

La figure 5.1 donne un aperçu global du fonctionnement du système MENTOR. En plus des trois modes de fonctionnement décrit précédemment, MENTOR inclut deux phases préliminaires de traitement des signaux physiologiques à savoir : (1) la capture des signaux

EEG et (2) le calcul des index cérébraux. Dans la première phase, les données EEG brutes sont lues à partir des senseurs, filtrées et transformées afin qu'elles puissent être exploitées pour le calcul des deux index cérébraux. L'indice d'engagement est dérivé directement à partir des signaux EEG capturés dans la première phase préliminaire de traitement. L'index de charge mentale de travail est dérivé quant à lui à partir des signaux EEG, et du modèle entraîné à travers le mode de fonctionnement *Entrainement* de MENTOR. Les index produits par cette deuxième phase peuvent être alors soit intégrés et utilisés par l'environnement d'apprentissage adaptatif si MENTOR est utilisé en mode *Apprentissage*, ou présentés directement à l'écran si le système est utilisé en mode *Affichage*.

La conception du système MENTOR repose sur une organisation modulaire ; chaque module est responsable d'un aspect différent de traitement (capture des signaux EEG, calcul des index cérébraux, module d'analyse, module décisionnel, etc.). Nous présentons dans ce qui suit les fonctionnalités de ces différents modules.<sup>1</sup>

---

<sup>1</sup> Le système MENTOR a été entièrement développé en utilisant le langage C# (plateforme Microsoft.Net version 4.0.30319). Le choix de ce langage a été motivé principalement par rapport au fait qu'il est supporté par le langage utilisé par le kit de développement de notre matériel EEG (C ++). Le langage C# a par ailleurs l'avantage de faciliter la création des interfaces et des composantes graphiques, qui représentent un aspect essentiel dans ce genre d'applications à caractère interactif.

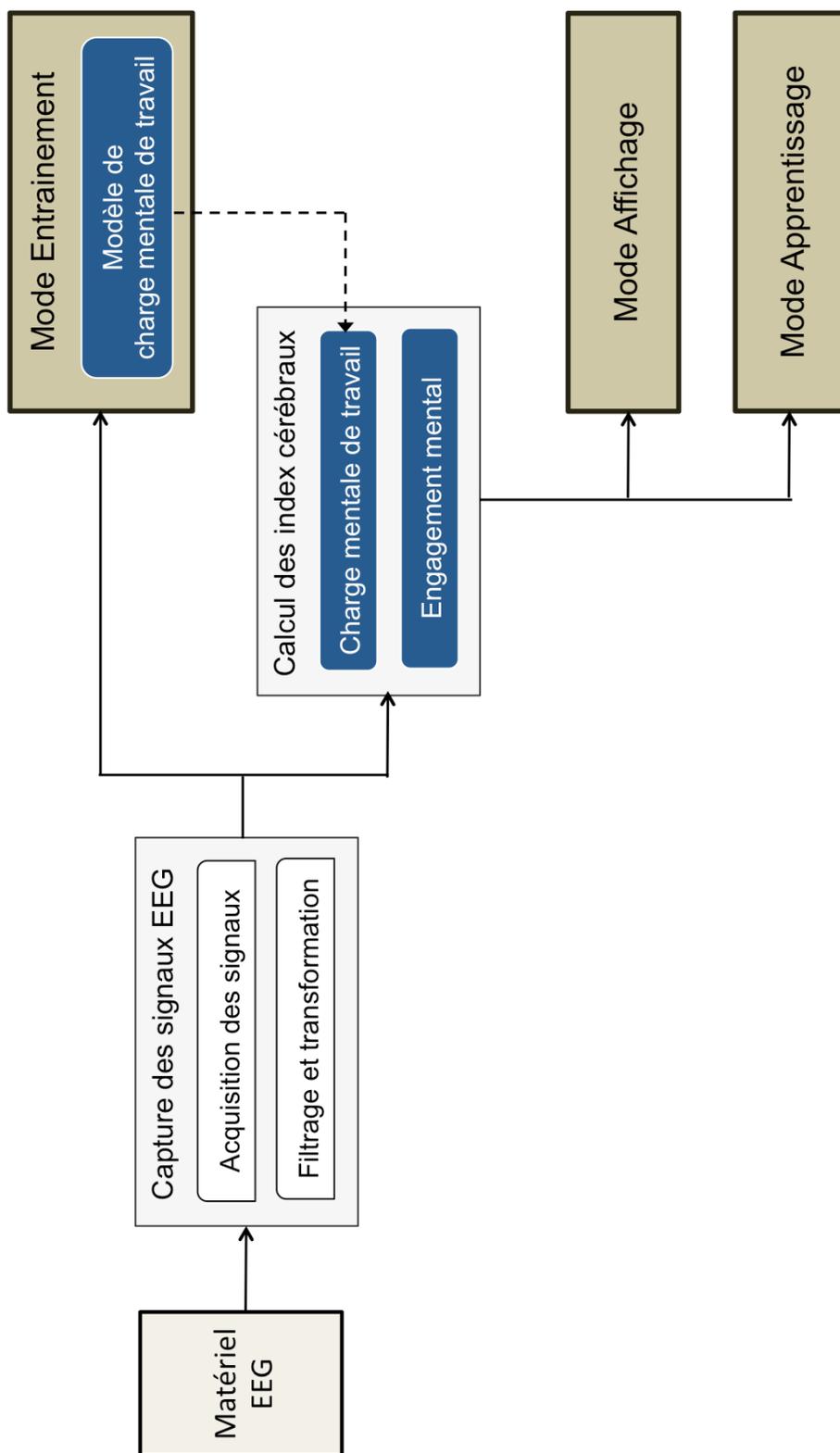


Figure 5.1 – Aperçu du fonctionnement général du système MENTOR.

## 5.2 Capture des signaux EEG

La phase de capture des signaux EEG est la première étape de base sur laquelle repose tout le fonctionnement du système MENTOR. Cette phase fait intervenir un processus *d'acquisition* en temps réel des signaux physiologiques à partir du *dispositif matériel EEG*. Les données EEG recueillies sont ensuite *filtrées* et *transformées*.



Figure 5.2 – Dispositif EEG Emotiv Epoc utilisé. De gauche à droite : le casque EEG sans fil, le récepteur USB et le kit d'hydratation des senseurs.<sup>2</sup>

### 5.2.1 Dispositif matériel EEG

Le dispositif d'acquisition est constitué du matériel EEG Emotiv Epoc, à savoir : un casque EEG, un récepteur USB sans fil (WiFi) et un pack d'hydratation pour les senseurs EEG (voir figure 5.2). Nous avons opté pour ce matériel pour trois raisons principalement. D'abord, le casque peut se connecter sans fil à n'importe quelle machine à travers le récepteur (clé) USB. Ensuite, contrairement à d'autres dispositifs EEG disponibles sur le marché (et en l'occurrence celui que nous avons utilisé dans les deux chapitres précédents), le casque Emotiv est simple à manœuvrer et ne demande pas de préparation ou de configurations matérielles particulières (telles que l'assemblage des fils pour établir une connexion avec un encodeur, le paramétrage du système et des canaux virtuels de réception, etc.). Enfin, le casque Emotiv est facile à porter, léger, et la connexion sans fil utilisée permet à l'utilisateur une plus grande liberté de mouvement.

---

<sup>2</sup> <http://www.emotiv.com/>

Le casque Emotiv contient en tout 16 électrodes (voir figure 5.3), situées selon la norme internationale 10-20 de positionnement des électrodes sur le scalp (Jasper, 1958). Il permet de mesurer jusqu'à 14 régions simultanément (*O1*, *O2*, *P7*, *P8*, *T7*, *T8*, *FC5*, *FC6*, *F3*, *F4*, *F7*, *F8*, *AF3* et *AF4*). Les deux autres électrodes sont utilisées comme des références ; elles correspondent respectivement à la région *P3* (appelée CMS pour « Common Mode Sense ») et la région *P4* (appelée DRL pour « Driven Right Leg »). Le système MENTOR utilise l'intégralité de ces senseurs ; l'enregistrement des données EEG se fait avec un taux d'échantillonnage de 128 Hz.

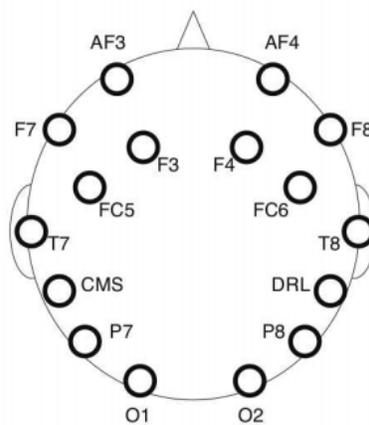


Figure 5.3 – Positionnement des 16 électrodes du casque EEG

Avant chaque utilisation, l'impédance des électrodes doit être réduite au niveau requis par le système Emotiv Epoc en imbibant les senseurs avec une solution saline.<sup>3</sup> Le suivi du niveau d'impédance se fait par l'intermédiaire de l'application « Emotiv TestBench » qui est fournie avec le système Emotiv Epoc (voir figure 5.4). Ce programme permet de vérifier la qualité des signaux EEG des différentes régions mesurées, ainsi que le niveau restant de la batterie du casque dont l'autonomie est de 12 heures d'utilisation approximativement.

---

<sup>3</sup> Cette procédure de diminution de l'impédance est requise dans la plupart des dispositifs EEG. Toutefois contrairement à beaucoup d'autres dispositifs qui nécessitent le recours à une pâte ou un gel sur le cuir chevelu (tel que le matériel Procomp infinity utilisé dans les chapitres 3 et 4), cette solution saline recommandée par Emotiv Epoc est beaucoup moins contraignante pour les utilisateurs.

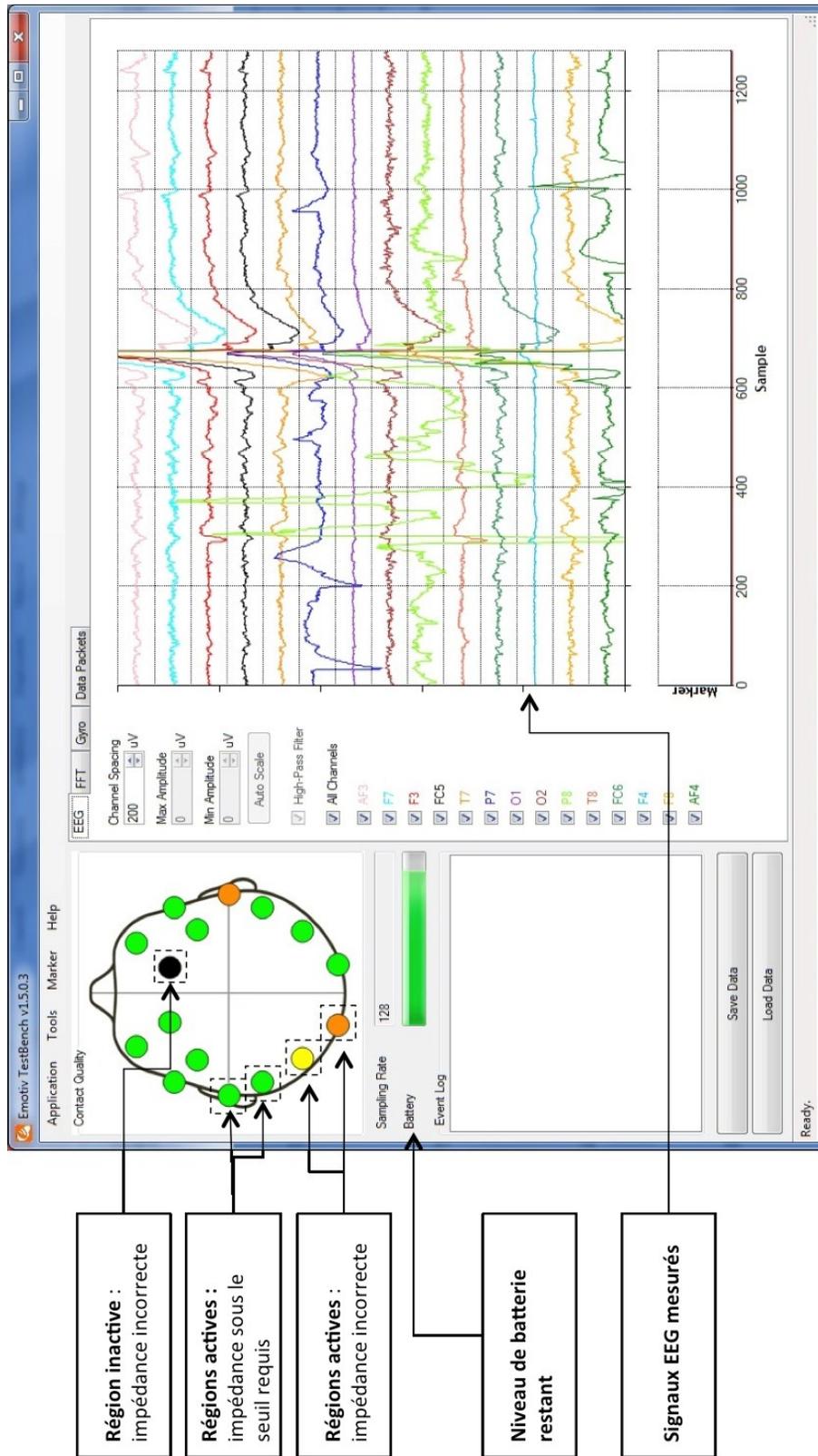


Figure 5.4 – Suivi de l'impédance des électrodes à travers l'application Emotiv TestBench.

## 5.2.2 Acquisition des données EEG

Pour accéder aux données EEG, le système MENTOR utilise l'interface de programmation « Application Programming Interface » (API) de l'édition de recherche du kit de développement du logiciel d'Emotiv (version 2010). Cette interface contient un ensemble de méthodes permettant l'accès à un vecteur de données EEG brutes. Le module d'acquisition de données de MENTOR intègre ces méthodes dans un processus léger (« thread ») qui consulte périodiquement la mémoire tampon dans laquelle se trouvent les données provenant de chacune des 14 régions mesurées. Ces données se présentent sous la forme d'une série temporelle de valeurs de voltage mesurées à partir de l'électrode située sur chaque région. La boucle de lecture du module d'acquisition constitue alors une structure de 14 blocs de données EEG (un bloc par région). Ces blocs EEG sont ensuite transmis au module de filtrage et de transformation pour être prétraités.

La boucle de lecture vérifie à chaque itération l'état de la connexion avec le casque afin de permettre au système MENTOR de détecter et de gérer les pertes de connexion.

## 5.2.3 Filtrage et transformation des données EEG

Le flux de données EEG collectées par le module d'acquisition est décomposé en segments contenant 128 points de données ; ce qui correspond à une seconde d'échantillonnage. Afin de détecter les bruits causés par les mouvements musculaires et les clignements des yeux, une procédure a été implémentée conformément à l'heuristique recommandée par (Prinzel et al., 2000b). Dans cette procédure, l'amplitude des voltages des données EEG du segment relatif à la région *AF4* est examinée (cette région du cerveau étant très sensible aux mouvements musculaires et clignements des yeux). Si cette amplitude dépasse un certain seuil dans plus de 25 % des points du segment, ce dernier est considéré comme bruité, et tout le bloc contenant les 14 segments est automatiquement rejeté. Dans le cas contraire, l'ensemble des segments du bloc de données est ajouté aux autres segments déjà collectés. Ainsi, le système MENTOR prévoit à chaque seconde, un segment de données pour chaque région, si aucun segment bruité n'est détecté (voir figure 5.5 [a]).

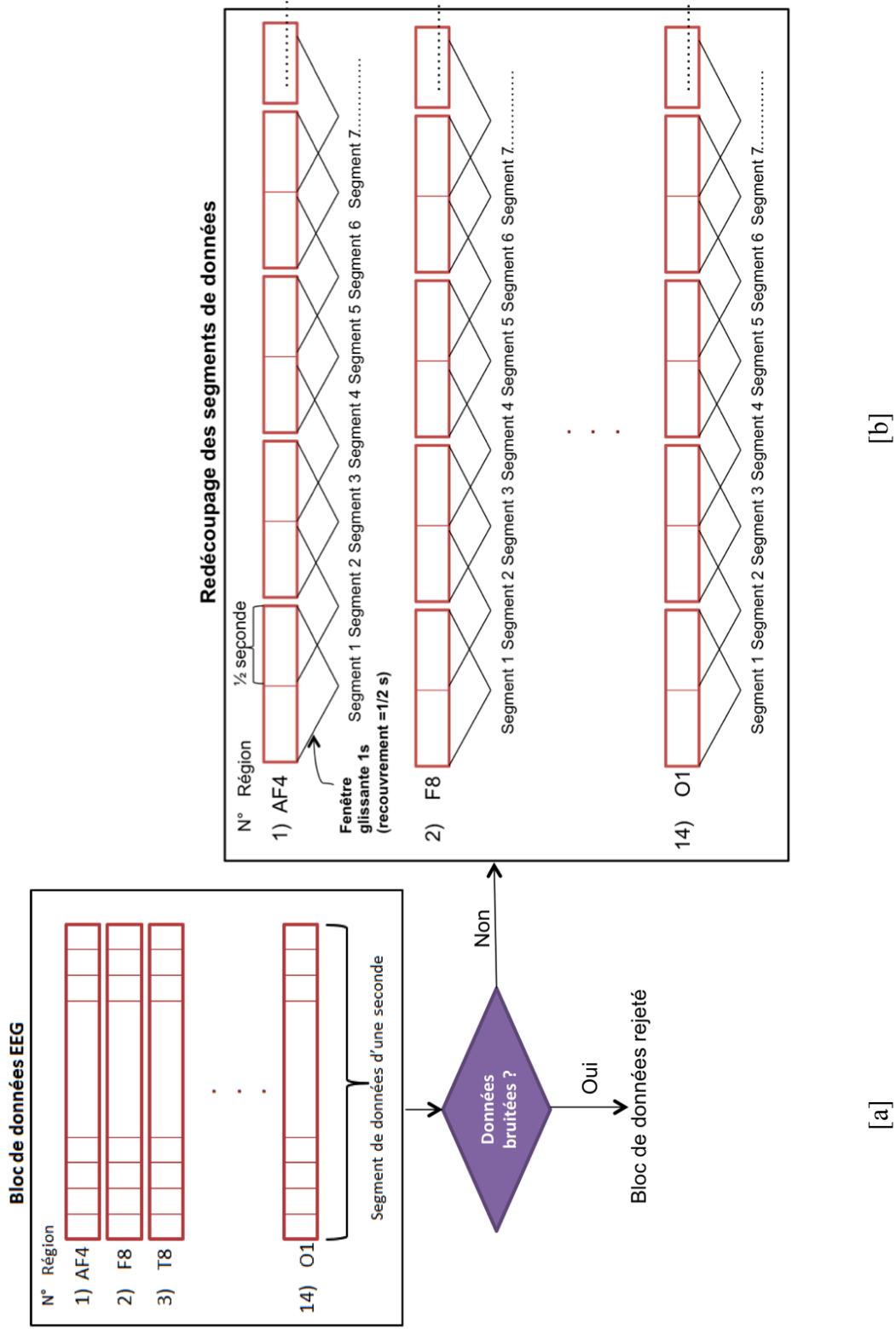


Figure 5.5 – Filtrage et transformation des données EEG : [a] Formation d'un bloc EEG et élimination du bruit.  
[b] Redécoupage des segments EEG.

À l'arrivée de chaque segment, le système procède à son découpage en deux parties d'une demi-seconde (64 points de données). Une fenêtre temporelle glissante avec un recouvrement égal à une demi-seconde est appliquée afin de reconstruire un nouveau segment d'une taille d'une seconde. Ce nouveau segment va regrouper la première moitié du segment découpé avec la deuxième moitié du segment précédent (correspondant à la seconde précédente). La figure 5.5 [b] illustre cette procédure.

Les points de données contenus dans les nouveaux segments sont ensuite multipliés par une fenêtre de Hamming d'une taille de 128 points, et une Transformée de Fourier Rapide (TFR) est appliquée. Cette transformation temps/fréquence résulte en un tableau comportant un spectre de fréquences où la largeur de chaque bande de fréquence est égale à 1 Hz (taux d'échantillonnage 128 Hz / 128 points de données par segment). Enfin, le module de transformation extrait les fréquences allant de 4 Hz à 43 Hz pour chaque région, produisant ainsi une matrice de données de taille égale à (44 x 14) qui regroupe les données EEG de toutes les régions mesurées.

Ainsi, cette première étape de capture permet au système d'accéder aux données EEG en temps réel, de les prétraiter, et de les transformer pour extraire les caractéristiques du signal EEG (sous la forme d'une matrice contenant un ensemble de fréquences par région). Ces données EEG seront ensuite utilisées par les autres modules de MENTOR (voir figure 5.1). En particulier, le mode Entraînement pourra entraîner un modèle de charge mentale de travail propre à chaque apprenant, le module Calcul des index pourra dériver en temps réel l'engagement cérébral et la charge mentale de l'apprenant, le mode Affichage pourra présenter le tableau de bord de l'état mental de l'apprenant et enfin, le module Apprentissage pourra adapter l'environnement d'apprentissage en fonction de l'état mental de l'apprenant. La section suivante est consacrée à la description du mode Entraînement.

### **5.3 Le mode Entraînement de MENTOR**

Un des objectifs du système MENTOR est de fournir une estimation, en temps réel, de la charge mentale de travail de l'utilisateur en se basant sur les données EEG. Cette fonctionnalité permet au système, dans un premier temps, de réaliser un suivi et une analyse

quantitative du comportement de cet indicateur cérébral relativement aux exigences mentales des différentes tâches et activités exécutées, et dans un deuxième temps, d'ajuster son interaction pour l'adapter à l'état mental de l'apprenant.

Le mode Entraînement du système MENTOR est dédié à entraîner et à tester un modèle de charge mentale de travail. Ce modèle sera utilisé par le mode Apprentissage (le module adaptatif de MENTOR) ou le mode Affichage (le tableau de bord de MENTOR) pour dériver l'index de charge mentale de l'apprenant à partir des caractéristiques extraites des données EEG.

