



HAL
open science

Etude de l'usage du stylo numérique en recherche en sciences de l'éducation : traitement et représentation des données temps issues de la trace d'écriture avec un stylo numérique.

Zaara Barhoumi

► **To cite this version:**

Zaara Barhoumi. Etude de l'usage du stylo numérique en recherche en sciences de l'éducation : traitement et représentation des données temps issues de la trace d'écriture avec un stylo numérique.. Education. École normale supérieure de Cachan - ENS Cachan, 2013. Français. NNT : 2013DENS0015 . tel-00845673

HAL Id: tel-00845673

<https://theses.hal.science/tel-00845673>

Submitted on 17 Jul 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ENSC-(n° d'ordre)

**THESE DE DOCTORAT
DE L'ECOLE NORMALE SUPERIEURE DE CACHAN**

Présentée par

Madame Zaara BARHOUMI

pour obtenir le grade de

DOCTEUR DE L'ECOLE NORMALE SUPERIEURE DE CACHAN

Domaine :

SCIENCES DE L'EDUCATION

**Etude de l'usage du stylo numérique en recherche en sciences de
l'éducation : Traitement et représentation des données temps issues
de la trace d'écriture avec un stylo numérique.**

Thèse présentée et soutenue à Cachan le 24 avril 2013 devant le jury composé de :

Jean-Marc LABAT	Professeur en informatique	Rapporteur
Georges-Louis BARON	Professeur en Sciences de l'éducation	Rapporteur
Jacques CRINON	Professeur en Sciences de l'éducation	Examineur
Eric BRUILLARD	Professeur en Informatique	Directeur de thèse
Françoise TORT	Maître de conférences	Co-directrice de thèse

ENS CACHAN/UMR STEF

61, avenue du Président Wilson, 94235 CACHAN CEDEX (France)

Résumé

La technologie de numérisation des données manuscrites a atteint son essor après l'invention du stylo numérique. Les recherches se sont concentrées alors sur le développement des applications IHM qui facilitent la manipulation des données encre numérique.

Dans un contexte de recherche, la technologie du papier et stylo numériques s'avère utile pour la collecte des données expérimentales dans différents domaines. Notre but est de fournir une aide pour le traitement et la visualisation des données, notamment les données temps, recueillies avec un stylo numérique dans deux cas d'usage en sciences de l'éducation à savoir les tests d'évaluation et la prise de notes d'observation.

Notre démarche consiste à étudier ces cas d'usage en terme de besoin dans quatre expériences concrètes et à développer, en participation avec l'utilisateur, des prototypes de système de traitement et de visualisation des données. Dans le but de concevoir un modèle général de tel système, nous avons analysé l'interface d'utilisation des prototypes en terme de fonctionnalités, de visualisation et d'interaction. Notre modèle décrit le processus de visualisation en mettant en œuvre trois types de tâches d'utilisateur à savoir les tâches d'analyses, les tâches de manipulation et les tâches d'interaction.

Mots clés : Conception centrée utilisateur, papier et stylo numériques, traitement de l'encre numérique, données temps, système de visualisation, représentations graphiques interactives.

Table des matières

Résumé.....	2
Table des matières.....	3
Chapitre 1. Contextes et problématique.....	6
1.1. Introduction.....	6
1.2. Contexte technologique.....	7
1.2.1. Les dispositifs de capture.....	8
1.2.2. L'encre numérique.....	10
1.2.3. Systèmes de reconnaissance.....	11
1.2.4. Traitement et visualisation de l'encre.....	12
1.2.5. La technologie Anoto et son fonctionnement.....	13
1.2.6. Les outils Kayentis.....	15
1.2.7. La plateforme de Kayentis Forms2.0.....	17
1.3. Contexte scientifique : usage du stylo et papier numériques.....	19
1.3.1. Interface utilisateur interactive.....	19
1.3.2. En éducation.....	21
1.3.3. Prise de note pour l'observation scientifique.....	24
1.4. Synthèse.....	26
1.5. Objectifs de la recherche.....	27
1.6. Méthodologie et outil de la recherche.....	28
1.7. Organisation du mémoire.....	30
Chapitre 2. L'usage du stylo numérique pour un test d'évaluation ordonné....	32
2.1. Introduction.....	32
2.2. Analyse préalable.....	32
2.3. Contexte et objectif de l'expérience.....	34
2.4. Organisation de l'expérience.....	36
2.5. La conception et l'instrumentation des tests.....	37
2.5.1. L'administration des tests.....	40
2.6. Le traitement des données.....	41
2.6.1. Les indicateurs temps.....	42
2.6.2. Exploration théorique.....	42
2.6.3. Analyse de la trace.....	43
2.6.4. Les données traitées.....	45
2.6.5. Les « indicateurs » personnalisés.....	46
2.7. La visualisation des données.....	50
2.7.1. Les représentations individuelles.....	51
2.7.2. Les représentations collectives.....	55
2.7.3. Exemple d'analyse.....	62
2.8. Feedback de l'expérience.....	64
2.8.1. Instrumentation des formulaires.....	64
2.8.2. L'interface d'utilisation.....	66
2.9. Conclusion.....	67
Chapitre 3. L'usage du stylo numérique pour des tests non ordonnés.....	68
3.1. Introduction.....	68
3.2. L'expérience EUSPRIG.....	68