Notre système utilise une approche individualisée pour la création de ce modèle : le résultat de la phase d'entraînement est un modèle propre à un apprenant particulier. Par conséquent, ce modèle ne peut pas servir à estimer la charge mentale de travail d'un autre utilisateur ; ce qui rend cette phase d'entraînement obligatoire avant l'utilisation des deux autres modes de fonctionnement du système.

Pour construire le modèle de charge mentale de travail, le mode Entraînement se base sur les mêmes principes de l'approche utilisée dans le chapitre 4 à savoir :

- Le recours à des exercices d'entraînement cérébral faisant intervenir des tâches purement cognitives.
- L'utilisation de l'instrument NASA\_TLX pour l'évaluation subjective de la charge mentale de travail.
- L'utilisation de l'Analyse en Composantes Principales (ACP) pour la réduction de la dimensionnalité des données EEG, et de l'algorithme de régression par processus gaussiens (RPG) pour la création du modèle de charge mentale de travail.

### **5.3.1 Exercices d'entraînement cérébral**

Le mode Entraînement de MENTOR intègre trois types de tests cognitifs à savoir : le test Empan Numérique (EN), le Test Empan Numérique Inversé (ENI), et des exercices de Calcul Mental (CM) : avec principalement des exercices d'addition et de soustraction.

Comme l'objectif de cette phase est de solliciter différents niveaux de charge mentale de travail chez l'apprenant, et que ceci est réalisé à travers la manipulation du niveau de difficulté de la tâche, le système donne la possibilité de paramétrer ces tests. Ainsi, le niveau de difficulté des tests Empan Numérique et Empan Numérique Inversé peut être augmenté (ou diminué) en paramétrant le nombre de chiffres de la séquence à retenir. De la même manière, la difficulté des exercices de Calcul Mental est manipulée en augmentant le nombre des chiffres à additionner ou à soustraire. Le système permet aussi de programmer le nombre de répétitions des différentes tâches à exécuter pour chaque activité cognitive. Par exemple, le mode Entraînement peut être configuré pour répéter six fois l'empan d'une séquence de 5 chiffres, puis quatre fois l'empan de 6 chiffres, ensuite deux fois l'empan de 7 chiffres, etc. Les chiffres utilisés dans les trois tests (EN, ENI et CM) sont générés par le système de façon aléatoire.

### **5.3.2 Collecte des données dans le mode Entraînement**

À l'ajout d'un nouvel utilisateur, un fichier d'identification est créé pour conserver ses informations personnelles (âge, sexe, niveau d'étude, etc.). Pour gérer les données des différents utilisateurs, le système se base sur des identifiants uniques et chaque utilisateur dispose d'un répertoire contenant toutes les données qui le concernent. En particulier, les données collectées pour l'entraînement du modèle de charge mentale de travail sont rassemblées dans différents fichiers ; chaque fichier comprend les données relatives à une activité particulière d'un test cognitif (soit un niveau de EN, ENI ou CM).

La figure 5.6 décrit cette procédure de collecte des données d'entraînement. Dès qu'une activité cognitive est lancée, le processus de capture des signaux EEG décrit précédemment (section 5.2) est automatiquement enclenché. Les données EEG collectées et prétraitées sont directement stockées dans un nouveau fichier propre à cette activité (fichier EEG).

Le processus de capture est arrêté à la fin de chaque activité, et l'apprenant procède à l'évaluation subjective de la charge mentale de travail en utilisant l'instrument de mesure NASA\_TLX (Hart et al., 1988a) administré par le système MENTOR. Un fichier NASA\_TLX sera alors créé pour chaque activité. En outre, une fois que toutes les activités cognitives de

chaque test d'entraînement cérébral sont terminées (c.-à-d. EN, ENI et CM), l'utilisateur doit procéder à la pondération du score NASA\_TLX conformément aux recommandations des concepteurs du test (Hart et al., 1988a) : soit une série de 15 comparaisons des 6 dimensions du NASA\_TLX (exigence mentale, exigence physique, exigence temporelle, performance, effort et niveau de frustration). Les scores pondérés sont alors enregistrés dans les fichiers d'évaluations NASA\_TLX des différentes activités réalisées.

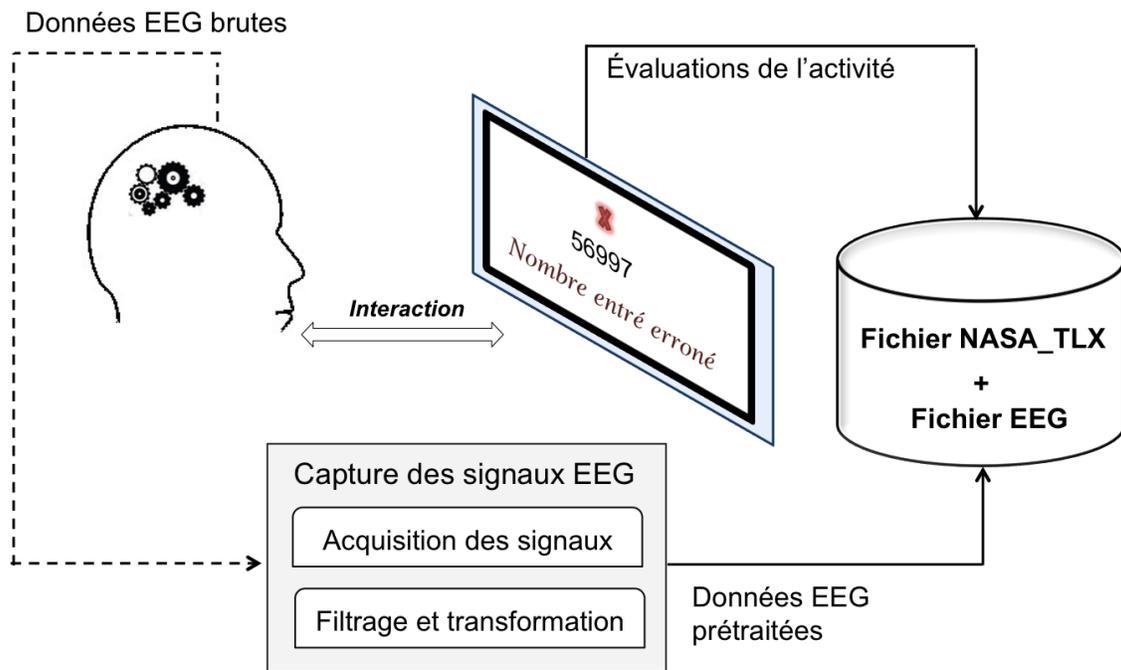


Figure 5.6 – Procédure de collecte des données d'entraînement du modèle de charge mentale de travail.

Cette phase d'exercices d'entraînement cérébral a pour objectif de collecter suffisamment de données pour que le système soit en mesure d'entraîner un modèle individuel de charge mentale de travail. Il est évident que plus il y a de tests et de niveaux variés de difficultés, plus le modèle entraîné est précis. Il est important toutefois de garder un compromis entre le volume des données à collecter et le temps requis par le système pour l'entraînement des modèles. Nous présentons dans ce qui suit la démarche utilisée pour la création de ces modèles.

### 5.3.3 Réduction de la dimensionnalité des données

Comme mentionnée précédemment, la création des modèles individuels de charge mentale de travail dans le mode Entraînement utilise l'algorithme d'Analyse en Composantes Principales (ACP) pour réduire la dimensionnalité des données EEG collectées. Ainsi, dès qu'un utilisateur achève ses exercices cognitifs, il est possible de sélectionner ses fichiers EEG, de les charger dans le système, et d'appliquer l'algorithme d'ACP.

L'algorithme commence par calculer la variance expliquée par chacune des composantes principales extraites à partir des données EEG : une composante correspond à une bande de fréquence de 1 Hz (ou « bin »). Ces composantes sont initialement au nombre de 616, soit le nombre total de bandes de fréquence de 1 Hz prétraitées par le module de capture (44 Hz en tout x 14 régions mesurées). Ces composantes seront affichées à l'écran selon un ordre décroissant du pourcentage de la variance expliquée par chacune d'elles, par rapport à l'ensemble de la matrice des données EEG (la composante d'ordre 1 explique le plus grand pourcentage de variance, la composante d'ordre 2 explique le second plus grand pourcentage et ainsi de suite). Le système calcule également le pourcentage de variance cumulée pour chaque composante.

La variance cumulée pour une composante  $C$  d'ordre  $K$  correspond à la somme de sa variance avec celle des composantes d'ordre inférieur :

$$\text{Variance\_cumulée}(C_k) = \sum_{i=1}^k \text{Variance}(C_i)$$

L'objectif est alors de réduire le nombre de ces composantes en retenant un certain nombre restreint de bandes de fréquences qui expliquent le mieux la variance des données EEG. Pour ce faire, il faut préciser le nombre de composantes que l'on désire conserver. Le système permet de fixer ce nombre de deux manières différentes : en le spécifiant directement, ou indirectement en indiquant un seuil de variance total que l'on désire atteindre. Dans le premier cas, si l'on spécifie un nombre égal à  $n$ , le système sélectionne les  $n$  premières composantes calculées par l'algorithme ; la variance totale expliquée est égale à la variance

cumulée à la *nième* composante. Dans le deuxième cas, si l'on spécifie un seuil de variance égale à  $\nu$ , le système choisit séquentiellement les composantes en commençant par la première, et s'arrête dès que la variance cumulée de la composante dépasse la valeur  $\nu$ .

Une fois le nombre de composantes fixé, le système transforme les données EEG collectées en les projetant selon les composantes sélectionnées, réduisant ainsi leur dimensionnalité. La figure 5.7 récapitule les principales étapes de la réduction de la dimensionnalité du mode Entraînement de MENTOR.

### 5.3.4 Apprentissage du modèle de charge mentale de travail

L'apprentissage du modèle de charge mentale de travail est réalisé en entraînant une fonction de régression sur les données EEG dont la dimensionnalité a été réduite par l'algorithme d'ACP. Les données d'entraînement sont formées de l'ensemble  $\{x_i, y_i\}_{i=1}^n$ , où  $x_i$  est un vecteur EEG correspondant à une seconde d'une activité cognitive, et  $y_i$  la valeur cible correspondant au score NASA\_TLX pondéré attribué par l'utilisateur.

Dans notre approche, nous utilisons l'algorithme de régression par processus Gaussiens (RPG) qui permet d'apprendre une fonction de régression  $f(x_i)$ . Cette fonction calcule la valeur de la charge mentale de travail  $y_i$  avec  $y_i = f(x_i) + \varepsilon_i$ , où  $\varepsilon_i$  est un bruit gaussien défini par une moyenne égale à zéro et une variance  $\sigma_i$ .

Pour un système comme MENTOR qui utilise un modèle de prédiction de charge mentale de travail *individualisé*, le temps requis pour la phase d'apprentissage du modèle prédictif va s'ajouter au temps d'exécution des différentes activités cognitives (exercices d'entraînement cérébral). Cette phase d'entraînement étant faite pour chaque participant, contrairement à l'approche de création de modèles *généralisés*, il est donc important que le temps d'apprentissage de la fonction de régression soit assez rapide. De plus, étant donné que MENTOR utilise ce modèle de régression pour prédire en temps réel la charge mentale de travail de l'utilisateur, le temps de prédiction du modèle doit également être assez rapide.

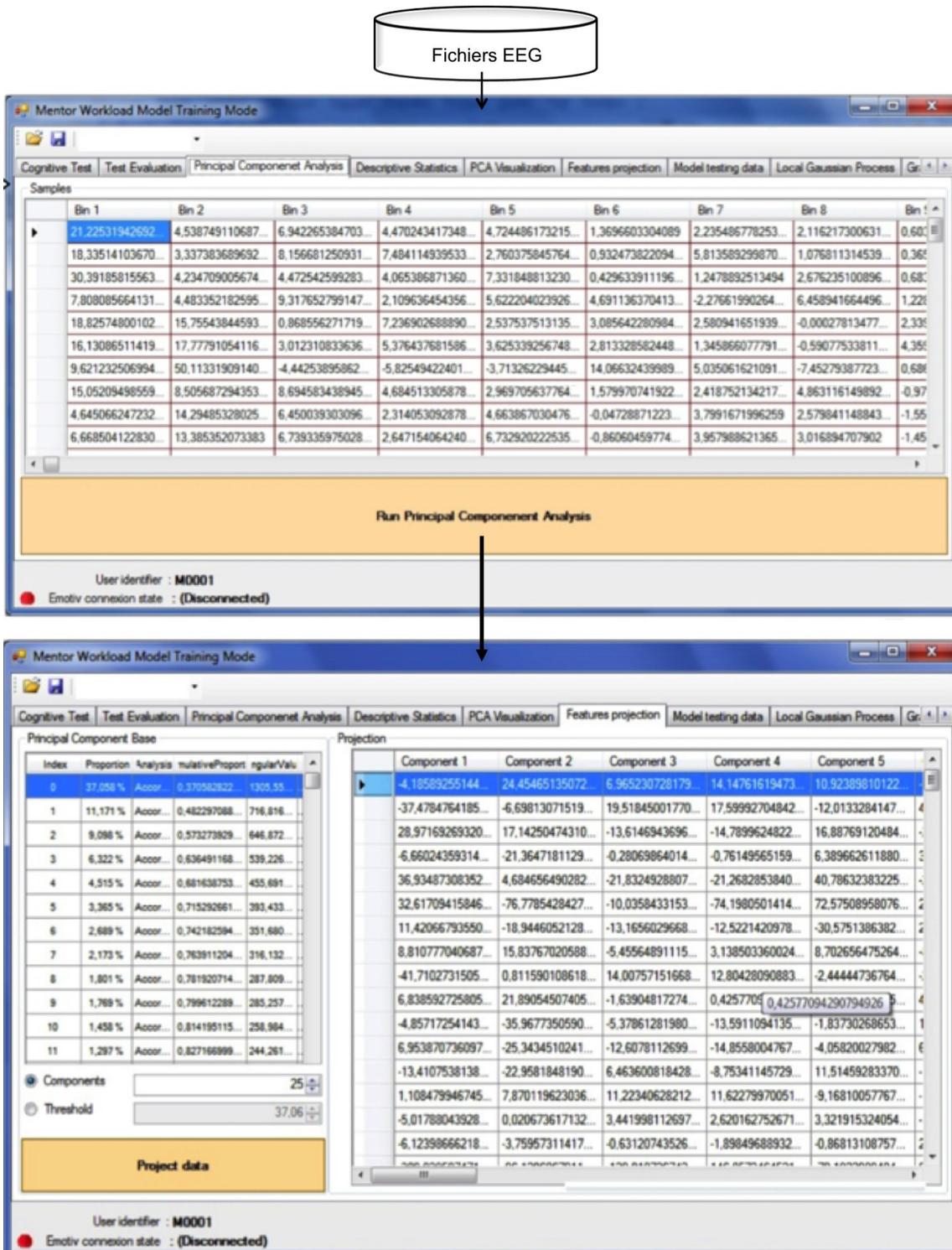


Figure 5.7 – Réduction de dimensionnalité des données EEG dans le mode Entraînement : la première interface (en haut) illustre la matrice des bandes de fréquences de 1 Hz (Bins) extraites du signal EEG, la deuxième interface (en bas) illustre les composantes du signal EEG retenu par l’algorithme d’ACP

En dépit de la grande précision de prédiction des processus gaussiens et de leur capacité à gérer les données de grande dimensionnalité, la complexité de l'algorithme de régression constitue une contrainte majeure. En effet, l'algorithme RPG a une complexité cubique en fonction des exemples d'apprentissage (Rasmussen, 2006). L'utilisation de cette technique dans MENTOR est donc problématique, dans la mesure où les exemples d'entraînement sont assez volumineux (par exemple, pour une phase d'exercices cognitifs d'une heure, un ensemble d'environ 3600 vecteurs de données EEG peuvent être collectés).

Pour remédier à cette limitation et accélérer la phase d'entraînement, nous avons intégré dans notre système une version rapide de l'algorithme RPG, recommandée pour les applications qui fonctionnent en temps réel, appelée régression par processus gaussiens locaux (RPGL) (Nguyen-Tuong et al., 2008); cette version repose sur le principe des techniques d'apprentissage machine dites aussi locales, qui ont pour objectif d'accélérer les procédures d'entraînement et de prédiction.

L'algorithme RPGL *partitionne* les exemples d'entraînement en régions (ou cliques) locales et *entraîne* une fonction de régression sur chacune de ces régions : étape de partition et d'apprentissage des modèles locaux. Puis, l'algorithme calcule une moyenne pondérée des valeurs prédites par les différentes fonctions locales entraînées, pour prédire la valeur cible d'un vecteur de données : étape de prédiction.

### **Étape de partition et d'apprentissage des modèles locaux**

Dans la phase de partitionnement des données, l'algorithme utilise, pour mesurer la distance, la même fonction noyau utilisée pour l'apprentissage des fonctions de régression locales. Nous avons basé le fonctionnement de cet algorithme dans MENTOR sur l'utilisation d'un noyau gaussien qui mesure la distance  $w_k$  avec la fonction suivante :

$$w_k = \exp\left(-\frac{1}{2}(x - c_k)^T W(x - c_k)\right)$$

Avec  $c_k$  le vecteur centre du  $k^{\text{ième}}$  modèle local et  $W$  un hyper-paramètre qui représente la largeur du noyau gaussien. L'algorithme de partition commence avec un modèle local contenant un seul vecteur de données qui est lui-même le centre du modèle. Puis, pour chaque nouveau vecteur de données dans l'ensemble d'entraînement, l'algorithme crée un nouveau modèle local avec un nouveau centre si la distance entre ce vecteur et tous les centres des régions locales est supérieure à un seuil de génération  $w_{gen}$  initialement fixé. Sinon, le vecteur de données est ajouté aux données du modèle local le plus proche, et le centre de cette région est mis à jour en calculant la nouvelle moyenne des vecteurs de données de cette région. À la fin de cette phase de partition, un ensemble de fonctions de régression sont apprises sur des régions contenant un ensemble de données relativement petit ce qui accélère cette phase d'apprentissage.

### Étape de prédiction :

La prédiction de la valeur cible de charge mentale de travail  $\hat{y}$  à partir d'un vecteur de données EEG  $x$  se fait en calculant la moyenne pondérée de chacune des prédictions  $\bar{y}$  faites par toutes les  $M$  fonctions locales apprises. La pondération de cette moyenne se base sur la distance de ce vecteur de données par rapport à chaque région.

La fonction de prédiction est donnée par :

$$\hat{y} = \frac{\sum_{k=1}^M w_k \bar{y}_k}{\sum_{k=1}^M w_k}$$

Avec  $w_k$  la distance du vecteur centre Avec  $c_k$  du  $k^{\text{ième}}$  modèle local.

Ainsi, ces deux étapes de l'algorithme de régression sont implémentées dans MENTOR. La première étape de partitionnement et d'apprentissage du modèle est exclusivement utilisée dans le mode Entraînement alors que la deuxième étape de prédiction est utilisée par le mode

Affichage et le mode Apprentissage pour la prédiction en temps réel de la charge mentale de travail.

Le mode Entraînement permet de sélectionner un ensemble de données d'apprentissage, de réduire leur dimensionnalité avec l'algorithme ACP, d'entraîner le modèle de charge mentale de travail avec les processus gaussiens locaux, et de le tester sur un ensemble de test. Le système permet aussi, lors de cette phase d'entraînement, de contrôler et de spécifier les paramètres et les hyper-paramètres du modèle tels que le seuil de génération  $w_{gen}$ , la largeur du noyau gaussien  $W$ , ou encore la variance du bruit gaussien  $\sigma_i$ .

Pour des fins de test et de comparaison, une version standard des processus gaussiens a été aussi intégrée dans le mode Entraînement. La figure 5.8 illustre une interface de ce mode ce mode de fonctionnement de MENTOR.

Une fois la phase d'apprentissage du modèle de charge mental terminée, le système permet de le sauvegarder afin qu'il soit utilisable par les autres modes du système. Il est à noter que la sauvegarde comprend en plus du modèle de régression utilisé pour la prédiction, les coefficients utilisés par l'algorithme d'ACP, ainsi que le nombre des composantes choisies pour la réduction de la dimensionnalités.

Avant de présenter les autres modes de fonctionnement du système MENTOR, nous décrivons ci-dessous la phase de calcul des index cérébraux.



Figure 5.8 – Interface du mode Entrainement de MENTOR

## **5.4 Calcul des index cérébraux**

Lors du démarrage du système MENTOR, si le mode choisi est le mode Affichage, ou Apprentissage, les vecteurs de données EEG acquis et prétraités par le module de capture seront dirigés vers le module de calcul des index cérébraux. Ce module se chargera d'extraire les index d'engagement et de charge mentale de travail. Cela suppose aussi qu'un modèle de charge mentale de travail a été déjà entraîné pour l'utilisateur.

### **5.4.1 L'index de charge mentale de travail**

Le système récupère le modèle de charge mentale de travail issu de la phase de traitement, et procède au paramétrage de deux fonctions principales :

- 1- La fonction de transformation : cette fonction utilise les paramètres enregistrés de l'algorithme ACP, afin de transformer le vecteur EEG en un ensemble de composantes. En particulier, les coefficients de transformation ainsi que le nombre de composantes retenues permettront de créer un nouveau vecteur de données de taille réduite.
- 2- La fonction de prédiction : Cette fonction infère la valeur de charge mentale de travail selon l'algorithme de régression par processus gaussiens décrit précédemment.

Ainsi, l'extraction de la valeur de charge mentale de travail se fait en deux étapes : la première étape consiste à transformer les données EEG en un vecteur de composantes ; la deuxième étape détermine la valeur de l'index de charge mentale de travail en utilisant la fonction de prédiction.

### **5.4.2 L'index d'engagement**

L'index d'engagement est directement extrait à partir du vecteur de données EEG et ne nécessite pas le recours à un modèle de calcul ou de prédiction spécifique. Les bandes de fréquences  $\theta$  (4-8 Hz),  $\alpha$  (8-13 Hz) et  $\beta$  (13-22 Hz), en provenance de chaque région, sont additionnées afin de produire l'index d'engagement (Pope et al., 1995) :

$$Eng\_Index = \frac{\sum_{i=1}^{14} \beta_i}{\sum_{i=1}^{14} \alpha_i + \theta_i}$$

La phase de calcul des index cérébraux reçoit ainsi un vecteur de données EEG, et produit les index cérébraux de charge mentale de travail et d'engagement chaque seconde. Une moyenne mobile avec une fenêtre glissante de 20 secondes est utilisée, afin de lisser les valeurs des index calculés et de réduire l'impact des fluctuations non significatives dans la visualisation et l'analyse de ces index.

Les index seront ensuite traités selon le mode de fonctionnement du système : d'une façon passive dans le mode Affichage, ou d'une manière active si le système interagit avec l'utilisateur dans le mode Apprentissage.

## 5.5 Le mode Affichage de MENTOR

Le mode Affichage de MENTOR permet de mettre en œuvre de l'un des principes de l'informatique physiologique (« physiological computing »), qui est le suivi de l'état interne d'un utilisateur à partir des informations extraites des signaux physiologiques. L'objectif de ce mode est donc de fournir une sorte de *tableau de bord* (« dashboard ») pour surveiller les valeurs des index mentaux en temps réel. L'état interne (mental) de l'utilisateur est inféré à travers l'analyse de ces index dans le temps.

Cette version non adaptative du système MENTOR a été développée pour visualiser et enregistrer l'évolution des index mentaux d'un apprenant lors de la réalisation d'un ensemble d'activités (par exemple : suivi d'un tutoriel, résolution d'un problème, lecture d'une solution, navigation sur Internet, etc.).

Le mode Affichage utilise un *module d'analyse* permettant d'évaluer l'état mental de l'apprenant selon ses indicateurs cérébraux.

### 5.5.1 Le module d'analyse de MENTOR

Pour évaluer l'état mental d'un apprenant lors d'une activité, le système vérifie le comportement de ses index d'engagement et de charge mentale de travail. Deux états possibles sont considérés pour chacun de ces index, à savoir : *positif* ou *négatif*.

- Index d'engagement positif : le niveau d'engagement de l'apprenant est jugé suffisant par le système durant l'exécution de l'activité.
- Index d'engagement négatif : le système considère que l'apprenant est désengagé pendant l'activité ; le niveau d'engagement de l'apprenant est jugé insuffisant par le système.
- Index de charge mentale de travail positif : le niveau global de charge mentale de travail est jugé adéquat.
- Index de charge mentale de travail négatif : le système détecte une surcharge « overload » ou une sous-charge « underload » mentale durant l'exécution de l'activité.

Le système considère alors que l'état mental d'un apprenant est positif si les valeurs des deux index mentaux sont positives. À l'inverse, si le système détecte un état de désengagement, de surcharge mentale, ou de sous-charge mentale, l'état mental de l'utilisateur est considéré comme négatif.

L'analyse des deux index mentaux lors d'une activité, c.-à-d. déterminer si le comportement d'un index est positif ou négatif, peut se faire selon l'une des deux méthodes suivantes, qui sont communément utilisées dans l'informatique physiologique, à savoir : la méthode de comparaison par rapport à la valeur de référence « Baseline » ou la méthode de la pente.

**La méthode de comparaison par rapport à la valeur de référence « Baseline ».** Ces valeurs de références sont établies en moyennant les valeurs de chaque index sur une période de temps durant laquelle l'apprenant doit être dans l'état le plus neutre possible. L'index d'engagement est considéré comme positif, si sa valeur moyenne lors de toute l'activité est supérieure à la valeur de référence. Autrement, l'index d'engagement est considéré comme

négatif. Pour la charge mentale de travail, les deux états à surveiller sont l'état de surcharge « overload » et de sous-charge « underload » mentale. La moyenne de l'index de charge mentale durant l'activité est comparée à la valeur de référence, plus ou moins un seuil défini par la variance de cette référence. Pour que le comportement de l'index de charge mentale de travail soit considéré comme positif, sa valeur moyenne lors de l'activité ne doit pas dépasser la valeur de référence de cet index plus deux fois sa variance. Sinon un état de surcharge est détecté. Si en revanche cette valeur est inférieure à la valeur de référence moins deux fois la valeur de sa variance, un état de sous-charge mentale est signalé.

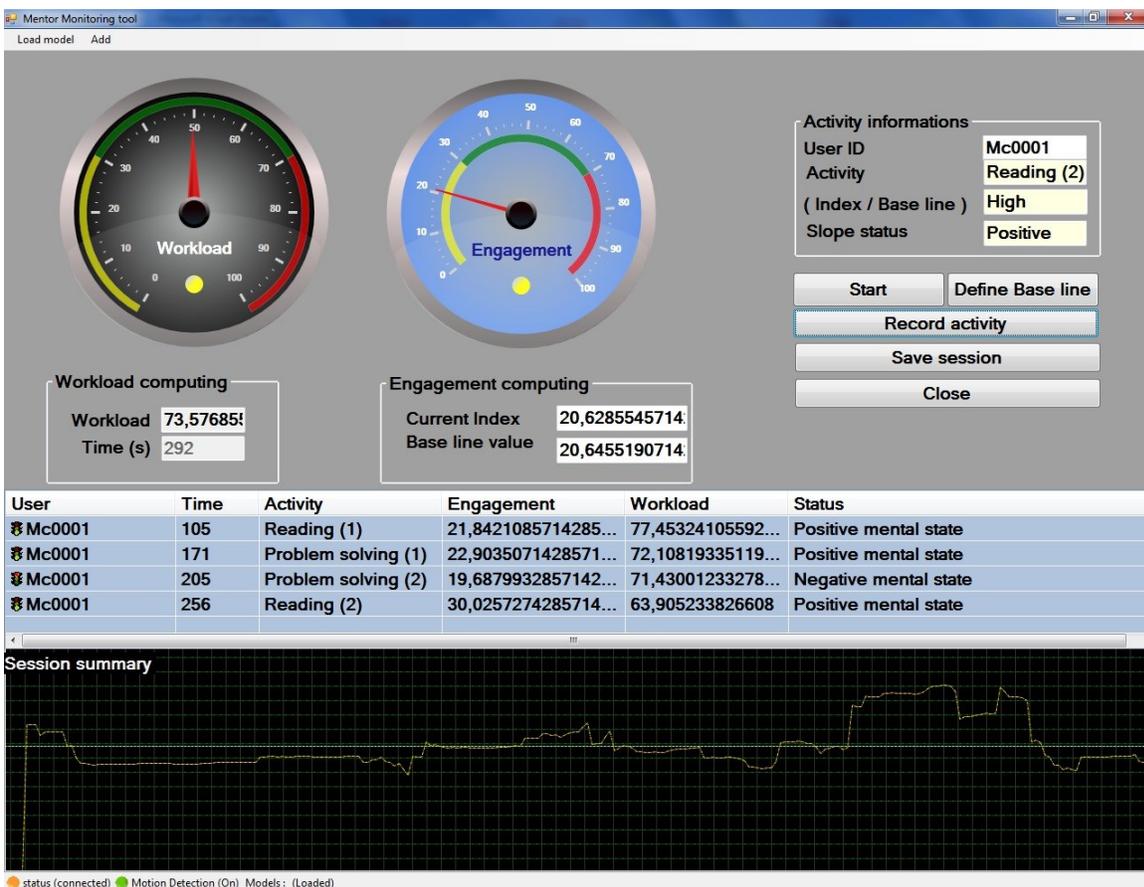


Figure 5.9 – Interface du mode Affichage de MENTOR

Outre son rôle de tableau de bord décrit précédemment, le mode Affichage a été également développé pour des fins d'expérimentation et d'évaluation des index utilisés. Ceci

nous permet entre autres d'examiner les comportements des index en temps réel en fonction de différents types d'activités.

**La méthode de la pente.** Cette approche utilise la méthode des moindres carrés pour déterminer une fonction linéaire  $f(x) = ax + b$  qui minimise la somme des carrés des écarts des valeurs de l'index durant l'activité avec la droite  $y = f(x)$ . Pour l'index d'engagement, si la valeur de la pente  $a > 0$  alors est positive, alors l'état d'engagement est considéré comme positif, sinon l'état d'engagement est considéré comme négatif. Pour la charge de travail, la valeur de cette pente ne doit pas dépasser un seuil d'accroissement  $+\omega$  sinon une surcharge mentale est détectée. La valeur de la pente ne doit pas aussi baisser sous un seuil  $-\omega$  sinon un état de sous-charge mentale est signalé.

La figure 5.9 présente l'interface du mode Affichage de MENTOR. À chaque changement d'activité, le système enregistre les valeurs des index, la durée de cette activité ainsi que l'évaluation de l'état mental de l'apprenant fourni par le système selon une des méthodes d'analyse précédemment décrites. Cette interface nous permet aussi de configurer les paramètres de la méthode d'analyse choisie (durée utilisée pour établir les valeurs de référence, valeur de la borne  $\omega$  pour la méthode de la pente, etc.) et d'observer son impact en temps réel.

Nous décrivons dans ce qui suit le troisième et dernier mode de fonctionnement de MENTOR, qui représente une application directe de notre approche dans le domaine éducatif.

## 5.6 Le mode Apprentissage de MENTOR

Le mode Apprentissage de MENTOR représente la partie adaptative du système. Il offre un environnement d'apprentissage capable d'interagir avec les apprenants en fonction de leurs paramètres mentaux. Contrairement au mode Affichage, qui se limite à une simple analyse du

comportement des index, ce mode dispose d'une logique d'adaptation lui permettant de gérer les étapes de l'apprentissage en fonction de l'analyse de ces deux index.

Notre objectif, à travers ce mode Apprentissage du système MENTOR est de présenter une nouvelle vision des STI augmentés par des capacités d'adaptation à l'état mental d'un apprenant. Les principales innovations que nous proposons se situent au niveau du modèle apprenant et du modèle tuteur (voir figure 5.10).

Le modèle apprenant est enrichi avec des informations sur l'engagement et la charge mentale de travail de l'apprenant ainsi que du module d'analyse décrit précédemment, qui permet d'inférer son état mental. Le modèle tuteur quant à lui dispose d'un module décisionnel qui choisit la meilleure action à entreprendre en fonction de l'état mental de l'apprenant ainsi que des autres connaissances disponibles dans le modèle apprenant. Les décisions du module tuteur sont également basées sur une consultation du curriculum.

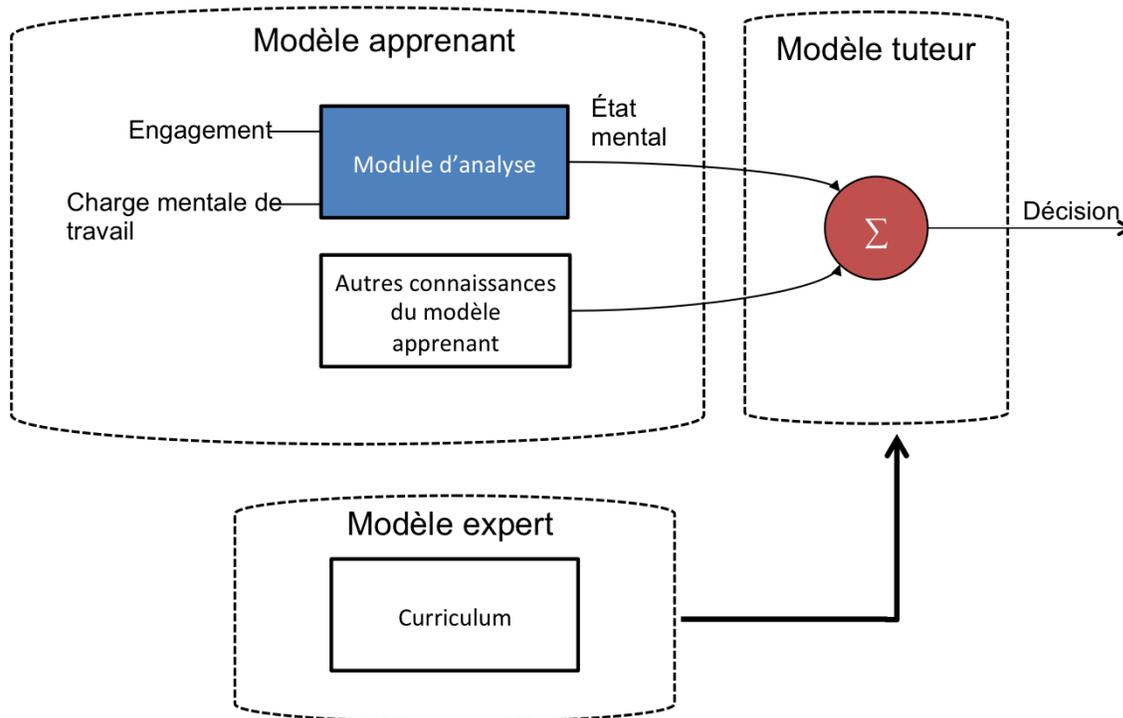


Figure 5.10 – Architecture d'un STI augmentée par les indicateurs d'états mentaux

### 5.6.1 Le curriculum de MENTOR

Le mode Apprentissage du système MENTOR a été créé pour enseigner une leçon sur la notation postfixée. Cette leçon est constituée de quatre parties présentant les notions de base de cette notation :

- 1- La première partie de la leçon présente une série de définitions formelles portant sur les expressions algébriques ainsi que les différents éléments qui la constituent. L'objectif pédagogique est que l'apprenant soit capable de reconnaître et de constituer une expression algébrique correcte à partir d'un alphabet composé d'un ensemble de variables, de constantes, d'opérateur et de fonctions.
- 2- La deuxième partie de la leçon expose la notion d'ambiguïté dans l'évaluation des expressions algébriques en l'absence de parenthèses. La notion de l'ordre de priorité des opérateurs est présentée dans cette partie, dont l'objectif principal est que l'apprenant soit capable d'évaluer correctement une expression. L'apprenant doit également être en mesure de transformer correctement une expression en enlevant ses parenthèses.
- 3- La troisième partie de la leçon porte sur le concept de la notation postfixée. Cette partie explique l'utilité de cette notion dans la désambiguïsation des expressions arithmétiques. Le principe de construction d'une expression postfixée est présenté sous la forme d'un ensemble de règles permettant aux apprenants de comprendre la logique de cette notation. L'objectif de cette leçon est que l'apprenant soit capable de reconnaître une expression postfixée correctement construite sans nécessairement être capable de l'évaluer.
- 4- La quatrième partie de la leçon aborde l'évaluation d'une expression postfixée en utilisant la notion de piles. La structure d'une pile ainsi que ses opérations élémentaires (empiler et dépiler) sont détaillées. L'algorithme d'évaluation est décrit sous forme d'étapes. L'exécution de cet algorithme est présentée sur un exemple détaillé. À la fin de cette partie, l'apprenant doit être capable d'évaluer correctement une expression postfixée.

Le système inclut aussi un ensemble de ressources pédagogiques dont la vocation est de permettre à l'apprenant de mieux maîtriser les notions et les concepts vus dans chaque partie. Deux types de ressources ont été considérés :

**Les questions :** elles représentent un élément fondamental dans l'interaction du système avec l'apprenant. Pour chaque question, le système dispose d'une solution ainsi qu'un ensemble d'astuces et d'indications qui peuvent aider l'apprenant dans sa réflexion. Le système prévoit aussi, pour certaines questions, l'ensemble de toutes les solutions équivalentes, afin d'évaluer et de réagir correctement aux réponses de l'apprenant. L'analyse de la réponse à une question consiste à vérifier l'exactitude de la réponse de l'apprenant et à afficher la bonne solution à l'apprenant si sa réponse est erronée.

**Les exemples résolus :** il s'agit d'illustrer les notions et les concepts vus dans les différentes parties de la leçon à l'aide de questions déjà résolues. L'objectif de ces exemples est d'enseigner la manière de résolution. L'exemple contient alors une question suivie d'une démarche détaillée et expliquée du processus de résolution. La tâche de l'apprenant consiste tout simplement à lire et à comprendre ces exemples. Le système prévoit un ensemble d'exemples qui traitent un aspect différent de chaque partie de la leçon.

Chacune des parties enseignées par le système est soutenue par quatre ressources pédagogiques qui couvrent un sous-ensemble ou la totalité des notions présentes. Chaque ressource peut être donc soit une question que le système pose à l'apprenant, soit un exemple résolu qu'il fournit. Ainsi chaque partie dispose d'un ensemble de quatre questions et de quatre exemples résolus organisés sous la forme d'une paire question-exemple portant sur la même notion. Dès qu'un apprenant termine la lecture d'une partie, le système doit décider du choix du type de la ressource pédagogique à fournir (question ou exemple résolu). Pour chaque partie de la leçon, le système a donc quatre décisions à prendre. Ce choix pédagogique est établi par le module décisionnel du modèle tuteur dont nous allons détailler le fonctionnement dans la section suivante.

Bien qu'il ait été créé pour enseigner une leçon spécifique d'informatique, le mode Apprentissage de MENTOR peut être paramétré pour enseigner n'importe quelle autre leçon.

Le paramétrage se fait à travers un fichier spécifique que le système lit au démarrage de ce mode. Ce fichier contient les informations sur les parties de la leçon (emplacement du fichier, ordre de la partie dans la leçon), ainsi que les informations sur ses différentes ressources pédagogiques (emplacement du fichier, ordre de la ressource relativement à la partie, la ou les réponses si la ressource est une question et les astuces si la ressource est une question). Les différentes parties des leçons ainsi que les exemples résolus doivent être sous le format PDF « Portable document Format ». Les questions peuvent être soit dans un format texte ou RTF « Rich Text Format ».

### **5.6.2 Le module décisionnel de MENTOR**

Le module décisionnel de MENTOR contient la logique d'adaptation du système. Il ajuste le processus d'apprentissage en fonction des informations qu'il reçoit sur l'état mental de l'apprenant durant l'exécution d'une activité d'apprentissage. Par une activité, nous entendons une des trois situations suivantes : la lecture d'une partie de la leçon, la réponse à une question ou la lecture d'un exemple résolu. L'objectif de ce module est de décider de la forme de la ressource pédagogique à fournir. Serait-il mieux pour le processus d'apprentissage de l'apprenant que le système présente un exemple typique détaillé sur la manière de résoudre un problème relatif à un concept clé de la leçon ? Ou serait-il plus approprié de poser le problème sous la forme d'une question, de laisser l'apprenant réfléchir par lui-même en lui donnant quelques indications, et finalement présenter la solution ?

**Dualité exemple-question.** Cette dualité entre exemples et questions a été souvent discutée dans l'étude du processus d'apprentissage (Kalyuga et al., 2001). Les exemples ont tendance à réduire l'effort mental comparé aux questions. L'effort produit par l'apprenant consiste à comprendre et à assimiler les étapes de résolution présentée dans l'exemple résolu. En revanche, si l'on pose le problème sous forme d'une question, l'apprenant doit fournir un effort de résolution en plus d'un autre effort produit pour comprendre et analyser la solution de ce problème dans le cas où celle-ci est différente de la sienne. Dans des domaines tels que la programmation, la physique ou les mathématiques, les exemples résolus peuvent avoir un impact plus important sur l'apprentissage que les problèmes à résoudre, surtout lorsque

l'apprenant est aux premières étapes du processus d'acquisition de compétences (Kalyuga et al., 2001).

D'un autre côté, le recours aux exemples résolus peut également avoir des répercussions négatives sur le processus d'apprentissage. L'apprenant peut ne pas identifier les éléments pertinents de l'exemple et se focaliser plutôt sur des informations inutiles ou secondaires. Un autre phénomène qui peut se produire lorsque l'apprentissage est basé sur les exemples est *l'illusion de compréhension* « illusion of understanding », situation dans laquelle l'apprenant pense avoir compris alors que ce n'est pas vraiment le cas. Ce phénomène fréquent survient lorsque l'apprenant parcourt superficiellement les éléments de l'exemple sans produire l'effort mental requis pour construire un schéma de résolution (Scott, 1971).

La présentation d'un exemple résolu ne garantira pas non plus que l'apprenant sera capable de généraliser à partir de l'exemple montré. En effet, certains apprenants n'engagent pas spontanément un effort mental d'élaboration de l'exemple résolu en essayant d'analyser, reproduire et comparer les étapes de résolution avec celles qu'ils auraient pu faire s'ils avaient à résoudre le problème.

L'avantage de recourir aux questions est donc de réduire ces risques dans le processus d'apprentissage dans la mesure où l'apprenant est amené à raisonner sur le problème, à produire une solution et à la comparer avec la solution du problème. Les questions demeurent toujours un moyen pédagogique efficace pour jauger la progression et le développement des compétences de l'apprenant. Toutefois, utiliser uniquement des questions peut aussi être un obstacle pour le processus d'apprentissage. L'effort mental étant plus important – comparé toujours à la lecture des exemples résolus —, l'apprenant peut facilement être fatigué et surchargé mentalement. De plus, dans certains types de problèmes, l'apprenant peut raisonner d'une manière inefficace ou non optimale. Ainsi même si sa réponse à la question est correcte, l'effort produit est inadéquat et les compétences développées sont incomplètes et insuffisantes. De même, si l'apprenant n'arrive pas à résoudre les problèmes correctement, il peut se démotiver et se désengager. Les exemples résolus dans ce cas permettent de maintenir l'apprenant un peu plus en confiance.

**La logique d'adaptation.** La stratégie pédagogique de MENTOR est fondée sur une analyse continue de l'état d'engagement et de charge mentale de travail de l'apprenant. L'objectif du module décisionnel est de maintenir l'apprenant dans un état mental positif. En d'autres termes, le système vise à garder l'apprenant dans un bon niveau d'engagement et à éviter de le surcharger ou le sous-charger tout au long du processus d'apprentissage. Si le système détecte l'occurrence d'état mental négatif dans une activité, il cherchera alors à rectifier cet état en changeant le type de ressource pédagogique à fournir.

Nous avons construit la logique d'adaptation du module décisionnel de MENTOR sur un ensemble de sept règles de base :

- (1) Si l'état mental de l'apprenant est positif (engagé avec absence de surcharge ou de sous-charge mentale), le système opte pour une question comme prochaine activité à fournir, quelle que soit l'activité courante.
- (2) Si à la fin d'une question, le système diagnostique un état mental négatif (désengagement, surcharge ou sous-charge mentale), il fournit un exemple résolu comme prochaine activité.
- (3) Si à la fin d'un exemple résolu, le système diagnostique un état mental négatif dû à un désengagement ou une sous-charge mentale, il fournit une question comme prochaine activité.
- (4) Si à la fin d'un exemple résolu, le système diagnostique un état mental négatif dû à une surcharge mentale, il fournit encore un exemple résolu comme prochaine activité.
- (5) Si à la fin de la lecture d'une partie de leçon, le système diagnostique un état mental négatif dû à un désengagement ou une sous-charge mentale, il fournit une question comme prochaine activité.
- (6) Si à la fin de la lecture d'une partie de leçon, le système diagnostique un état mental négatif dû à une surcharge mentale, il fournit un exemple résolu comme prochaine activité.
- (7) Quel que soit son état mental, si l'apprenant répond incorrectement à une question, le système fournit un exemple résolu comme prochaine activité.