3.2.1.	Contexte et objectif de l'expérience	68
3.2.2.	Conception et instrumentation.....	69
3.2.3.	Administration du test	72
3.2.4.	Traitement des données.....	73
3.2.5.	Les représentations graphiques.....	74
3.2.6.	Feedback de l'expérience.....	80
3.3.	L'expérience RJC.....	81
3.3.1.	Contexte et objectif de l'expérience	81
3.3.2.	Conception et instrumentation du test.....	82
3.3.3.	Administration du test	84
3.3.4.	Traitement des données.....	85
3.3.5.	Les représentations graphiques.....	86
3.3.6.	Feedback de l'expérience.....	90
3.4.	Conclusion.....	92
Chapitre 4.	L'usage du stylo numérique pour la prise de notes d'observation	93
4.1.	Introduction.....	93
4.2.	L'objectif et le contexte de l'expérience	93
4.3.	L'analyse des besoins.....	95
4.4.	La conception et l'instrumentation des grilles d'observation	96
4.5.	Le traitement des données.....	99
4.6.	Les représentations graphiques.....	100
4.7.	Feedback de l'expérience.....	105
4.8.	Conclusion.....	107
Chapitre 5.	L'implémentation et l'analyse de la plateforme SEGELL	108
5.1.	Introduction.....	108
5.2.	Spécification fonctionnelle.....	108
5.2.1.	Le but.....	108
5.2.2.	Le principe de la solution.....	109
5.2.3.	Les processus de fonctionnement.....	109
5.3.	Scénario d'utilisation.....	110
5.4.	Modélisation de l'architecture et de la base des données.....	135
5.4.1.	L'architecture de l'application.....	135
5.4.2.	La base des données	136
5.4.3.	Calcul des indicateurs temps	140
5.5.	Caractéristiques de l'interface d'utilisation	141
5.6.	Conclusion.....	147
Chapitre 6.	Méthode d'analyse et modèle générique.....	149
6.1.	Introduction.....	149
6.2.	La méthode d'analyse	149
6.2.1.	L'instrumentation et le recueil des données.....	150
6.2.1.1.	Contexte d'usage.....	150
6.2.1.2.	L'instrumentation.....	150
6.2.1.3.	Les données recueillies	150
6.2.2.	Le système de traitement et de visualisation.....	151
6.3.	Notre modèle	156
6.4.	Conclusion.....	159
Chapitre 7.	Conclusion et perspectives.....	160
7.1.	Résumé des résultats	160
7.1.1.	Contribution théorique	160
7.1.2.	Outil développé.....	160
7.2.	Réalisations futures et perspectives	161

Références bibliographiques.....	163
Annexe A.....	168
Annexe B.....	175
Table des illustrations	178

Chapitre 1. Contextes et problématique

1.1. Introduction

Nos travaux de recherche portent sur la conception et la réalisation d'un outil et d'un modèle générique pour le traitement et la représentation des données temps issues de la trace d'écriture sur papier augmenté avec un stylo numérique. Nos travaux sont pluridisciplinaires et s'inscrivent dans deux principaux domaines de recherche : l'interaction homme machine et les sciences de l'éducation. Plus précisément, ces travaux touchent les axes de recherches suivants : la conception des interfaces utilisateur, la représentation des données temporelles, le traitement de la trace numérique et l'usage du stylo numérique pour la collecte de données de recherche en sciences de l'éducation.

En suivant une méthode de conception et de recherche participative, ces travaux reposent sur des expériences menées en collaboration avec les chercheurs en sciences de l'éducation au sein du laboratoire Sciences et Technologies en Éducation et Formation (STEF) et dans le cadre du projet ENEIDE¹.