Le principe général du fonctionnement de MENTOR à travers l'utilisation de ces règles est présenté ci-dessous.

***Décision après la lecture d'une partie de la leçon :*** le système utilise les questions comme moyen principal d'interaction avec l'apprenant. La règle (1) fait que le système opte automatiquement pour une question comme ressource pédagogique à fournir si l'état de l'apprenant est positif. L'hypothèse derrière cette règle est que si l'apprenant lit une leçon en maintenant un état positif, c'est que globalement il n'éprouve pas des difficultés de compréhension des notions présentées. Ainsi le fait de lui donner un problème sous forme de question permet de vérifier son état de maîtrise de ce qui a été montré dans la partie de la leçon.

Si en revanche son état mental est négatif, le système examine la cause de cet état. Si cet état négatif est causé par une surcharge mentale, la règle (6) fait que le système choisit un exemple résolu comme prochaine ressource pédagogique. L'hypothèse est qu'une surcharge est synonyme de difficultés cognitives, d'incertitude ou de manque d'aisance avec ce qui a été montré. L'apprenant investit un grand effort mental pour comprendre ce qui a été présenté. Enchaîner cette activité par un exemple résolu peut l'aider à mieux comprendre avec un minimum d'effort comparé à une question. Nous pensons que si dans ce cas l'apprenant s'engage dans la résolution d'un problème, son état de surcharge peut s'aggraver et ceci peut perturber son processus d'apprentissage. Un exemple résolu lui permet de mettre en pratique les notions qu'il vient de voir, de dissiper ses incertitudes et de le mettre plus en confiance.

À l'opposé, si l'état négatif de l'apprenant est dû à un désengagement ou à une sous-charge mentale, le système opte pour une question par le biais de la règle (5). Dans ce cas, nous supposons que soit l'apprenant maîtrise parfaitement ce qui a été présenté, d'où son manque d'implication mentale, ou que cet état négatif est dû à un désintérêt par rapport à ce qui vient d'être présenté. Dans les deux cas, une question peut être un moyen plus stimulant pour l'apprenant et pourra susciter davantage d'implication mentale.

***Décision après une question :*** à la fin d'une question, si le système ne détecte pas d'état négatif, il opte encore une fois pour une question comme prochaine activité en utilisant la

règle (1). Nous supposons dans ce cas que l'apprenant réagit bien mentalement et que la stratégie basée sur les questions est bien adaptée.

Il est à noter que l'utilisation de la règle (1) est limitée par la règle (7). En cas de mauvaise réponse, le système change automatiquement le type de ressource même si l'état mental de l'apprenant est positif afin de corriger ou anticiper un état négatif dû à une succession de mauvaises réponses.

Ce changement d'activité est également adopté si le système détecte un état mental négatif chez l'apprenant, et ce même si sa réponse est correcte avec la règle (2). L'idée est que si l'apprenant manifeste un état négatif suite à la résolution d'un problème, le changement d'activité peut être bénéfique dans tous les cas. Ainsi, si l'apprenant est surchargé, l'enchaînement avec un exemple résolu peut corriger cet état ou à la rigueur ne pas l'empirer comparé à l'enchaînement avec un autre problème. Si en revanche cet état est causé par un désengagement ou une sous-charge mentale, le changement de type d'activité en lui-même peut constituer une motivation pour l'apprenant et corriger cet état négatif.

***Décision après un exemple résolu :*** à la suite d'un exemple résolu montré à un apprenant, le système opte pour une question comme prochaine étape si l'état mental de l'apprenant est positif en utilisant la règle (1). Cette stratégie est adoptée pour obtenir un effet connu sous le nom *d'effet d'achèvement de problème* « problem completion effect » (Paas, 1992). Cet effet obtenu en faisant suivre un exemple résolu par une question permet généralement d'augmenter les performances d'apprentissage et améliore la motivation de l'apprenant (Sweller, 2003). Pour cette raison, nous décidons d'enchaîner avec une question, plutôt que de continuer avec un exemple résolu.

La règle (3) du module décisionnel indique que le système choisit une question comme prochaine activité, s'il détecte un état négatif chez l'apprenant causé par un désengagement ou une sous-charge mentale. Tout comme pour la décision prise à la fin de l'activité de lecture de la leçon, ce choix est également justifié par l'impact que peut avoir une question dans l'accroissement de l'intérêt et de la motivation de l'apprenant. De plus, ceci peut donner la

possibilité d'obtenir l'effet d'achèvement de problème qui peut augmenter également la motivation et l'intérêt.

Enfin, si à la suite d'un exemple résolu, on détecte un état mental négatif chez l'apprenant causé par une surcharge mentale, MENTOR, par l'intermédiaire de la règle (4), continue de présenter un exemple résolu comme prochaine activité. L'idée derrière cette règle est que si un apprenant éprouve des difficultés cognitives pour comprendre l'exemple, manque de certitudes ou d'aisance par rapport à ce qui a été présenté ou s'il est tout simplement mentalement fatigué, il peut manifester cet état mental de surcharge. Par conséquent, il ne serait pas très judicieux de lui fournir un problème à résoudre vu que ça peut aggraver cet état et influencer négativement le reste du processus d'apprentissage. Un autre exemple résolu peut le mettre plus en confiance, renforcer ses connaissances sans trop de sollicitation mentale. Il est vrai que si cette surcharge est symptomatique d'un manque de confiance, une bonne réponse à une question peut également mettre l'apprenant plus en confiance et augmenter sa motivation (effet d'achèvement de problème). Toutefois, si l'apprenant éprouve des difficultés relativement au problème présenté et n'arrive pas à le résoudre correctement, son état négatif pourra s'aggraver davantage et ceci pourra alors nuire à la suite de son processus d'apprentissage.

Ainsi le mode apprentissage utilise ces règles pour établir des décisions quant aux choix des ressources pédagogiques à fournir à l'apprenant. Vu que le processus d'apprentissage programmé dans MENTOR prévoit quatre ressources pédagogiques après chaque partie présentée de la leçon, il n'existe pas de décision à la quatrième ressource présentée dans la mesure où le système passe à la partie suivante de la leçon, et donc les règles (1, 2, 3 et 4) ne s'appliquent pas. Pour le curriculum que nous avons créé, le module décisionnel établira alors un total de 16 décisions.

Nous présentons dans ce qui suit l'interface du mode Apprentissage du système.

### 5.6.3 Fonctionnement du mode Apprentissage

L'interface du mode Apprentissage représente le moyen par lequel le système fournit le matériel et les ressources pédagogiques. Son rôle est de présenter les différentes parties de la leçon à apprendre ainsi que les différentes ressources pédagogiques.

Lorsqu'une session d'apprentissage débute, cette interface se charge de présenter en ordre les différentes parties de la leçon à proposer. Dès qu'un apprenant termine la lecture d'une partie et demande l'activité suivante, l'interface se charge d'aviser le module décisionnel de ce fait. Ce dernier récupère les informations sur l'état mental de l'apprenant à partir du module d'analyse du système, décide du type de la prochaine ressource pédagogique à fournir (question ou exemple), consulte son emplacement à partir du curriculum et la transmet à l'interface qui se chargera de l'afficher à l'écran. Dans le cas où le type de cette activité est un exemple résolu, l'interface se contente de l'afficher et d'attendre que l'apprenant sollicite la suite des étapes d'apprentissage pour refaire ce même processus de fonctionnement. En revanche, si le type d'activité proposé par le module décisionnel est une question, l'interface affiche la question, donne des astuces dès que l'apprenant les sollicite, récupère la réponse de l'apprenant, informe l'apprenant sur l'exactitude de sa réponse et affiche la solution. De même, dès que l'apprenant termine la lecture de la solution et demande la suite, le même processus est réactivé. La figure 5.11 présente l'interface du mode Apprentissage du système MENTOR.

Le système utilise deux types de fichier journal pour enregistrer les informations sur les différentes activités établies par l'apprenant. Le premier fichier enregistre, chaque seconde, les valeurs des différents index mentaux de l'apprenant. Le deuxième fichier, quant à lui, contient des informations sur la réalisation de l'activité (telles que le type d'activité, la durée, la réponse si c'est une question, etc.).

Nous avons ainsi couvert les principaux aspects fonctionnels du prototype développé. Nous rappelons que la conception de ce système a pour objectif de montrer une nouvelle manière d'adaptation des STI qui tient compte de l'état mental de l'apprenant dans la gestion

du processus d'apprentissage. La suite de ce chapitre est dédiée à l'évaluation de notre système.

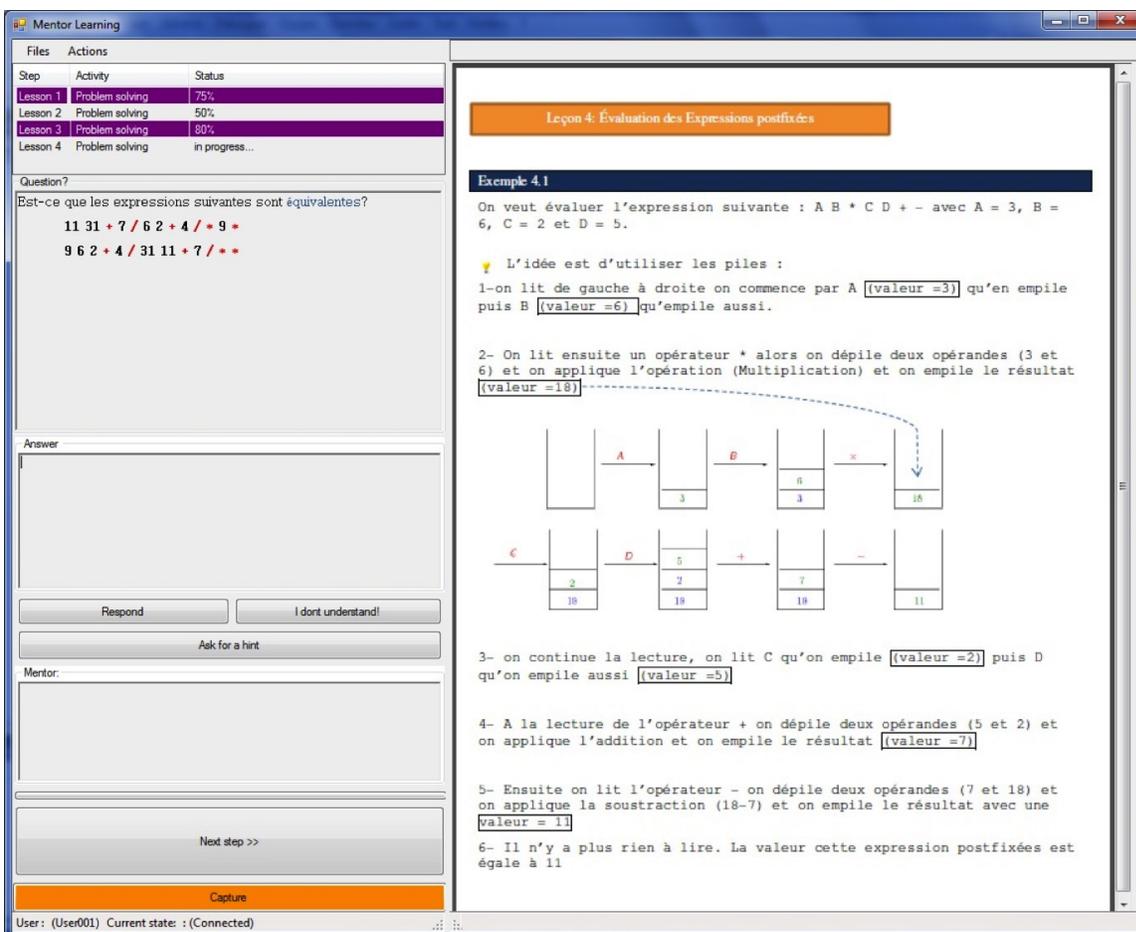


Figure 5.11 – Interface du mode Apprentissage du système MENTOR

## Partie II — Expérience 3 : Évaluation du système MENTOR

Un des objectifs des STI est de fournir un enseignement adapté aux apprenants. Cette adaptation peut s'opérer selon plusieurs considérations (cognitives, pédagogiques, affectives, sociales, etc.) et peut toucher différents aspects de l'interaction du système avec l'apprenant (choix de la prochaine étape d'apprentissage, du type de rétroaction ou « feedback », de l'aide à fournir, etc.).

Le mode Apprentissage du système que nous venons de décrire base toute sa logique d'adaptation, pour la sélection du type de ressources pédagogiques à fournir (question ou exemple), sur l'analyse des index d'engagement et de charge mentale de travail. L'objectif du raisonnement du système est de choisir le type de ressources qui conviendrait le plus à l'état mental de l'apprenant et qui aura le meilleur impact sur le reste du parcours.

Une troisième étude expérimentale a été réalisée afin de vérifier les hypothèses suivantes :

- Dans un premier temps, nous voulons nous assurer que l'intégration de nos index cérébraux dans un STI peut avoir un réel impact sur les performances d'apprentissage. Notre hypothèse de base est que si le système gère la dynamique d'apprentissage en fonction de ces indicateurs mentaux, l'ensemble de ses décisions peut aider l'apprenant à mieux assimiler le contenu à apprendre.
- Dans un deuxième temps, nous pensons que l'intégration de l'état mental dans ce type d'adaptation peut améliorer l'expérience des apprenants vis-à-vis du déroulement de leur apprentissage. En d'autres termes, nous estimons que si le système est conscient des difficultés mentales des apprenants, ceci peut se refléter au niveau de leur satisfaction vis-à-vis du système, mais aussi au niveau de leurs réactions émotionnelles.

Nous décrivons dans la suite la méthodologie expérimentale utilisée pour vérifier nos deux hypothèses.

## **5.7 Méthodologie et protocole expérimental**

Pour mettre en évidence l'impact de l'intégration des indicateurs mentaux dans la gestion des ressources pédagogiques à fournir, notre étude expérimentale se base sur l'utilisation de deux différentes versions du système. La différence entre les deux versions se situe uniquement au niveau de la logique d'adaptation du module décisionnel de MENTOR.

La première version du système laisse intacte la logique d'adaptation en conservant les sept règles de bases décrites dans la section 5.6.2. Le choix du type de ressources à fournir se fait en fonction de l'évolution de l'état mental de l'apprenant. En particulier, le système tend à

privilégier les questions dans le cas où l'état mental est positif. Dans le cas inverse, le choix du type de ressources se fait conformément à des heuristiques visant à ce que l'apprenant regagne un état mental positif.

La seconde version du système ne prend pas en compte les index d'engagement et de charge mentale de travail dans la sélection du type de ressource à fournir. Seule la règle (7) est conservée dans la logique d'adaptation de MENTOR, et les six autres règles sont ignorées. Le principe de cette version est simple : à la fin de la lecture des parties de la leçon, le système choisit de poser une question à l'apprenant. Tant que la réponse à la question posée est correcte, le système continue à adopter cette même stratégie. En revanche, si une mauvaise réponse est donnée, le système choisit immédiatement un exemple résolu comme prochaine activité afin de corriger le raisonnement de l'apprenant. Dès que l'apprenant termine la lecture de la question, le système enchaîne automatiquement avec une question pour augmenter sa motivation et obtenir un effet d'achèvement de problème. Ainsi le paramètre qui peut déclencher une action d'adaptation dans cette version est la réponse incorrecte d'un apprenant.

Les deux versions utilisées partagent un point commun dans leur fonctionnement : si les paramètres d'adaptation sont positifs, les deux versions optent pour les questions pour la prochaine étape. La version adaptative du système représente alors une version augmentée de la deuxième version dans la mesure où en plus de tenir compte, dans son fonctionnement, de l'exactitude de la réponse (à travers la 7<sup>e</sup> règle), elle applique également d'autres actions d'adaptation en fonction des paramètres mentaux.

En résumé, nous allons comparer deux versions du système dont la première utilise, dans sa logique d'adaptation, l'analyse des index mentaux en plus de la réponse de l'apprenant, et la deuxième se base uniquement sur la réponse de l'apprenant. Les deux versions utiliseront, dans le même ordre, exactement les mêmes ressources pédagogiques ; c'est-à-dire que les deux versions auront à choisir entre les mêmes paires de ressources formées d'une question et d'un exemple résolu. La différence entre les deux versions se fera donc uniquement au niveau du choix de la ressource à sélectionner. Ainsi, les deux versions peuvent opter pour le même type de ressources ou pour deux types de ressources différents.

### 5.7.1 Participants

14 participants en tout ont pris part à notre étude. Les participants sélectionnés sont des étudiants de l'Université de Montréal dans un même programme de certificat en informatique appliquée.<sup>4</sup>

Tous les participants ont été mis au courant du déroulement et du contexte de l'expérience, et ont signé un formulaire de consentement. Chaque participant a été affecté d'une manière aléatoire à l'un des deux groupes suivants :

1. Le groupe expérimental ( $N = 7$ ) : ce groupe va utiliser la version adaptative du système MENTOR. Donc les types d'activités d'apprentissage seront adaptés en temps réel selon leurs index cérébraux et leurs réponses.
2. Le groupe de contrôle ( $N = 7$ ) : ce groupe utilise la deuxième version du système qui ne prend en compte dans son adaptation que la réponse de l'apprenant.

### 5.7.2 Déroulement de l'expérience

Pour chaque apprenant, l'expérience s'est déroulée sur deux jours successifs. Lors du premier jour, l'apprenant utilise le mode Entraînement de MENTOR pour la création de son modèle individuel de charge mentale de travail. Dans cette phase, qui dure à peu près une heure, l'apprenant effectue un ensemble de 40 exercices cognitifs selon notre approche décrite dans les sections 5.3.1 et 5.3.2 de ce chapitre. Le tableau 5.1 montre la répartition des différents exercices réalisés.

Une fois les données d'entraînement collectées, nous procédons à la réduction de leur dimensionnalité en utilisant l'Analyse en Composantes Principales (décrite dans la section 5.3.3), puis l'algorithme de régression par processus gaussiens locaux (section 5.3.4) est exécuté pour l'apprentissage du modèle de charge mentale de travail.

---

<sup>4</sup> Nous nous sommes assurés lors de la phase de recrutement que les participants ne disposaient pas de connaissances sur la notation postfixée.

Tableau 5.1 – Exercices cognitifs réalisés durant le premier jour de l’expérience

<b>Empan Numérique</b> (total de 15 exercices)									
Nombre de chiffres	2	3	4	5	6	7	8	9	10
Nombre de répétitions	3	2	2	2	2	1	1	1	1
<b>Empan Numérique Inversé</b> (total de 15 exercices)									
Nombre de chiffres	2	3	4	5	6	7	8	9	10
Nombre de répétitions	3	2	2	2	2	1	1	1	1
<b>Calcul Mental</b> (total de 10 exercices)									
Nombre de chiffres	2	3	4	5	6				
Nombre de répétitions	2	2	2	2	2				

Lors du deuxième jour de l’expérience, l’apprenant utilise une des deux versions du système MENTOR pour suivre la session d’apprentissage prévue. La durée de cette phase est d’approximativement une heure dont 20 à 30 minutes sont consacrées à l’apprentissage de la leçon. Comme on peut le voir dans la figure 5.12, le déroulement du processus expérimental commence par un premier test (phase de pré-test). L’apprenant utilise ensuite MENTOR pour apprendre la leçon. Puis, l’apprenant repasse un deuxième test (phase de post-test). Enfin, une phase de débriefage « debriefing » est établie à la fin de cette expérience, où l’apprenant est amené à donner un bilan de son expérience d’apprentissage et de son interaction avec le système. Deux pauses de 5 minutes séparant le pré-test, l’activité d’apprentissage et le post-test ont été programmées dans notre processus expérimental afin que l’apprenant se relaxe et se vide l’esprit avant d’entamer la prochaine étape.

**Le test.** Les phases de pré-test et de post-test consistent à répondre à une même série de 16 questions portant sur les concepts de la leçon. Chacune des quatre parties de la leçon est

concernée par quatre questions dans ce test. À chaque question, l'apprenant peut répondre par vrai ou faux ou ne pas répondre. Un exemple type de question est de vérifier si deux expressions postfixées sont équivalentes. Le score dans chaque phase de test est calculé en fonction des réponses : une bonne réponse vaut 1 point alors qu'une mauvaise réponse (ou une non-réponse) vaut 0. Ainsi, le score maximal que l'apprenant peut avoir est 16 points et le score minimal est de 0.

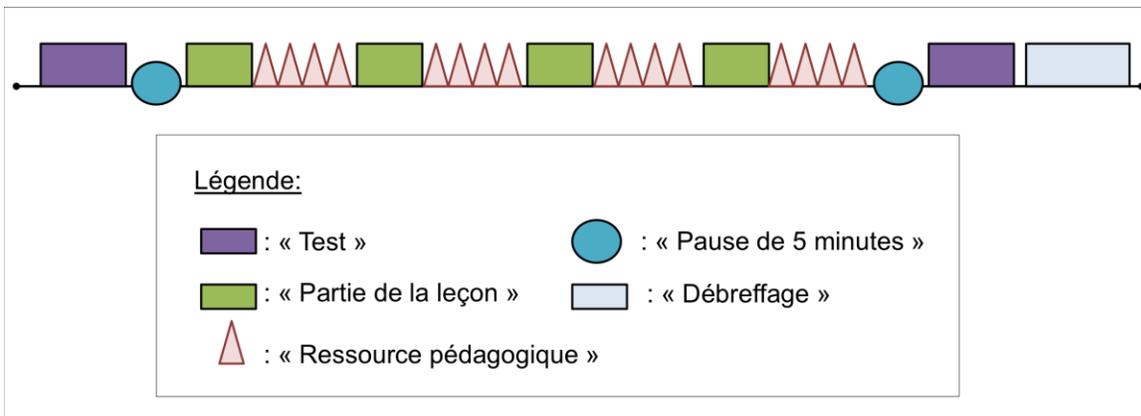


Figure 5.12 – Déroulement du processus expérimental de la deuxième journée des expérimentations

L'objectif du recours au pré-test est de connaître à priori le niveau de connaissances de l'apprenant sur la matière couverte par le cours. Le post-test, quant à lui, détermine le niveau des connaissances acquises après la séance d'apprentissage. Ceci nous permet d'évaluer la progression de l'apprenant (entre le pré-test avant la leçon et le post-test après la leçon). L'utilisation de ce procédé nous permet ainsi de vérifier notre première hypothèse quant à l'impact de l'utilisation des index mentaux dans l'adaptation du processus d'apprentissage, en analysant s'il existe une réelle différence, en termes d'acquisition de connaissances, entre les deux versions du système que nous utilisons dans nos expérimentations.

**Le débriefage.** Pendant la phase de débriefage, l'apprenant évalue la qualité de l'enseignement prodigué par la version du système qu'il a utilisé et donne son bilan de son expérience d'interaction avec l'environnement d'apprentissage. Cette évaluation porte sur le niveau général de satisfaction de l'apprenant par rapport au cours, ainsi que sur sa perception du niveau de pertinence des actions du système.

Bien que les apprenants ne connaissent pas la façon avec laquelle le système choisit les ressources pédagogiques, ils ont été informés avant le début de la séance d'apprentissage que les exemples et les questions utilisés pour supporter les parties de la leçon sont automatiquement choisis par le système. En revanche, les apprenants n'étaient pas informés du fait qu'ils allaient établir ce type d'évaluation à la fin de l'expérience afin de ne pas les influencer.

Pour évaluer **le niveau de satisfaction**, l'apprenant indique sur une échelle d'évaluation de 7 grades allant de 1 (fortement en désaccord) à 7 (fortement en accord) à quel point il est en accord avec l'énoncé suivant :

*« Dans l'ensemble, je suis satisfait de la facilité du déroulement de mon expérience d'apprentissage avec le système »*

La deuxième dimension d'évaluation concerne la perception de l'apprenant de l'utilité ou de la pertinence des exemples et questions fournis par le système. Il s'agit donc d'évaluer la qualité d'adaptation à travers la perception de l'apprenant du degré de pertinence des ressources pédagogiques présentées par le système. Cette dimension concerne ainsi la pertinence (ou la perspicacité) des décisions du tuteur. Pour évaluer **le niveau de pertinence** des activités présentées, une échelle de 7 grades allant de 1 (fortement en désaccord) à 7 (fortement en accord) a été également utilisée. L'apprenant indique à quel point il est en accord avec l'énoncé suivant :

*« Dans l'ensemble, je suis satisfait des activités d'apprentissages (exemples et questions) choisis par le système. Les exemples et les questions sont présentés aux bons moments et m'aident à comprendre. Le choix fait entre poser la question ou présenter un exemple est adapté à mon niveau de compréhension. »*

Les deux échelles que nous venons de présenter ont été extraites et adaptées à partir du questionnaire après-scénario « The After-Scenario Questionnaire » (ASQ) (Lewis, 1991) utilisé pour évaluer la satisfaction de l'utilisateur par rapport à l'utilisation d'un système

informatique. Ce questionnaire comporte les deux échelles que nous venons de décrire en plus d'une troisième qui n'est pas applicable dans notre étude.<sup>5</sup>

Le recours à ces deux échelles nous permet ainsi de vérifier notre deuxième hypothèse qui se rapporte à l'impact de l'utilisation des index mentaux sur le niveau de satisfaction globale des apprenants vis-à-vis du système.

Nous avons également mesuré, lors des différentes activités, les réactions émotionnelles des apprenants afin d'évaluer l'impact de MENTOR sur l'expérience émotionnelle des apprenants lors de leur interaction avec le système.

### **5.7.3 Enregistrement des émotions**

À la fin de chaque activité de la session d'apprentissage, une fenêtre apparaît à l'apprenant pour lui permettre d'identifier son état émotionnel. L'apprenant choisit parmi quatre ensembles d'émotions, l'ensemble qui correspond le plus à son état actuel. Le choix des émotions permet de situer son état émotionnel dans un des quatre quadrants conformément au modèle bidimensionnel valence/activation des émotions utilisé dans les chapitres 3 et 4 :

- (Q1) Valence positive et activation forte
- (Q2) Valence négative et activation forte
- (Q3) Valence positive et activation basse
- (Q4) Valence négative et activation basse

---

<sup>5</sup> Cette troisième échelle se rapporte à la satisfaction par rapport au temps alloué pour la tâche. Nous n'utilisons pas cette dimension dans notre expérience puisque nous n'avons pas de contrainte temporelle.

#### **5.7.4 Paramétrage du module d'analyse**

Pour analyser en temps réel le comportement des index d'engagement et de charge mentale de travail, nous avons eu recours à la méthode de la pente (décrite dans la section 5.5.1 de ce chapitre). Les seuils (maximal et minimal) utilisés pour signaler une surcharge ou une sous-charge mentale sont respectivement : 0.03 (une augmentation de 3 % de l'index correspond à une surcharge mentale) et -0.03 (une diminution de 3 % de l'index correspond à une sous-charge mentale).

### **5.8 Résultats et discussion**

Dans cette partie, nous exposons les résultats de nos expérimentations. Nous commençons d'abord par analyser l'impact de l'utilisation des index mentaux sur l'apprentissage en comparant les performances et la progression des deux groupes considérés (groupe expérimental vs. groupe de contrôle) entre le pré-test et le post-test. Nous analyserons ensuite l'impact de l'utilisation des deux versions du système sur le niveau de satisfaction des apprenants. Puis nous présentons une étude comparative des réponses émotionnelles entre les deux groupes de participants. Nous analyserons enfin le comportement des index d'engagement de charge mentale suite aux interventions initiées par MENTOR.

#### **5.8.1 Analyse des performances d'apprentissage**

Une analyse de variance (ANOVA) mixte 2 (*groupe* : expérimental vs. contrôle)  $\times$  2 (*temps* : pré-test vs. post-test) a été réalisée afin de comparer les performances des apprenants des deux groupes en termes de scores réalisés dans les deux tests. La variable *groupe* est un facteur inter-sujet (pour comparer les scores entre les deux conditions expérimentales), alors que la variable *temps* est un facteur intra-sujet (pour analyser l'évolution ou de la variation des scores entre le pré-test et le post-test pour chaque apprenant de façon individuelle).

Les résultats ont d'abord montré une différence significative des scores réalisés par les apprenants entre le pré-test et le post-test pour les deux groupes (effet principal de la variable

temps) :  $F(1, 12) = 2253.353$   $p < 0.001$ . Ainsi, indépendamment du groupe, et donc de la version du système utilisée, il existe un gain d'apprentissage après l'utilisation du système.

Un effet significatif de l'interaction des deux facteurs (groupe  $\times$  temps) sur les performances d'apprentissage a également été trouvé :  $F(1, 12) = 29.824$ ,  $p < 0.001$ . L'analyse des résultats a révélé qu'au fil du temps, c'est-à-dire entre le pré-test et le post-test, les apprenants du groupe expérimental obtiennent significativement de meilleures performances d'apprentissage comparé au groupe de contrôle. Les moyennes des scores obtenus dans le pré-test et le post-test pour les deux groupes sont listées dans le Tableau 5.2.

Tableau 5.2 – Performances des apprenants dans les deux groupes avant et après le tutoriel. Les valeurs qui diffèrent significativement ont des indices différents ( $p < 0.05$ )

	Pré-test	Post-test
Groupe expérimental		
M	4.86 <sub>a</sub>	13.86 <sub>b</sub>
SD	1.07	0.70
Groupe de contrôle		
M	3.57 <sub>a</sub>	10.71 <sub>c</sub>
SD	1.27	0.95

L'ANOVA établie pour comparer les scores des apprenants dans le groupe expérimental et le groupe de contrôle a révélé qu'il n'existe pas de différence statistiquement significative entre les deux groupes dans le pré-test :  $F(1, 12) = 4.190$ ,  $p = n.s.$  La moyenne générale pour tous les apprenants dans le pré-test était de  $M = 4.21$  ( $SD = 1.31$ ). En revanche, l'ANOVA établie pour comparer les scores des apprenants dans le post-test, a montré que les scores réalisés dans le groupe expérimental étaient significativement plus élevés que le groupe de contrôle :  $F(1, 12) = 50.069$ ,  $p < 0.001$ . La moyenne des scores du groupe expérimental était de  $M = 13.86$  ( $SD = 0.67$ ) contre  $M = 10.71$  ( $SD = 0.95$ ) pour le groupe de contrôle.

Ce résultat confirme notre première hypothèse à savoir : l'intégration de l'index de charge mentale de travail et d'engagement dans la gestion des activités à fournir aux

apprenants peut avoir un impact positif sur les performances d'apprentissage. Les apprenants dont les ressources pédagogiques étaient choisies selon leurs états mentaux arrivaient à fournir en moyenne 86,6 % de bonnes réponses après la séance d'apprentissage. Une augmentation de 22.7 % en termes de performance d'apprentissage a été réalisée en utilisant cette heuristique d'adaptation.

### **5.8.2 Analyse des mesures subjectives**

Une ANOVA a été réalisée pour comparer les évaluations du niveau de satisfaction des apprenants entre le groupe expérimental et le groupe de contrôle. Cette ANOVA a montré une différence presque significative entre les deux groupes :  $F(1, 12) = 4.545, p = 0.054$ . Les apprenants du groupe expérimental ont exprimé plus de satisfaction ( $M = 5.71, SD = 1.604$ ) par rapport au groupe expérimental ( $M = 4.29, SD = 0.756$ ).

Une deuxième ANOVA a été réalisée pour comparer les évaluations du niveau de pertinence des activités proposées par le système dans les deux groupes. Ces évaluations étaient significativement meilleures dans le groupe expérimental ( $M = 5, SD = 1.414$ ) par rapport au groupe de contrôle ( $M = 2.43, SD = 0.787$ ),  $F(1, 12) = 17.673, p < 0.05$ .

Ces résultats viennent ainsi valider le fait que l'augmentation de la logique d'adaptation avec les index d'engagement et de charge mentale de travail a une influence positive sur la satisfaction des apprenants par rapport à leur expérience d'apprentissage de manière générale, et leur perception de la pertinence des actions initiées par le système (dans le choix des ressources pédagogiques).

### **5.8.3 Comparaison des réactions émotionnelles**

Afin de voir également l'impact de notre approche sur les émotions exprimées durant l'interaction des apprenants avec les deux versions de MENTOR, nous avons calculé la proportion (en pourcentage) de présence de chaque quadrant du modèle bidimensionnel valence/activation durant la séance d'apprentissage.

Une ANOVA multivariée (MANOVA) a été réalisée afin d'analyser les différences entre le groupe expérimental et le groupe de contrôle, avec comme variable indépendante le groupe, et comme variable dépendante la combinaison des proportions dans chaque quadrant (Q1, Q2, Q3 et Q4).

Tableau 5.3 – Statistiques descriptives des proportions des émotions pour chaque condition expérimentale

	Groupe de contrôle		Groupe expérimental	
	Moyenne	Écart-type	Moyenne	Écart-type
<b>Q1</b>	0.25	0.19	0.51	0.20
<b>Q2</b>	0.20	0.15	0.07	0.063
<b>Q3</b>	0.20	0.05	0.33	0.13
<b>Q4</b>	0.35	0.19	0.09	0.07

Les résultats de la MANOVA ont montré qu'il existe une différence significative entre les deux groupes en termes de proportions de quadrants :  $F(3, 10) = 8.665, p < 0.05$ . L'analyse spécifique de chaque variable dépendante à part (c.-à-d. de chaque quadrant), en utilisant 4 différentes ANOVA, avec une correction de Bonferroni, a démontré que les deux groupes étaient statistiquement différents sur le plan émotionnel. Les résultats suivants ont été trouvés pour chaque quadrant :

- Q1 (valence positive et activation forte) :  $F(1, 12) = 5.945, p < 0.05$  ; les proportions de Q1 dans le groupe expérimental sont significativement supérieures à ceux du groupe de contrôle.
- Q2 (valence négative et activation forte) : pas de différence significative pour les proportions de Q2 :  $F(1, 12) = 4.101, p = n.s$  ; toutefois, les proportions de Q2 dans le groupe expérimental sont inférieures à ceux du groupe de contrôle.
- Q3 (valence positive et activation basse) :  $F(1, 12) = 5.37, p < 0.05$  ; les proportions de Q3 dans le groupe expérimental sont également significativement supérieures à ceux du groupe de contrôle.

- Q4 (valence négative et activation basse) :  $F(1, 12) = 10.8, p < 0.05$  ; les proportions de Q4 dans le groupe expérimental sont significativement inférieures à ceux du groupe de contrôle.

Les moyennes des proportions des émotions sont données dans le tableau 5.3.

L'étude des émotions rapportées par les apprenants des deux groupes confirme l'impact positif de l'utilisation de la version adaptative de MENTOR augmentée par les indicateurs cérébraux sur l'expérience émotionnelle des apprenants. En effet, avec une forte présence du quadrant Q1 (incluant des émotions telles que l'intérêt, la curiosité, l'enthousiasme et l'inspiration) et du quadrant Q3 (calme, relaxe, satisfaction et sérénité), les apprenants du groupe expérimental avaient tendance à rapporter plus d'émotions positives lorsque le système prenait en compte leurs états mentaux lors du choix de ses interventions. De même, les émotions négatives du quadrant Q4 (ennui et désengagement) étaient significativement plus présentes dans le groupe de contrôle.

L'augmentation de la logique d'adaptation avec les index mentaux semble donc réduire considérablement la présence de ces émotions négatives chez les apprenants (Q4) et favoriser des conditions émotionnelles positives (Q1 et Q3).

#### **5.8.4 Analyse du comportement des index cérébraux**

Dans cette partie, nous nous intéressons à l'analyse du comportement des index mentaux lorsque le système signale un état mental négatif chez l'apprenant. Plus précisément, nous voulons valider l'efficacité de notre stratégie d'analyse du comportement mental dans le déclenchement des interventions de MENTOR. Ainsi, nous voulons comparer comment se comportent ces index, avant et après le signalement d'un état négatif causé par une chute d'engagement ou une augmentation (ou diminution) conséquente de l'index de charge mentale de travail.

Pour ce faire, nous nous intéressons aux apprenants du groupe de contrôle qui interagissent avec la deuxième version du système, dont la stratégie d'adaptation dépend uniquement de l'exactitude de la réponse. La question que nous nous posons alors est la suivante : comment

se comportent les index mentaux d'engagement et de charge mentale de travail chez un apprenant, dans les cas où le module d'analyse détecte un état mental négatif et la version utilisée du système ne la prend pas en compte ? En d'autres termes, nous voulons voir le comportement de ces deux index mentaux, s'il existe une divergence entre les deux stratégies d'adaptation sur le type de ressources à fournir comme prochaine activité.

Nous commençons d'abord, par les cas où le système détecte une chute d'engagement causant un état mental négatif et aurait proposé une ressource pédagogique différente de celle choisie effectivement par la version du système utilisée dans le groupe de contrôle. Ces cas sont survenus 15 fois sur les 112 (7 participants \*16 choix) décisions prises par cette version du système. Une ANOVA à mesures répétées a été réalisée, avec la variable temps comme facteur intra-sujets (avant et après l'activité choisie), le numéro (ID) des participants comme facteur inter-sujet, et la moyenne de l'index d'engagement comme variable dépendante (la mesure répétée).

Les résultats ont montré que la moyenne de l'index de l'engagement était significativement plus élevée avant les points d'interventions détectés par MENTOR, par rapport à la moyenne de l'indice de l'engagement après ces points d'interventions :  $F(1, 8) = 21.156$   $p < 0.05$ . Ceci révèle que le module d'analyse identifie correctement les chutes d'engagement et qu'en l'absence d'adaptation adéquate cet indice continue à chuter.

Les deuxièmes cas concernent les divergences de décisions pour les apprenants du groupe de contrôle lorsque le système détecte une surcharge mentale. Ces cas de figure sont survenus 15 fois au cours de notre expérimentation. La même ANOVA à mesures répétées a été réalisée, mais cette fois avec la moyenne de l'index de charge mentale de travail comme variable dépendante. L'ANOVA a révélé une différence significative dans les moyennes de l'index avant et après l'activité proposée par le système :  $F(1, 5) = 40.866$ ,  $p < 0.05$  ; avec une moyenne significativement plus élevée après le choix d'activité fait par le système. Ainsi, si le système ne choisit pas une activité qui prend en compte cette surcharge mentale, la valeur de cet index risque d'augmenter et l'état négatif de l'apprenant peut empirer.

Les troisièmes cas de figure concernent les situations de sous-charge mentale. De la même manière, nous voulons voir comment se comporte cet index si le module d'analyse détecte une sous-charge et que l'activité proposée par la version du système est différente de celle qu'aurait proposée la version adaptative de MENTOR, qui prend en compte l'état mental de l'apprenant. Ces cas sont survenus 8 fois durant l'expérimentation. Une troisième ANOVA à mesures répétées, avec la moyenne de l'index de charge mentale de travail comme variable dépendante, a révélé une différence significative dans les moyennes de l'index avant et après l'activité proposée par le système :  $F(1, 2) = 33.597, p < 0.05$  ; avec une valeur moyenne de cet index encore plus basse après l'activité proposée par le système.

Ainsi, cette analyse nous a confirmé dans un premier temps que le module d'analyse de MENTOR arrive à détecter correctement les états mentaux critiques de baisse d'engagement, de sous-charge ou de surcharge cognitive ; et dans un deuxième temps, qu'il est nécessaire d'établir des actions d'adaptation qui permettent de corriger ces états.

## 5.9 Synthèse et conclusion

Dans ce chapitre, nous avons présenté notre système tutoriel adaptatif MENTOR « MENTal tuTOR », qui représente notre approche de l'intégration des indicateurs d'états mentaux et plus précisément les index d'engagement et de charge mentale de travail dans un STI. L'objectif de notre approche est de fournir une nouvelle manière de voir le fonctionnement des Systèmes Tutoriels Émotionnellement Intelligents en augmentant leur logique d'adaptation par ces indicateurs cérébraux, en vue d'enrichir la dynamique d'interaction avec l'apprenant et améliorer l'expérience d'apprentissage.

Dans la première partie du chapitre, nous avons présenté les fonctionnalités de notre système ainsi que ses différents modules de traitement. MENTOR permet d'acquérir en temps réel les données EEG de l'utilisateur et de les traiter en vue de les utiliser selon trois différents modes de fonctionnement.

Premièrement, le mode *Entrainement* utilise différents types d'exercices cognitifs paramétrables d'entraînement cérébral pour construire un modèle de charge mentale de travail

pour un nouvel utilisateur. Ce modèle, qui utilise une version rapide (locale) de l'algorithme de régression par processus gaussiens, permet de dériver en temps réel la charge mentale de travail de l'utilisateur à partir de ses signaux EEG.

Deuxièmement, le mode *Affichage* fournit un tableau de bord des indicateurs cérébraux de l'utilisateur, ainsi qu'une évaluation de son état mental, permettant de détecter, via un *module d'analyse* dédié, les états critiques de baisse d'engagement, de surcharge mentale ou de sous-charge mentale.

Bien qu'ils aient été développés pour le contexte d'apprentissage (et plus particulièrement les STI), les modes *Entrainement* et *Affichage* de MENTOR ainsi que le module d'analyse peuvent très bien convenir pour mesurer et évaluer l'état mental de l'utilisateur dans d'autres types d'Interactions Homme-Machine.

Enfin, le troisième mode de fonctionnement de notre système : le mode *Apprentissage* fournit un environnement éducatif qui s'adapte en temps réel en fonction des indicateurs cérébraux de l'apprenant. Le curriculum actuel de notre système porte sur l'enseignement de la notation postfixée. Ce curriculum a ceci dit été conçu de façon paramétrable de manière à permettre l'enseignement de n'importe quelle autre leçon. Ce paramétrage pourra se faire à travers un format de fichier spécifique que le système pourra lire au démarrage. La logique d'adaptation de l'environnement est axée sur une analyse des index d'engagement et de charge mentale de travail des apprenants. En fonction de cette analyse, un *module décisionnel* permet d'adapter les activités d'apprentissage. En particulier, nous nous sommes attaqués à un problème récurrent dans les STI qui est le choix des ressources pédagogiques à fournir à l'apprenant. Notre idée est de choisir le type d'activités qui conviendrait le plus à l'état mental actuel de l'apprenant.

Dans la deuxième partie de ce chapitre, nous avons présenté l'étude expérimentale que nous avons réalisée pour évaluer notre système. Nous avons pu démontrer les résultats suivants :

- (1) Le système MENTOR permet d'améliorer significativement les performances des apprenants en terme de gains d'apprentissage avant et après l'utilisation du

l'environnement d'apprentissage, par rapport un groupe de contrôle, où une deuxième version du système, qui ne tient pas compte des indicateurs mentaux des apprenants, a été utilisée.

- (2) La logique d'adaptation selon les états mentaux de MENTOR a également un impact positif sur l'expérience d'interaction des apprenants en terme de satisfaction (par rapport au groupe de contrôle), mais aussi en terme de réactions émotionnelles positives lors de l'utilisation de l'environnement d'apprentissage.
- (3) Le module d'analyse de MENTOR a pu correctement détecter et corriger les états mentaux critiques de baisse d'engagement, de sous-charge et de surcharge mentale.

Dans le prochain chapitre, nous allons clôturer cette thèse en résumant les principales contributions de nos recherches et en présentant nos perspectives pour nos travaux futurs.



## Chapitre 6

# Conclusion

### 6.1 Contributions

La première contribution de notre recherche a été de remédier aux limites actuelles des STEI dans la modélisation des réactions émotionnelles de l'apprenant à travers l'intégration d'indicateurs d'engagement et de charge mentale de travail. Pour cela, nous avons réalisé deux études expérimentales. Dans la première, nous avons étudié le comportement d'un index d'engagement développé dans la NASA pour évaluer le niveau d'attention des pilotes dans le cadre de simulations de vols (Pope et al., 1995). Cet index extrait à partir des signaux EEG, a été analysé dans un environnement d'apprentissage de type jeu-concours. Les états émotionnels des apprenants déterminés à partir de senseurs physiologiques (la pression du volume sanguin/artériel et de la conductivité de la peau) ont été associés avec les différentes valeurs de l'index. Les résultats de cette étude nous ont permis de montrer les points suivants :

- Cet index d'engagement peut aider les STEI à suivre le comportement des apprenants dans la mesure où les apprenants qui étaient engagés sur la durée de l'expérience avaient tendance à mieux déceler les pièges dans les questions et donc à avoir de meilleurs résultats que ceux qui n'étaient pas engagés selon cet index.

- L'index d'engagement pouvait également renseigner sur les états émotionnels des apprenants. En particulier, les valeurs les plus élevées de cet index ont été enregistrées avec les émotions avec forte activation selon l'approche dimensionnelle des émotions.

Dans la deuxième étude expérimentale, nous avons utilisé une approche de modélisation de la charge mentale basée sur des modèles d'apprentissage machine (principalement l'analyse en composantes principales pour la réduction de la dimensionnalité des données EEG et les processus gaussiens pour la régression) qui utilisent les plages de fréquences du signal EEG comme entrée ainsi que le score du NASA\_TLX comme cible. Notre approche consistait à entraîner dans un premier temps le modèle sur les données issues d'activités purement cognitives (mémorisation, calcul mental, etc.) et de le valider, dans un deuxième temps, sur les données relatives à la session d'apprentissage. L'index d'engagement ainsi que les états émotionnels des apprenants ont été également enregistrés au cours de cette session d'apprentissage. L'analyse des résultats nous a permis de tirer les conclusions suivantes :

- Notre approche pour la modélisation de la charge mentale de travail à partir des données EEG peut être pertinente dans un contexte d'apprentissage. Plus précisément, l'index que nous avons développé a montré un comportement cohérent avec la conception des activités d'apprentissages, les mesures objectives et subjectives de la charge mentale de travail et les caractéristiques des apprenants.
- L'analyse de l'index d'engagement dans un environnement d'apprentissage plus riche (avec des leçons et des problèmes à résoudre) comparé à l'environnement du jeu-concours a également reconfirmé ses capacités à renseigner sur le comportement des apprenants.
- L'analyse de l'index d'engagement et de charge mentale de travail peut également renseigner sur l'état émotionnel de l'apprenant en termes de valence et d'activation. Par exemple l'état émotionnel caractérisé par une valence positive et une activation positive (état pouvant contenir des émotions comme la joie, le plaisir, etc.) se caractérise par un fort niveau d'engagement ainsi qu'une faible valeur de l'index de charge mentale de travail.

Notre deuxième contribution a été de proposer une approche d'intégration de ces deux indicateurs en temps réel dans un STEI. Pour cela nous avons développé un système nommé MENTOR qui dispose de trois modes de fonctionnements :

- Le mode Entraînement: permet d'entraîner un modèle de charge mentale conformément à l'approche utilisée dans notre étude expérimentale. Plus précisément, ce mode permet de lancer des exercices d'entraînement cérébral faisant intervenir des tâches purement cognitives. L'évaluation subjective de ces exercices se fait à travers l'instrument NASA\_TLX. Ce mode contient également une implémentation de l'algorithme d'analyse en composantes principales et de l'algorithme de régression par processus gaussiens pour l'entraînement du modèle de charge mentale de travail à partir des données EEG.
- Le mode affichage : fournit un tableau de bord permettant de détecter et de suivre les valeurs des index mentaux en temps réel. Le mode affichage permet de discrétiser l'index d'engagement en deux états (engagé et désengagé) et l'index de charge mentale de travail en trois (surchargé, neutre et sous-chargé). En fonction des états des indicateurs cérébraux, ce mode peut inférer l'état interne global (positif ou négatif) de l'apprenant.
- Le mode apprentissage : représente système tutoriel capable de détecter de leurs paramètres mentaux. Contrairement au mode Affichage, qui se limite à une simple analyse passive du comportement des index, ce mode dispose d'une logique d'adaptation lui permettant de gérer les étapes de l'apprentissage en fonction de l'analyse de ces deux index.

À travers ces trois modes du système MENTOR nous avons développé un système fonctionnel dont l'objectif est de fournir un exemple de conception globale d'un STEI qui intègre ces paramètres mentaux. Nous avons également montré comment ces index peuvent être contribués dans l'amélioration du fonctionnement des systèmes tutoriels, en s'attaquant à un problème récurrent dans les STI qui est le choix des ressources pédagogiques à fournir à l'apprenant. Pour cela, nous avons mis en place une logique d'adaptation qui se base sur un

ensemble de règles simples qui permettent de choisir la ressource qui conviendrait le mieux en fonction de l'état d'engagement et de charge mentale de travail.

Une troisième étude expérimentale comparant deux versions du système MENTOR (une version qui s'adapte en fonction des états mentaux contre une version qui s'adapte uniquement aux performances des apprenants) nous a permis de démontrer les points suivants :

- L'intégration des indicateurs d'engagement et de charge mentale de travail permet d'améliorer significativement les performances d'apprentissage.
- L'adaptation en fonction de ces indicateurs cérébraux permet d'avoir un impact positif sur l'expérience d'interaction des apprenants en termes de satisfaction et d'états émotionnels ce qui répond à un objectif principal des STEI.
- Le module d'analyse de MENTOR a pu correctement intervenir sur les états mentaux critiques de baisse d'engagement, de sous-charge et de surcharge mentale.

## **6.2 Travaux futurs**

Le développement de notre approche à travers la mise en place du système MENTOR nous a ouvert un certain nombre de voies pour des travaux futurs. En voici quelques exemples :

- Pour des raisons de simplicité, nous avons limité le fonctionnement du module décisionnel du système MENTOR à un choix entre deux types d'activités qui représentent des éléments fondamentaux de l'enseignement à savoir les exemples et aux exercices. Nous pensons toutefois qu'il est possible d'augmenter la logique d'adaptation du système pour prendre en compte d'autres types d'actions et d'activités.
- Nous envisageons également d'intégrer dans le système MENTOR des fonctionnalités qui lui permettent d'analyser et de gérer en parallèle et en temps réel les indicateurs de mentaux d'un groupe d'apprenants lors d'un apprentissage collaboratif en classe ou en ligne.
- Une autre voie qui nous semble intéressante est celle d'expérimenter le système MENTOR dans un domaine différent tel que les jeux vidéo ou la robotique.

- Nous projetons également d'améliorer le fonctionnement du mode affichage (c'est-à-dire le mode passif) de MENTOR par l'intégration de mécanismes de suivi et de raisonnement qui permettent de retracer les difficultés mentales rencontrées par l'apprenant au cours de la tâche de la tâche réaliser. Par exemple, au cours d'une tâche de lecture d'un texte, le système doit être capable de déterminer les moments ou même les parties du texte dans lesquelles l'apprenant a éprouvé une certaine difficulté.
- Nous voulons également améliorer la modélisation de la charge mentale de travail en comparant plusieurs types d'algorithmes d'apprentissage machine.

# Bibliographie

- Adamo-Villani, N., & Wright, K. 2007. *SMILE: an immersive learning game for deaf and hearing children*. Paper presented at the ACM SIGGRAPH 2007 educators program.
- Aïmeur, E., Frasson, C., & Lalonde, M. 2001. The role of conflicts in the learning process. *ACM Journal of SIGCUE Outlook, Special interest Group in Computer Uses in Educational Psychology Review*, 27(3), 12-27.
- Alexander, S., Sarrafzadeh, A., & Fan, C. 2003. *Pay attention! The computer is watching: Affective tutoring systems*. Paper presented at the World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education.
- Allanson, J., & Fairclough, S. H. 2004. A research agenda for physiological computing. *Interacting with Computers*, 16(5), 857-878.
- Anderson, E. W., Potter, K. C., Matzen, L. E., Shepherd, J. F., Preston, G. A., & Silva, C. T. 2011. *A user study of visualization effectiveness using EEG and cognitive load*. Paper presented at the Computer Graphics Forum.
- Anderson, J. R. 1993. *Rules of the Mind*. Philadelphia: Lawrence Erlbaum Associates, Inc.
- Anderson, J. R., Boyle, C. F., Corbett, A. T., & Lewis, M. W. 1990. Cognitive modeling and intelligent tutoring. *Artif. Intell.*, 42(1), 7-49. doi: [http://dx.doi.org/10.1016/0004-3702\(90\)90093-F](http://dx.doi.org/10.1016/0004-3702(90)90093-F)
- Ang, J., Dhillon, R., Krupski, A., Shriberg, E., & Stolcke, A. 2002. Prosody-based automatic detection of annoyance and frustration in human-computer dialog. *7th International Conference on Spoken Language Processing (ICSLP 2002)* (pp. 2037–2040). Denver, Colorado.
- Arroyo-Palacios, J., & Romano, D. 2008. *Towards a standardization in the use of physiological signals for affective recognition systems*. Paper presented at the Measuring Behavior.
- Arroyo, I., Cooper, D. G., Bursleson, W., Woolf, B. P., Muldner, K., & Christopherson, R. 2009. Emotion Sensors Go To School. *Artificial Intelligence in Education* (pp. 17-24): IOS Press.
- Arroyo, I., Ferguson, K., Johns, J., Dragon, T., Meheranian, H., Fisher, D., et al. 2007. Repairing Disengagement With Non-Invasive Interventions. *Artificial Intelligence in Education* (pp. 195-202 ). Amsterdam, The Netherlands: IOS Press.

- Ashby, F. G., & Isen, A. M. 1999. A neuropsychological theory of positive affect and its influence on cognition. *Psychological review*, 106(3), 529.
- Asteriadis, S., Tzouveli, P., Karpouzis, K., & Kollias, S. 2009. Estimation of behavioral user state based on eye gaze and head pose—application in an e-learning environment. *Multimedia Tools and Applications*, 41(3), 469-493.
- Bachorowski, J.-A., & Owren, M. J. 1995. Vocal Expression of Emotion: Acoustic Properties of Speech Are Associated with Emotional Intensity and Context. *Psychological Science*, 6(4), 219-224. doi: 10.2307/40063021
- Bailey, B. P., & Iqbal, S. T. 2008. Understanding changes in mental workload during execution of goal-directed tasks and its application for interruption management. *ACM Trans. Comput.-Hum. Interact.*, 14(4), 1-28. doi: 10.1145/1314683.1314689
- Baker, R. S. J. d., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. 2010. Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, 68(4), 223-241. doi: 10.1016/j.ijhcs.2009.12.003
- Ball, G., & Breese, J. 2000. Emotion and personality in a conversational agent. In J. Cassel, J. Sullivan, S. Prevost & E. Churchill (Eds.), *Embodied conversational agents* (pp. 189-219): MIT Press.
- Banda, N., & Robinson, P. 2011. Multimodal Affect Recognition in Intelligent Tutoring Systems In S. K. D'Mello, A. C. Graesser, B. r. Schuller & J.-C. Martin (Eds.), *4th International Conference on Affective Computing and Intelligent Interaction (ACII 2011)* (pp. 200-207): Springer Berlin Heidelberg.
- Baron-Cohen, S., Wheelwright, S., Jolliffe, & Therese. 1997. Is there a " language of the eyes"? Evidence from normal adults, and adults with autism or Asperger syndrome. *Visual Cognition*, 4(3), 311-331.
- Barrett, L. F., & Russell, J. A. 1998. Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology - PSP*, 74(4), 967-984.
- Barrouillet, P. 1996. Ressources, capacités cognitives et mémoire de travail: postulats, métaphores et modèles: La charge mentale. *Psychologie française*, 41(4), 319-338.
- Beal, C. R., & Lee, H. 2005. Creating a pedagogical model that uses student self reports of motivation and mood to adapt ITS instruction. *Workshop on Motivation and Affect in Educational Software, at the 12th International Conference on Artificial Intelligence in Education* (pp. 39-46).
- Beal, C. R., Qu, L., & Lee, H. 2006. *Classifying learner engagement through integration of multiple data sources*. Paper presented at the Proceedings of the National Conference on Artificial Intelligence.
- Beale, R., & Creed, C. 2009. Affective interaction: How emotional agents affect users. *International Journal of Human-Computer Studies*, 67(9), 755-776.
- Beck, J. E. 2005. Engagement tracing: using response times to model student disengagement. *Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology* (pp. 88-95): IOS Press.
- Berka, C., Levendowski, D. J., Cvetinovic, M. M., Petrovic, M. M., Davis, G., Lumicao, M. N., et al. 2004. Real-time analysis of EEG indexes of alertness, cognition, and memory acquired with a wireless EEG headset. *International Journal of Human-Computer Interaction*, 17(2), 151-170.

- Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V., et al. 2007. EEG Correlates of Task Engagement and Mental Workload in Vigilance, Learning, and Memory Tasks. *Aviation, Space, and Environmental Medicine*, 78(5), B231-B244.
- Berka, C., Levendowski, D. J., Ramsey, C. K., Davis, G., Lumicao, M. N., Stanney, K., et al. 2005. Evaluation of an EEG workload model in an Aegis simulation environment. *Defense and Security* (pp. 90-99): International Society for Optics and Photonics.
- Bradley, M. M., & Lang, P. J. 1994. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1), 49-59.
- Brody, L. 1999. *Gender, emotion, and the family*. Massachusetts: Harvard University Press.
- Brusilovsky, P., & Pesin, L. 1994. *ISIS-Tutor: An adaptive hypertext learning environment*. Paper presented at the Proceedings of JCKBSE.
- Brusilovsky, P., Schwarz, E., & Weber, G. 1996. *A tool for developing adaptive electronic textbooks on WWW*. Paper presented at the WebNet.
- Burleson, W., & Picard, R. W. 2004. Affective Agents: Sustaining Motivation to Learn Through Failure and a State of Stuck. *Social and Emotional Intelligence in Learning Environments, Workshop In conjunction with the 7th International Conference on Intelligent Tutoring Systems*. Maceio - Alagoas, Brasil.
- Burleson, W., & Picard, R. W. 2007. Gender-Specific Approaches to Developing Emotionally Intelligent Learning Companions. *IEEE Intelligent Systems*, 22(4), 62-69. doi: 10.1109/mis.2007.69
- Cacioppo, J. T., Berntson, G. C., Poehlmann, K. M., & Ito, T. A. 2000. The psychophysiology of Emotions. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of Emotions (2nd edition)*: The Guilford Press.
- Cain, B. 2007. A review of the mental workload literature: DTIC Document.
- Castor, M., Hanson, E., Svensson, E., Nählinder, S., LeBlaye, P., MacLeod, I., et al. 2003. GARTEUR Handbook of mental workload measurement. *GARTEUR, Group for Aeronautical Research and Technology in Europe, Flight Mechanics Action Group FM AG13*, 164.
- Chaffar, S., Derbali, L., & Frasson, C. 2009. Towards Emotional Regulation in Intelligent Tutoring Systems. In T. Bastiaens, J. Dron & C. Xin (Eds.), *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2009* (pp. 2428-2435). Vancouver, Canada: AACE.
- Chaffar, S., & Frasson, C. 2004. Inducing Optimal Emotional State for Learning in Intelligent Tutoring Systems. In J. C. Lester, R. M. Vicari & F. Paraguaçu (Eds.), *Intelligent Tutoring Systems* (Vol. 3220, pp. 7-21): Springer Berlin / Heidelberg.
- Chan, T. W. 1995. Artificial agents in distance learning. *International Journal of Educational Telecommunications*, 1(2-3), 263-282.
- Chanel, G., Ansari-Asl, K., & Pun, T. 2007. Valence-arousal evaluation using physiological signals in an emotion recall paradigm. *Systems, Man and Cybernetics, 2007. ISIC. IEEE International Conference on* (pp. 2662-2667).
- Chanel, G., Rebetez, C., Bétrancourt, M., & Pun, T. 2008. *Boredom, engagement and anxiety as indicators for adaptation to difficulty in games*. Paper presented at the Proceedings of the 12th international conference on Entertainment and media in the ubiquitous era.

- Chaouachi, M., Chalfoun, P., Jraidi, I., & Frasson, C. 2010. Affect and Mental Engagement: Towards Adaptability for Intelligent Systems. *23rd International FLAIRS Conference*. Daytona Beach, Florida, USA: AAAI Press.
- Chaouachi, M., Heraz, A., Jraidi, I., & Frasson, C. 2009. *Influence of Dominant Electrical Brainwaves on Learning Performance*. Paper presented at the World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2009, Vancouver, Canada.
- Chaouachi, M., Jraidi, I., & Frasson, C. 2011. Modeling Mental Workload Using EEG Features for Intelligent Systems. In J. Konstan, R. Conejo, J. L. Marzo & N. Oliver (Eds.), *User Modeling, Adaption and Personalization* (Vol. 6787, pp. 50-61): Springer Berlin Heidelberg.
- Conati, C. 2002. Probabilistic Assessment of User's Emotions in Educational Games. *Applied Artificial Intelligence*, 16, 555-575.
- Conati, C., & Merten, C. 2007. Eye-tracking for user modeling in exploratory learning environments: An empirical evaluation. *Knowledge-Based Systems*, 20(6), 557-574.
- Coulson, M. 2004. Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of nonverbal behavior*, 28(2), 117-139.
- Craig, S., Graesser, A., Sullins, J., & Gholson, B. 2004. Affect and learning: An exploratory look into the role of affect in learning with AutoTutor. *J. Edu. Med.*, 29(3), 241-250.
- Crowder, N. A. 1959. Automatic tutoring by means of intrinsic programming. *Automatic teaching: The state of the art*, 116.
- Csikszentmihalyi, M. 1990. *The psychology of optimal experience*. New York: Harper & Row.
- Csikszentmihalyi, M. 1997. *Finding flow: The psychology of engagement with everyday life*. New York: Basic Books.
- Cytowic, R. E. 1989. *Synesthesia: A Union of the Senses*. New York: Springer-Verlag.
- D'Mello, S., & Graesser, A. 2009. Automatic detection of learner's affect from gross body language. *Applied Artificial Intelligence*, 23, 123-150.
- D'Mello, S., & Graesser, A. 2011. The half-life of cognitive-affective states during complex learning. *Cognition & Emotion*, 25(7), 1299-1308.
- D'Mello, S., & Graesser, A. 2012a. Dynamics of affective states during complex learning. *Learning and Instruction*, 22(2), 145-157. doi: <http://dx.doi.org/10.1016/j.learninstruc.2011.10.001>
- D'Mello, S., Olney, A., Williams, C., & Hays, P. 2012b. Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal of human-computer studies*, 70(5), 377-398.
- D'Mello, S., Picard, R. W., & Graesser, A. 2007. Toward an Affect-Sensitive AutoTutor. *IEEE Intelligent Systems*, 22(4), 53-61. doi: <http://dx.doi.org/10.1109/MIS.2007.79>
- D'Mello, S., Taylor, R., Davidson, K., & Graesser, A. 2008. Self Versus Teacher Judgments of Learner Emotions During a Tutoring Session with AutoTutor. *Intelligent Tutoring Systems*. Montreal, Canada: Springer-Verlag.
- D'Mello, S. K., Craig, S. D., Gholson, B., Franklin, S., Picard, R. W., & Graesser, A. C. 2005. Integrating Affect Sensors in an Intelligent Tutoring System. *Affective Interactions: The Computer in the Affective Loop Workshop at International Conference on Intelligent User Interfaces* (pp. 7-13). New York: AMC Press.

- D'Mello, S. K., Lehman, B., & Person, N. 2010. Monitoring affect states during effortful problem solving activities. *International Journal of Artificial Intelligence in Education*, 20(4), 361-389.
- D'Mello, S., Chipman, P., & Graesser, A. 2007. *Posture as a predictor of learner's affective engagement*. Paper presented at the Proceedings of the 29th annual cognitive science society.
- D'Mello, S., Craig, S., Fike, K., & Graesser, A. 2009. Responding to learners' cognitive-affective states with supportive and shakeup dialogues. *Human-Computer Interaction. Ambient, Ubiquitous and Intelligent Interaction* (pp. 595-604): Springer.
- D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. 2014. Confusion can be beneficial for learning. *Learning and Instruction*, 29, 153-170.
- Damasio, A. 1994. *Descartes' Error: Emotion, Reason and the Human Brain*. New York: Grosset and Putnam.
- Davidson, R. J. 1988. EEG measures of cerebral asymmetry: Conceptual and methodological issues. *International Journal of Neuroscience*, 39(1-2), 71-89.
- Davis, R. A. 2006. Gaussian Process. *Encyclopedia of Environmetrics*: John Wiley & Sons, Ltd.
- de Vicente, A., & Pain, H. 2002. Informing the Detection of the Students' Motivational State: An Empirical Study. *6th International Conference on Intelligent Tutoring Systems* (pp. 933-943): Springer-Verlag.
- Demos, J. N. 2005. *Getting started with neurofeedback*: WW Norton & Company.
- Dimberg, U., Thunberg, M., & Elmehed, K. 2000. Unconscious facial reactions to emotional facial expressions. *Psychological science*, 11(1), 86-89.
- Donchin, E., Spencer, K. M., & Wijesinghe, R. 2000. The mental prosthesis: assessing the speed of a P300-based brain-computer interface. *Rehabilitation Engineering, IEEE Transactions on*, 8(2), 174-179.
- Dragon, T., Arroyo, I., Woolf, B. P., Burleson, W., El Kaliouby, R., & Eydgahi, H. 2008. *Viewing student affect and learning through classroom observation and physical sensors*. Paper presented at the Intelligent tutoring systems.
- Ehrlichman, H., & Wiener, M. S. 1980. EEG asymmetry during covert mental activity. *Psychophysiology*, 17(3), 228-235.
- Ekman, P. 1982. *Emotion in the Human Face, 2nd edition*. New York: Cambridge University Press.
- Ekman, P. 1999. Basic emotions. In T. Dalgleish & M. Power (Eds.), *Handbook of cognition and emotion* (Vol. 4, pp. 5-60). Sussex, UK: John Wiley & Sons.
- Ekman, P. 2004. *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*: Henry Holt & Co.
- Ekman, P., & Friesen, W. 1978. *The Facial Action Coding System: A Technique For The Measurement of Facial Movement*. San Francisco, CA: Consulting Psychologists Press, Inc.
- Ekman, P., & Friesen, W. 2003. *Unmasking the face*. Cambridge, MA.: Malor Books.
- el Kaliouby, R. 2005. *Mind-reading Machines: the automated inference of complex mental states from video*. PhD thesis, University of Cambridge.
- El Kaliouby, R., & Robinson, P. 2005. *Generalization of a vision-based computational model of mind-reading*. Paper presented at the First Int. Conf. Affective Computing and Intelligent Interaction, Beijing, China.

- Elliott, C., Rickel, J., & Lester, J. 1999. Lifelike pedagogical agents and affective computing: an exploratory synthesis. *Artificial intelligence today: recent trends and developments* (Vol. 1600, pp. 195-212): Springer-Verlag.
- Eppler, M. J., & Mengis, J. 2004. The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *The information society*, 20(5), 325-344.
- Fairclough, S. H. 2009. Fundamentals of physiological computing. *Interacting with Computers*, 21(1-2), 133-145. doi: <http://dx.doi.org/10.1016/j.intcom.2008.10.011>
- Fiedler, K. 1991. On the task, the measures and the mood in research on affect and social cognition. *Emotion and social judgments*, 83-104.
- Fiedler, K. 2001. Affective states trigger processes of assimilation and accommodation. *Theories of mood and cognition: A user's guidebook*, 85-98.
- Forbes-Riley, K., & Litman, D. 2009. *Adapting to Student Uncertainty Improves Tutoring Dialogues*. Paper presented at the AIED: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling.
- Forbes-Riley, K., Rotaru, M., & Litman, D. J. 2008. The relative impact of student affect on performance models in a spoken dialogue tutoring system. *User Modeling and User-Adapted Interaction*, 18(1-2), 11-43. doi: <http://dx.doi.org/10.1007/s11257-007-9038-5>
- Frasson, C. 1991. Systèmes tutoriels intelligents : états et perspectives en Amérique du Nord. *Génie Éducatif*, 1(1), 7-15.
- Fredrickson, B. L. 2001. The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American psychologist*, 56(3), 218.
- Fredrickson, B. L., & Branigan, C. 2005. Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition & Emotion*, 19(3), 313-332.
- Freeman, F. G., Mikulka, P. J., Prinzel, L. J., & Scerbo, M. W. 1999. Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. *Biological Psychology*, 50(1), 61-76.
- Freeman, F. G., Mikulka, P. J., Scerbo, M. W., Prinzel, L. J., & Clouatre, K. 2000. Evaluation of a psychophysiological controlled adaptive automation system, using performance on a tracking task. *Applied Psychophysiology and Biofeedback*, 25(2), 103-115.
- Frijda, N. 1986. *The Emotions*: Cambridge University Press.
- Gagné, R. M. 1985. *The conditions of learning and theory of instruction (4th ed.)*. New York, NY: Holt, Rinehart & Winston.
- Galán, F. C., & Beal, C. R. 2012. EEG estimates of engagement and cognitive workload predict math problem solving outcomes. *User Modeling, Adaptation, and Personalization* (pp. 51-62): Springer.
- Gee, J. P. 2004. *Situated Language and Learning: A Critique of Traditional Schooling*. London, UK: Routledge, Taylor & Francis.
- Gevins, A., & Smith, M. E. 2003. Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science*, 4( (1-2)), 113-131.
- Gil, S. 2009. Comment étudier les émotions en laboratoire. *Revue électronique de psychologie sociale*, 4, 15-24.
- Goleman, D. 1995. *Emotional Intelligence*. New York: Bantam Books.
- Goleman, D. 1996. *Emotional intelligence: why it can matter more than IQ*. London.

- Graesser, A. C., D'mello, S. K., Chipman, P., King, B., & McDaniel, B. 2007. Exploring Relationships Between Affect and Learning with AutoTutor. *Artificial Intelligence in Education* (pp. 16-23).
- Graesser, A. C., D'Mello, S. K., and Person, N., 2009. Meta-knowledge in tutoring. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 361-412). Mahwah, NJ: Erlbaum.
- Graesser, A. C., Wiemer-Hastings, K., Wiemer-Hastings, P., & Kreuz, R. 1999. AutoTutor: A simulation of a human tutor. *Cognitive Systems Research, 1*, 35-51.
- Grafsgaard, J., Boyer, K., Phillips, R., & Lester, J. 2011. Modeling Confusion: Facial Expression, Task, and Discourse in Task-Oriented Tutorial Dialogue. In G. Biswas, S. Bull, J. Kay & A. Mitrovic (Eds.), *Artificial Intelligence in Education* (Vol. 6738, pp. 98-105): Springer Berlin Heidelberg.
- Graham, S., & Weiner, B. 1996. Theories and principles of motivation. In D.C. Berliner & R. C. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 63-84). New York: Macmillan.
- Gunes, H., & Pantic, M. 2010. Automatic, dimensional and continuous emotion recognition. *International Journal of Synthetic Emotions (IJSE), 1*(1), 68-99.
- Hancock, P., & Chignell, M. H. 1986. Toward a Theory of Mental Work Load: Stress and Adaptability in Human-Machine Systems. *Proc. IEEE SMC 1986*, 378-383.
- Hankins, T. C., & Wilson, G. F. 1998. A comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight. *Aviation, Space, and Environmental Medicine, 69*(4), 360-367.
- Hart, S. G., & Staveland, L. E. 1988a. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload, 1*(3), 139-183.
- Hart, S. G., & Staveland, L. E. 1988b. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam: North-Holland.
- Healey, J. A. 2000. *Wearable and automotive systems for affect recognition from physiology*. Massachusetts Institute of Technology.
- Heger, D., Putze, F., & Schultz, T. 2010. Online workload recognition from EEG data during cognitive tests and human-machine interaction. *KI 2010: Advances in Artificial Intelligence* (pp. 410-417): Springer.
- Heraz, A., Jraidi, I., Chaouachi, M., & Frasson, C. 2009. Predicting Stress Level Variation from Learner Characteristics and Brainwaves. *AIED'2009: 14th International Conference on Artificial Intelligence in Education* (pp. 722-724). Brighton, UK: IOS Press.
- Hoc, J., & Amalberti, R. 2003. Adaptation et contrôle cognitif: supervision de situations dynamiques complexes. *Actes des Deuxièmes Journées d'Étude en Psychologie Ergonomique-ÉPIQUE*, 135-147.
- Hung, Y.-C. 2012. Combining Self-Explaining With Computer Architecture Diagrams to Enhance the Learning of Assembly Language Programming. *Education, IEEE Transactions on, 55*(4), 546-551.
- Hussain, M. S., AlZoubi, O., Calvo, R. A., & D'Mello, S. K. 2011. *Affect detection from multichannel physiology during learning sessions with AutoTutor*. Paper presented at the Artificial Intelligence in Education.

- Isen, A. 2000. Positive affect and decision making. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 417-435). New York: Guilford Press.
- Isen, A. M., Daubman, K. A., & Nowicki, G. P. 1987. Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology*, 52(6), 1122–1131.
- Izard, C. E. 1993. *The Differential Emotions Scale: DES IV-A; [a Method of Measuring the Meaning of Subjective Experience of Discrete Emotions]*: University of Delaware.
- Jasper, H. H. 1958. The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*(10), 371-375.
- Jraidi, I. 2014. Modélisation des émotions de l'apprenant et interventions implicites pour les Systèmes Tutoriels Intelligents.
- Jraidi, I., Chalfoun, P., & Frasson, C. 2012. Implicit Strategies for Intelligent Tutoring Systems. In S. Cerri, W. Clancey, G. Papadourakis & K. Panourgia (Eds.), *Intelligent Tutoring Systems* (Vol. 7315, pp. 1-10): Springer Berlin Heidelberg.
- Jraidi, I., Chaouachi, M., & Frasson, C. 2010a. Enhancing Learner Self-Esteem for Learning Improvements. *23rd International FLAIRS Conference*. Daytona Beach, Florida, USA: AAAI Press.
- Jraidi, I., Chaouachi, M., & Frasson, C. 2010b. Self-Esteem Conditioning for Learning Conditioning. *Proceedings of the 2010 10th IEEE International Conference on Advanced Learning Technologies*: IEEE Computer Society.
- Jraidi, I., Chaouachi, M., & Frasson, C. 2011. Automatic Detection of User' Uncertainty in Problem Solving Task: a Multimodal Approach. *FLAIRS Conference*.
- Jraidi, I., Chaouachi, M., & Frasson, C. 2013a. A hierarchical probabilistic framework for recognizing learners' interaction experience trends and emotions. *Submitted to Advances in Human-Computer Interaction*.
- Jraidi, I., & Frasson, C. 2010c. Subliminally Enhancing Self-esteem: Impact on Learner Performance and Affective State. In V. Aleven, J. Kay & J. Mostow (Eds.), *Intelligent Tutoring Systems* (Vol. 6095, pp. 11-20): Springer Berlin Heidelberg.
- Jraidi, I., & Frasson, C. 2013b. Student's Uncertainty Modeling through a Multimodal Sensor-Based Approach. *Journal of Educational Technology & Society*, 16(1), 219-230.
- Jraidi, I., Heraz, A., Chaouachi, M., & Frasson, C. 2009. New Architecture of a Multi Agent System which Measures the Learner Brainwaves to Predict his Stress Level Variation. In T. Bastiaens, J. Dron & C. Xin (Eds.), *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2009* (pp. 2726-2733). Vancouver, Canada: AACE.
- Jung, T.-P., & Makeig, S. 1994. *Estimating level of alertness from EEG*. Paper presented at the Engineering in Medicine and Biology Society, 1994. Engineering Advances: New Opportunities for Biomedical Engineers. Proceedings of the 16th Annual International Conference of the IEEE.
- Just, M. A., Carpenter, P. A., & Miyake, A. 2003. Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work. *Theoretical Issues in Ergonomics Science*, 4(1-2), 56-88.
- Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. 2001. When problem solving is superior to studying worked examples. *Journal of educational psychology*, 93(3), 579.
- Kapoor, A., Burleson, W., & Picard, R. W. 2007. Automatic prediction of frustration. *International Journal of Human-Computer Studies*, 65(8), 724-736. doi: 10.1016/j.ijhcs.2007.02.003

- Kapoor, A., Mota, S., & Picard, R. 2001. Toward a learning companion that recognizes affect. *AAAI Fall Symposium: Emotional and Intelligent*. North Falmouth, Massachusetts.
- Kaufmann, H., & Dünser, A. 2007. *Summary of usability evaluations of an educational augmented reality application*: Springer.
- Keller, J. M. 1987. Development and use of the ARCS model of instructional design. *Journal of instructional development*, 10(3), 2-10.
- Keltner, D., & Haidt, J. 2001. Social functions of emotions. In T. J. Mayne & G. A. Bonanno (Eds.), *Emotions: Current Issues and Future Directions* (pp. 192-213): Guilford Press.
- Kim, J. 2007. Bimodal emotion recognition using speech and physiological changes. *Robust speech recognition and understanding*, 265-280.
- Kim, J., & André, E. 2006. Emotion recognition using physiological and speech signal in short-term observation. *Perception and interactive technologies* (pp. 53-64): Springer.
- Kim, Y. 2005. Empathetic virtual peers enhanced learner interest and self-efficacy. *Workshop on Motivation and Affect in Educational Software, in conjunction with the 12th International Conference on Artificial Intelligence in Education*.
- Kirsh, D. 2000. A few thoughts on cognitive overload.
- Koedinger, K. R., & Sueker, E. L. 1996. PAT goes to college: Evaluating a cognitive tutor for developmental mathematics. *The 1996 international conference on Learning sciences* (pp. 180-187): International Society of the Learning Sciences.
- Kohlmorgen, J., Dornhege, G., Braun, M., Blankertz, B., Müller, K., Curio, G., et al. 2007. Improving human performance in a real operating environment through real-time mental workload detection. *Toward Brain-Computer Interfacing* (pp. 409-422). Cambridge, MA: MIT Press.
- Lane, H. C. 2006. Intelligent tutoring systems: Prospects for guided practice and efficient learning: Citeseer.
- Lang, P. J. 1995. The emotion probe: Studies of motivation and attention. *American Psychologist*, 50(5), 372-385. doi: 10.1037/0003-066x.50.5.372
- LeDoux, J. 1999. *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. London: Phoenix.
- Lee, H., Plass, J. L., & Homer, B. D. 2006. Optimizing cognitive load for learning from computer-based science simulations. *Journal of educational psychology*, 98(4), 902.
- Levenson, R. W. 1988. Emotion and the autonomic nervous system: A prospectus for research on autonomic specificity. *Social psychophysiology and emotion: Theory and clinical applications*, 17-42.
- Lewis, J. R. 1991. Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. *ACM SIGCHI Bulletin*, 23(1), 78-81.
- Liang, S., Lin, C., Wu, R., Chen, Y., Huang, T., & Jung, T. 2006. *Monitoring driver's alertness based on the driving performance estimation and the EEG power spectrum analysis*. Paper presented at the Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the.
- Liao, W., Zhang, W., Zhu, Z., Ji, Q., & Gray, W. D. 2006. Toward a decision-theoretic framework for affect recognition and user assistance. *International Journal of Human-Computer Studies - Human-computer interaction research in the management information systems discipline*, 64(9), 847-873. doi: 10.1016/j.ijhcs.2006.04.001

- Litman, D. J., & Forbes-Riley, K. 2004. Predicting student emotions in computer-human tutoring dialogues. *42nd Annual Meeting on Association for Computational Linguistics*. Barcelona, Spain: Association for Computational Linguistics.
- Litman, D. J., & Forbes-Riley, K. 2006. Recognizing student emotions and attitudes on the basis of utterances in spoken tutoring dialogues with both human and computer tutors. *Speech Communication*, 48(5), 559-590.
- Lu, S., Tsechpenakis, G., Metaxas, D. N., Jensen, M. L., & Kruse, J. 2005. *Blob Analysis of the Head and Hands: A Method for Deception Detection*. Paper presented at the HICSS.
- Lubar, J. F. 1991. Discourse on the development of EEG diagnostics and biofeedback for attention-deficit/hyperactivity disorders. *Biofeedback and Self-regulation*, 16(3), 201-225.
- Lubar, J. F., Swartwood, M. O., Swartwood, J. N., & O'Donnell, P. H. 1995. Evaluation of the effectiveness of EEG neurofeedback training for ADHD in a clinical setting as measured by changes in TOVA scores, behavioral ratings, and WISC-R performance. *Biofeedback and Self-regulation*, 20(1), 83-99.
- Mandler, G. 1984. *Mind and body: Psychology of emotion and stress*: WW Norton New York.
- Mandryk, R. L., Atkins, M. S., & Inkpen, K. M. 2006. *A continuous and objective evaluation of emotional experience with interactive play environments*. Paper presented at the Proceedings of the SIGCHI conference on Human Factors in computing systems.
- Markus, H. R., & Kitayama, S. 1991. Culture and the self: Implications for cognition, emotion, and motivation. *Psychological review*, 98(2), 224.
- Mayer, J. D., & Salovey, P. 1997. What is emotional intelligence? In P. Salovey & D. Sluyter (Eds.), *Emotional development and emotional intelligence: Implications for educators* (pp. 3-31). New York: Basic Books.
- Mayer, R. E., & Moreno, R. 2003. Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist*, 38(1), 43-52.
- McDaniel, B. T., D'Mello, S. K., King, B. G., Chipman, P., & Tapp, K. a. G., A.C. 2007. Facial features for affective state detection in learning environments. In D. S. McNamara & J. G. Trafton (Eds.), *29th Annual Cognitive Science Society* (pp. 467-472). Austin, TX: Cognitive Science Society.
- McQuiggan, S. W., Lee, S., & Lester, J. C. 2007a. Early Prediction of Student Frustration. *2nd international conference on Affective Computing and Intelligent Interaction*. Lisbon, Portugal: Springer-Verlag.
- Mcquiggan, S. W., Lee, S., & Lester, J. C. 2007b. Early prediction of student frustration. *Affective Computing and Intelligent Interaction* (pp. 698-709): Springer.
- McQuiggan, S. W., & Lester, J. C. 2006. Learning empathy: a data-driven framework for modeling empathetic companion agents. *5th international joint conference on Autonomous agents and multiagent systems* (pp. 961-968). Hakodate, Japan: ACM.
- McQuiggan, S. W., Robison, J. L., & Lester, J. C. 2008a. *Affective transitions in narrative-centered learning environments*. Paper presented at the Intelligent Tutoring Systems.
- McQuiggan, S. W., Robison, J. L., Phillips, R., & Lester, J. C. 2008b. *Modeling parallel and reactive empathy in virtual agents: An inductive approach*. Paper presented at the Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 1.

- Mehrabian, A. 1996. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in Temperament. *Current Psychology*, 14(4), 261-292. doi: 10.1007/bf02686918
- Mehrabian, A., & Russell, J. A. 1974. *An approach to environmental psychology*: the MIT Press.
- Meshkati, N., & Hancock, P. 2011. *Human mental workload*: Elsevier.
- Mitra, S., & Acharya, T. 2007. Gesture recognition: A survey. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 37(3), 311-324.
- Monk, C. A., Boehm-Davis, D. A., & Trafton, J. G. 2002. The attentional costs of interrupting task performance at various stages. *Human Factors and Ergonomics Society Annual Meeting* (Vol. 46, pp. 1824-1828): SAGE Publications.
- Mori, J., Prendinger, H., & Ishizuka, M. 2003. Evaluation of an Embodied Conversational Agent with Affective Behavior. In J. Rosenschein & M. Wooldridge (Eds.), *Proceedings of International joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), Workshop on Embodied Conversational Characters as Individuals* (pp. 58-61). Melbourne, Australia.
- Mota, S., & Picard, R. W. 2003. Automated Posture Analysis for Detecting Learner's Interest Level. *Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction*.
- Mulder, G., Mulder, L. J., Meijman, T. F., Veldman, J. B., & van Roon, A. M. 2000. A psychophysiological approach to working conditions. *Engineering psychophysiology: Issues and applications*, 139-159.
- Muse, L. A., Harris, S. G., & Feild, H. S. 2003. Has the inverted-U theory of stress and job performance had a fair test? *Human Performance*, 16(4), 349-364. doi: 10.1207/S15327043HUP1604\_2
- Nguyen-Tuong, D., Peters, J. R., & Seeger, M. 2008. *Local Gaussian process regression for real time online model learning*. Paper presented at the Advances in Neural Information Processing Systems.
- Nkambou, R. 2006. *A framework for affective intelligent tutoring systems*. Paper presented at the Information Technology Based Higher Education and Training, 2006. ITHET'06. 7th International Conference on.
- Nkambou, R., & Héritier, V. 2004. *Facial expression analysis for emotion recognition in ITS*. Paper presented at the ITS 2004 workshop on Emotional Intelligence proceedings.
- Nwana, H. S. 1990. Intelligent tutoring systems: an overview. *Artificial Intelligence Review*, 4(4), 251-277. doi: 10.1007/bf00168958
- Ochs, M., & Frasson, C. 2004. Emotionally Intelligent Tutoring Systems (EITS). *The Florida Artificial Intelligence Research Society (FLAIRS) Conference* (pp. 251-255). Miami, Floride, Etats-Unis.
- Offenloch, K., & Zahner, G. 1990. Computer aided physiological assessment of the functional state of pilots during simulated flight: DTIC Document.
- Ortony, A., Clore, G., & Collins, A. 1988. *The Cognitive Structure of Emotions*. Cambridge, MA: Cambridge University Press.
- Ortony, A. T., Terence J. 1990. What's basic about basic emotions? *Psychological Review*, 97(3), 315-331. doi: 10.1037/0033-295X.97.3.315

- Paas, F., Renkl, A., & Sweller, J. 2004. Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional science*, 32(1), 1-8.
- Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. 2003. Cognitive load measurement as a means to advance cognitive load theory. *Educational psychologist*, 38(1), 63-71.
- Paas, F. G. 1992. Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of educational psychology*, 84(4), 429.
- Panksepp, J. 1992. A critical role for "affective neuroscience" in resolving what is basic about basic emotions. *Psychological Review*, 99(3), 554-560.
- Pantic, M., & Rothkrantz, L. J. 2003. Toward an affect-sensitive multimodal human-computer interaction. *Proceedings of the IEEE*, 91(9), 1370-1390.
- Parasuraman, R., & Caggiano, D. 2002. Mental workload. *Encyclopedia of the human brain*, 3, 17-27.
- Pekrun, R. 1992. The Impact of Emotions on Learning and Achievement: Towards a Theory of Cognitive/Motivational Mediators. *Applied Psychology*, 41(4), 359-376. doi: 10.1111/j.1464-0597.1992.tb00712.x
- Pekrun, R. 2006. The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315-341.
- Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. 2011. Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, 36(1), 36-48.
- Petty, R. E., Cacioppo, J. T., Sedikides, C., & Strathman, A. J. 1988. Affect and persuasion: A contemporary perspective. *American Behavioral Scientist*.
- Phelps, E. A. 2006. Emotion and cognition: insights from studies of the human amygdala. *Annu. Rev. Psychol.*, 57, 27-53.
- Picard, R. 1997. *Affective Computing*: MIT Press.
- Picard, R. W., & Daily, S. B. 2005. *Evaluating affective interactions: Alternatives to asking what users feel*. Paper presented at the CHI Workshop on Evaluating Affective Interfaces: Innovative Approaches.
- Picard, R. W., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., et al. 2004. Affective Learning - A Manifesto. *BT Tech. J.*, 22(4), 253-269.
- Plutchik, R. 1980. A general psychoevolutionary theory of emotion. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience: Vol. 1. Theories of emotion* (pp. 3-33): Academic press.
- Pope, A. T., Bogart, E. H., & Bartolome, D. S. 1995. Biocybernetic system evaluates indices of operator engagement in automated task. *Biological psychology*, 40(1), 187-195.
- Prendinger, H., & Ishizuka, M. 2005a. The empathic companion: A character-based interface that addresses users' affective states. *Applied Artificial Intelligence*, 19(3-4), 267-285.
- Prendinger, H., Mori, J., & Ishizuka, M. 2005b. Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game. *International journal of human-computer studies*, 62(2), 231-245.
- Prinzel Iii, L. J., Pope, A. T., & Freeman, F. G. 2002. Physiological self-regulation and adaptive automation. *The International Journal of Aviation Psychology*, 12(2), 179-196.

- Prinzel, L. J., Freeman, F. G., & Scerbo, M. W. 2000a. A Closed-Loop System for Examining Psychophysiological Measures for Adaptive Task Allocation. *International Journal of Aviation Psychology*, 10(4), 393-410.
- Prinzel, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. 2000b. A closed-loop system for examining psychophysiological measures for adaptive task allocation. *Int J Aviat Psychol*, 10(4), 393-410. doi: 10.1207/s15327108ijap1004\_6
- Prinzel, L. J., Scerbo, M. W., Freeman, F. G., & Mikulka, P. J. 1995. *A bio-cybernetic system for adaptive automation*. Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.
- Pсотка, J., Massey, L. D., & Mutter, S. A. 1988a. *Intelligent tutoring systems : lessons learned*: L. Erlbaum Associates.
- Pсотка, J., Massey, L. D., & Mutter, S. A. 1988b. *Intelligent tutoring systems: Lessons learned*: Psychology Press.
- Qi, Y., Reynolds, C., & Picard, R. W. 2001. The Bayes Point Machine for computer-user frustration detection via pressuremouse. *2001 workshop on Perceptive user interfaces* (pp. 1-5). Orlando, Florida: ACM.
- Rasmussen, C. E. 2006. Gaussian processes for machine learning.
- Reed, G. F. 1977. Obsessional cognition: performance on two numerical tasks. *The British Journal of Psychiatry*, 130(2), 184-185.
- Reid, G. B., & Nygren, T. 1988. The subjective workload assessment technique: A scaling procedure for measuring mental workload. *Human mental workload*, 185, 218.
- Reilly, E. L. 1999. EEG recording and operation of the apparatus. *Electroencephalography: basic principles, clinical applications and related fields*, 122-142.
- Reynolds, C. 2001. *The Sensing and Measurement of Frustration with Computers*. SM Thesis, Massachusetts Institute of Technology
- Robison, J., McQuiggan, S., & Lester, J. 2009a. Evaluating the consequences of affective feedback in intelligent tutoring systems. *ACII 2009. 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops* (pp. 1-6).
- Robison, J. L., Mcquiggan, S. W., & Lester, J. C. 2009b. Modeling Task-Based vs. Affect-based Feedback Behavior in Pedagogical Agents: An Inductive Approach. *Artificial Intelligence in Education* (pp. 25-32): IOS Press.
- Rodrigo, M., Baker, R., D'Mello, S., Gonzalez, M., Lagud, M., Lim, S., et al. 2008. Comparing Learners' Affect While Using an Intelligent Tutoring System and a Simulation Problem Solving Game. *Intelligent Tutoring Systems* (pp. 40-49).
- Rowe, J., Mott, B., McQuiggan, S., Robison, J., Lee, S., & Lester, J. 2009. *Crystal island: A narrative-centered learning environment for eighth grade microbiology*. Paper presented at the Workshop on Intelligent Educational Games at the 14th International Conference on Artificial Intelligence in Education, Brighton, UK.
- Rubio, S., Díaz, E., Martín, J., & Puente, J. M. 2004. Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology*, 53(1), 61-86.
- Russell, C. A., Wilson, G. F., Rizki, M. M., Webb, T. S., & Gustafson, S. C. 2005. Comparing classifiers for real time estimation of cognitive workload. *Foundations of Augmented Cognition*. Las Vegas, NV: Lawrence Erlbaum Associates Inc.
- Russell, J. A. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology - PSP*, 39, 1161-1178.

- Russell, J. A. 2003. Core affect and the psychological construction of emotion. *Psychological Review*, 110(1), 145-172. doi: 10.1037/0033-295X.110.1.145
- Russell, J. A., Bachorowski, J.-A., & Fernandez-Dols, J. M. 2003. Facial and vocal expressions of emotion. *Annual review of psychology*, 54(1), 329-349.
- Russell, J. A., Weiss, A., & Mendelsohn, G. A. 1989. Affect grid: a single-item scale of pleasure and arousal. *Journal of personality and social psychology*, 57(3), 493.
- Salovey, P., & Mayer, J. D. 1990. Emotional intelligence. *Imagination, Cognition and Personality*, 9(3), 185-211.
- Samsonovich, A. V., Jong, K. A. D., Kitsantas, A., Peters, E. E., Dabbagh, N., & Kalbfleisch, M. L. 2008. *Cognitive Constructor: An Intelligent Tutoring System based on a Biologically Inspired Cognitive Architecture (BICA)*. Paper presented at the Artificial General Intelligence 2008: Proceedings of the First AGI Conference.
- Sanghvi, J., Castellano, G., Leite, I., Pereira, A., McOwan, P. W., & Paiva, A. 2011. *Automatic analysis of affective postures and body motion to detect engagement with a game companion*. Paper presented at the Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference on.
- Sarrafzadeh, A., Fan, C., Dadgostar, F., Alexander, S., & Messom, C. 2004. *Frown gives game away: affect sensitive systems for elementary mathematics*. Paper presented at the Systems, Man and Cybernetics, 2004 IEEE International Conference on.
- Scherer, K. 2000. Emotion. In M. Hewstone & W. Stroebe (Eds.), *Introduction to Social Psychology : A European perspective* (pp. 151–191): Oxford Blackwell Publishers.
- Scherer, K. 2005. What are emotions? And how can they be measured? *Social Science Information*, 44(4), 695-729.
- Scott, D. 1971. On engendering an illusion of understanding. *The Journal of Philosophy*, 787-807.
- Shen, L., Wang, M., & Shen, R. 2009. Affective e-Learning: Using " Emotional " Data to Improve Learning in Pervasive Learning Environment. *Educational Technology & Society*, 12(2), 176-189.
- Shouse, E. 2005. Feeling, Emotion, Affect. *M/C Journal*, 8(6).
- Shute, V. J. 1995. SMART: Student modeling approach for responsive tutoring. *User Modeling and User-Adapted Interaction*, 5(1), 1-44.
- Sirevaag, E. J., & Stern, J. A. 2000. Ocular measures of fatigue and cognitive factors. *Engineering psychophysiology: Issues and applications*, 269-287.
- Sobol-Shikler, T. 2011. Automatic inference of complex affective states. *Computer Speech & Language*, 25(1), 45-62.
- Speier, C., Valacich, J. S., & Vessey, I. 1999. The influence of task interruption on individual decision making: An information overload perspective. *Decision Sciences*, 30(2), 337-360.
- Stamenović, J., Đurić, S., Jolić, M., Živadinović, B., & Đurić, V. 2005. Examination of cognitive functions in patients with Parkinsons disease. *Facta Universitatis*, 11, 104-113.
- Stern, M. K., & Woolf, B. P. 1998. *Curriculum sequencing in a web-based tutor*. Paper presented at the Intelligent Tutoring Systems.
- Stevens, R., Galloway, T., & Berka, C. 2006. Integrating EEG Models of Cognitive Load with Machine Learning Models of Scientific Problem Solving. In D. Schmorow, K.

- Stanney & L. Reeves (Eds.), *Augmented Cognition: Past, Present and Future* (pp. 55-65): Strategic Analysis, Inc., Arlington.
- Stevens, R., Galloway, T., & Berka, C. 2007a. EEG-Related Changes in Cognitive Workload, Engagement and Distraction as Students Acquire Problem Solving Skills. In C. Conati, K. McCoy & G. Paliouras (Eds.), *User Modeling 2007* (Vol. 4511, pp. 187-196): Springer Berlin / Heidelberg.
- Stevens, R. H., Galloway, T., & Berka, C. 2007b. *Integrating innovative neuro-educational technologies (I-Net) into K-12 science classrooms*. Paper presented at the Proceedings of the 3rd international conference on Foundations of augmented cognition, Beijing, China.
- Stipacek, A., Grabner, R., Neuper, C., Fink, A., & Neubauer, A. 2003. Sensitivity of human EEG alpha band desynchronization to different working memory components and increasing levels of memory load. *Neuroscience Letters*, 353(3), 193-196.
- Strain, A. C., & D'Mello, S. K. 2011. *Emotion regulation during learning*. Paper presented at the Artificial Intelligence in Education.
- Sweller, J. 1994. Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, 4(4), 295-312.
- Sweller, J. 2003. Evolution of human cognitive architecture. *Psychology of Learning and Motivation*, 43, 215-266.
- Sweller, J., Ayres, P., & Kalyuga, S. 2011. *Cognitive load theory* (Vol. 1): Springer.
- Symonds, W. 2001. Giving it the old online try. *Business Week*(Issue 3760, 3 December).
- Tao, J., & Tan, T. 2005. Affective computing: A review. *Affective computing and intelligent interaction* (pp. 981-995): Springer.
- Tchounikine, P. 2002. Quelques éléments sur la conception et l'ingénierie des EIAH. *Information Interaction Intelligence (Actes des 2ème assises nationales du GdR I3, Nancy, décembre 2002)*.
- Tomkins, S. S. 1984. Affect theory. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 163-195). Hillsdale, NJ: Erlbaum.
- Tricot, A., & Chanquoy, L. 1996. La charge mentale, «vertu dormitive» ou concept opérationnel?: La charge mentale. *Psychologie française*, 41(4), 313-318.
- Tsang, P. S., & Velazquez, V. L. 1996. Diagnosticity and multidimensional subjective workload ratings. *Ergonomics*, 39(3), 358-381.
- Valdés, P., Bosch, J., Grave, R., Hernandez, J., Riera, J., Pascual, R., et al. 1992. Frequency domain models of the EEG. *Brain topography*, 4(4), 309-319.
- VanLehn, K. 2011. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197-221.
- VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. B. 2003. Why Do Only Some Events Cause Learning During Human Tutoring? *Cognition and Instruction*, 21(3), 209-249.
- Vidulich, M. A., & Tsang, P. S. 2012. Mental workload and situation awareness. *Handbook of Human Factors and Ergonomics*, 243-273.
- Wagner, J., Kim, J., & André, E. 2005. *From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification*. Paper presented at the Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on.

- Wang, H., Chignell, M., & Ishizuka, M. 2006. *Empathic tutoring software agents using real-time eye tracking*. Paper presented at the Proceedings of the 2006 symposium on Eye tracking research & applications.
- Ward, R. D., & Marsden, P. H. 2003. Physiological responses to different WEB page designs. *International Journal of Human-Computer Studies - Application of affective computing in human-computer interaction*, 59(1-2), 199-212. doi: 10.1016/s1071-5819(03)00019-3
- Webber, C., Pesty, S., & Balacheff, N. 2002. *A multi-agent and emergent approach to learner modelling*. Paper presented at the ECAI.
- Welford, A. 1978. Mental work-load as a function of demand, capacity, strategy and skill. *Ergonomics*, 21(3), 151-167.
- Wenger, E. 1987. *Artificial intelligence and tutoring systems: computational and cognitive approaches to the communication of knowledge*. Los Altos, CA: Morgan Kaufmann.
- Whitehill, J., Bartlett, M., & Movellan, J. 2008. Automatic facial expression recognition for intelligent tutoring systems. *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW '08. IEEE Computer Society Conference on* (pp. 1-6).
- Wiederhold, B. K., Jang, D., Kaneda, M., Cabral, I., Lurie, Y., May, T., et al. 2001. An investigation into physiological responses in virtual environments: an objective measurement of presence. *Towards cyberpsychology: Mind, cognitions and society in the internet age*.
- Wilson, G. F. 2002. An analysis of mental workload in pilots during flight using multiple psychophysiological measures. *Int J Aviat Psychol*, 12, 3-18.
- Woolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D., & Picard, R. 2009. Affect aware tutors : recognising and responding to student affect. *International Journal of Learning Technology*, 4(3/4), 129-164.
- Yoo, S. K., Lee, C. K., Park, Y. J., Kim, N. H., Lee, B. C., & Jeong, K. S. 2005. Neural network based emotion estimation using heart rate variability and skin resistance. *Advances in Natural Computation* (pp. 818-824): Springer.
- Young, M. S., & Stanton, N. A. 2002. Malleable attentional resources theory: a new explanation for the effects of mental underload on performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 44(3), 365-375.
- Zakharov, K., Mitrovic, A., & Johnston, L. 2008. Towards Emotionally-Intelligent Pedagogical Agents. In B. Woolf, E. Aïmeur, R. Nkambou & S. Lajoie (Eds.), *Intelligent Tutoring Systems* (Vol. 5091, pp. 19-28): Springer Berlin / Heidelberg.
- Zhai, J., & Barreto, A. 2006. Stress Detection in Computer Users Based on Digital Signal Processing of Noninvasive Physiological Variables. *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE* (pp. 1355-1358).
- Zimmerman, B. J. 2000. Self-Efficacy: An Essential Motive to Learn. *Contemporary Educational Psychology*, 25(1), 82-91. doi: <http://dx.doi.org/10.1006/ceps.1999.1016>
- Zimmermann, P., Guttormsen, S., Danuser, B., & Gomez, P. 2003. Affective computing--a rationale for measuring mood with mouse and keyboard. *International journal of occupational safety and ergonomics*, 9(4), 539-551.



## Annexe A

# Régression par processus gaussiens

Un processus gaussien est un processus stochastique dont la distribution jointe de tout sous ensemble de ses variables  $f = \{f_i\}_{i=1}^n$  défini sur un ensemble d'index  $X = \{x_i\}_{i=1}^n$ , est par hypothèse *une distribution normale multivariée* à  $n$  dimensions. Un processus gaussien est défini par une fonction de covariance  $k$  et d'une fonction de moyennes  $\mu$  qui permettent d'obtenir la distribution :

$$p(f) \sim \mathcal{N}(\mu, \mathbf{K})$$

$\mathcal{N}$  étant la distribution normale,  $\mathbf{K}$  est la matrice  $n$  par  $n$  de covariance construite selon  $\mathbf{K}=k(x_i, x_j)$  et  $\mu$  est le vecteur de moyenne construit tel que  $\mu_i = \mu(x_i)$ .

La régression par les processus gaussiens permet d'estimer une fonction  $f$  expliquant au mieux un ensemble de données d'observation  $\{x_i, y_i\}_{i=1}^n$ . Avec  $x_i$  un vecteur d'entrée et  $y_i$  est la valeur cible. Le principe de base de cette approche bayésienne repose sur l'idée de définir une probabilité a priori dans l'espace de fonctions et d'appliquer un processus d'inférence qui permettra d'augmenter la probabilité des fonctions qui expliquent le mieux les données

d'observation et inversement réduire la probabilité des fonctions qui ne disposent pas de cette capacité d'explication.

Il existe plusieurs types de fonctions de covariances employées dans la régression par les processus gaussiens, une des plus fréquemment utilisées est le noyau gaussien (Rasmussen, 2006) :

$$k(x_i, x_j) = \sigma_s^2 \exp\left(-\frac{1}{2}(x_i - x_j)^T \mathbf{W}(x_i - x_j)\right)$$

$\sigma_s^2$  représente la variance du signal et  $\mathbf{W}$  est la largeur du noyau gaussien.

Pour la fonction de moyennes on a généralement recours une fonction à moyennes égales à zéro (Rasmussen, 2006). Ainsi en pratique on a souvent la formulation suivante de la probabilité *a priori* du modèle :

$$p(\mathbf{f}) \sim \mathcal{N}(\mathbf{0}, \mathbf{K})$$

De même, le vecteur  $\mathbf{f}$  ne pas être observé directement et les données  $\mathbf{y}$  sont généralement considérées comme bruitées. Dans les processus gaussiens, la définition d'un modèle de bruit conduit à une distribution de probabilité conditionnelle sur les observations  $p(\mathbf{y} | \mathbf{f})$  qui sera utilisé pour la mise à jour de la distribution de probabilité sur les fonctions.

Ainsi pour transformer un vecteur d'entrée  $x_i$  en une valeur cible  $y_i$ , la formulation du modèle de régression standard par processus gaussien est la suivante :

$$y_i = f(x_i) + \varepsilon_i$$

Avec  $\varepsilon_i$  un bruit blanc gaussien de variance  $\sigma_n$  et de moyenne égale à zéro. Cette hypothèse auquel est soumise la fonction de régression  $f(x)$  donne lieu à une probabilité conditionnelle sous forme gaussienne décrite par la formulation suivante :

$$p(\mathbf{y} | \mathbf{f}) \sim \mathcal{N}(\mathbf{f}, \sigma_n^2 \mathbf{I})$$

$\mathbf{I}$  est une matrice identité utilisée pour indiquer que le bruit utilisé est stationnaire.

L'évaluation de cette expression pour une fonction arbitraire nous renseigne sur la probabilité que les valeurs cibles sont générées par cette fonction. Il s'agit alors de la *vraisemblance* de la fonction que si l'on pondère par la probabilité a priori pour l'ensemble des fonctions, on obtient la vraisemblance marginale des valeurs cibles :

$$p(y) \sim \int p(y|f)p(f)df$$

$$p(y) \sim \mathcal{N}(0, \mathbf{K} + \sigma_n^2 \mathbf{I})$$

La prédiction d'une nouvelle entrée  $\mathbf{x}_*$  consiste à déterminer  $f(x_*)$ . La distribution jointe de cette prédiction  $f(x_*)$  avec les anciennes valeurs cibles observées est donnée par :

$$\begin{bmatrix} y \\ f(x_*) \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \mathbf{K} + \sigma_n^2 \mathbf{I} & k(X, x_*) \\ k(x_*, X) & k(x_*, x_*) \end{bmatrix}\right)$$

La distribution conditionnelle donne alors à la valeur à prédire  $f(x_*)$  avec la variance  $V(x_*)$  selon les deux équations suivantes :

$$f(x_*) = k_*^T (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} y = k_*^T \alpha$$

$$V(x_*) = k(x_*, x_*) - k_*^T (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} k_*$$

Avec  $k_* = k(X, x_*)$  et  $\alpha = (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} y$

## Annexe B

# L'instrument NASA TLX

Table de définition des dimensions du NASA\_TLX

Dimension	Pôle	Description
Exigence mentale	Faible / Élevé	Jusqu'à quel point un effort mental était requis par la tâche (réflexion, mémorisation, décision, calcul, etc.) ?
Exigence physique	Faible / Élevé	Jusqu'à quel point un effort physique était requis par la tâche (activer, contrôler, manipuler, etc.) ?
Exigence temporelle	Faible / Élevé	Jusqu'à quel point vous avez senti la pression du temps lors de l'exécution de la tâche ?
Performance	Bonne / Mauvaise	Jusqu'à quel point vous êtes satisfait avec votre performance dans l'exécution de la tâche ?
Effort	Faible / Élevé	Jusqu'à quel point avez-vous eu à travailler (mentalement ou physiquement) pour atteindre votre niveau de performance?
Frustration	Faible / Élevé	Jusqu'à quel point vous sentiez-vous non confiant, découragé, irrité, stressé et ennuyés vs confiant, avec plaisir, content, relaxe, satisfait de vous durant la tâche?

# Publications

## 2014

- Jraidi, I., Chaouachi, M., & Frasson, C. (2014). “A Hierarchical Probabilistic Framework for Recognizing Learners’ Interaction Experience Trends and Emotions”. *Advances in Human-Computer Interaction*, 2014.
- Ghali, R., Chaouachi, M., Derbali, L., Frasson, C. “Motivational strategies to support learners in Serious Games”, 5th International Conference on Agents and Artificial Intelligence ICAART, ScitePress, Anger, France

## 2013

- Jraidi, I., Chaouachi, M., Frasson, C. “A dynamic multimodal approach for assessing learners' interaction experience”, ICMI 2013: 15th ACM on International conference on multimodal interaction, Sydney. Australia, ACM. **Best paper award winner.**

## 2012

- Chaouachi, M., Frasson, C. “Mental workload, engagement and emotions: an exploratory study for intelligent tutoring systems”, ITS 2012: 11th International Conference on Intelligent Tutoring Systems, chania, Greece
- Chaouachi, M., Frasson, C. “Workload Aware Learning Environment and Self-Regulated Learning”, ITS 2012: Workshop on SRELT, chania, Greece

## 2011

- Chaouachi, M., Jraidi, I., Frasson, C. “Modeling Mental Workload using EEG Features for Intelligent Systems”, submitted in User Modeling, Adaptation and Personalization, Girona, Spain
- Jraidi, I., Chaouachi, M., Frasson, C. “Automatic Detection of User’s Uncertainty in Problem Solving Task: a Multimodal Approach”, 24th International FLAIRS Conference, Palm Beach, Florida, USA
- Jraidi, I., Chaouachi, M., Frasson, C. “Affective Computing for Student Uncertainty Modeling”, Artificial Intelligence in Education, Christchurch, New Zealand

## 2010

- Chaouachi, M., Frasson, C. “Exploring the Relationship between Learner EEG Mental Engagement and Affect”, ITS 2010: 10th International Conference on Intelligent Tutoring Systems, Pittsburgh, Pennsylvania, USA
- Jraidi, I., Chaouachi, M., Frasson, C. “Self-Esteem Conditioning for Learning Conditioning”, ICALT 2010: 10th IEEE International Conference on Advanced Learning Technologies, Sousse, Tunisia
- Jraidi, I., Chaouachi M., Frasson, C. “Enhancing Learner Self-Esteem for Learning Improvements”, 23rd International FLAIRS Conference, AAAI Press, Daytona Beach, Florida, USA
- Chaouachi, M., Chalfoun, P., Jraidi, I., Frasson, C. “Affect and Mental Engagement: Towards Adaptability for Intelligent Systems”, 23rd International FLAIRS Conference, AAAI Press, Daytona Beach, Florida, USA

## 2009

- Jraidi, I., Heraz, A., Chaouachi, M., Frasson, C. “New Architecture of a Multi Agent System which Measures the Learner Brainwaves to Predict his Stress Level Variation”, In T. Bastiaens, J. Dron & C. Xin (Eds.), World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2009 (pp. 2726-2733). Vancouver, Canada: AACE
- Chaouachi, M., Heraz, A., Jraidi, I., Frasson, C. “Influence of Dominant Electrical Brainwaves on Learning Performance”, In T. Bastiaens, J. Dron & C. Xin (Eds.), World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2009 (pp. 2448-2454). Vancouver, Canada: AACE
- Heraz, A., Jraidi, I., Chaouachi, M., Frasson, C. “Predicting Stress Level Variation from Learner Characteristics and Brainwaves”, AIED'2009: 14th International Conference on Artificial Intelligence in Education, IOS Press, Brighton, UK, 722-724