Si la technologie propose de nouveaux moyens de capture et de recueil de données, le traitement et l'exploitation approfondie de ces données restent coûteux pour les scientifiques. Nous nous intéressons au traitement des données temporelles fines collectées par les chercheurs à l'aide d'un stylo numérique et nous proposons des représentations interactives pour enrichir l'analyse de ces données.

Dans ce chapitre, nous commençons par décrire le contexte technologique dans lequel nos travaux se contraignent. Ainsi, nous présentons la technologie Anoto² du papier et stylo numériques et les solutions informatiques de visualisation et de prétraitement des données stylo numérique de l'entreprise Kayentis³, partenaire dans le projet ENEIDE. Ensuite, nous parlons de l'usage de cette technologie, du stylo et papier numériques, dans les domaines de

¹ Projet ENEIDE : <http://www.capdigital.com/eneide/>

² www.anoto.com

³ www.kayentis.fr

recherche notamment en éducation. Ensuite, nous présentons quelques systèmes de visualisation et de synchronisation des données stylo numérique utilisés pour l'observation scientifique. Enfin, nous décrivons notre problématique et notre méthode de recherche ainsi que la contribution de ce travail.

1.2. Contexte technologique

L'écriture manuscrite est une modalité d'entrée puissante et polyvalente et très familière pour la plupart des utilisateurs. Pour cela, les travaux relatifs en IHM (Interaction Homme Machine) se sont concentrés essentiellement sur l'amélioration des interfaces d'interaction avec les dispositifs numériques pour les rendre plus conviviales, ergonomiques et intuitives tel qu'un stylo et un papier.

Les travaux (Wellner, 1993), (Kobayashi & Koike, 1998), centrés sur l'augmentation du papier tout en profitant de ses caractéristiques ergonomiques, visent à faciliter la liaison entre les informations sur papier et les dispositifs numériques⁴.

Afin de faciliter l'utilisation des dispositifs avec de nouvelles modalités d'interaction permettant d'écrire sur des interfaces numériques ou bien papiers, les systèmes de capture, de reconnaissance de l'écriture manuscrite et de traitement de données ont évolué. Même si chacun de ces processus s'inscrit dans un domaine de recherche différent, ils sont complémentaires et indispensables pour le fonctionnement de ces interfaces.

Dans le schéma ci-dessous, nous présentons les phases nécessaires du processus pour l'utilisation des interfaces à interaction stylo.

⁴ Pour une description plus détaillée de la taxonomie et l'historique d'apparition de ces technologies, nous recommandons de voir la thèse de Beat Signer (Signer, 2005) élaborés dans le cadre du projet Paper ++.

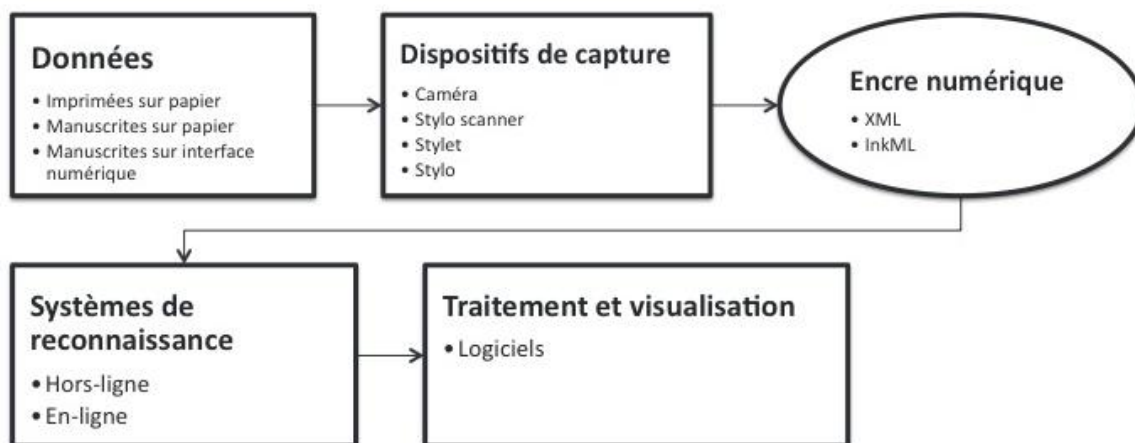


Figure 1 : Processus d'utilisation des interfaces et technologies à interaction stylo.

1.2.1. Les dispositifs de capture

Pour la capture des données (manuscrites sur interface numérique, manuscrite sur papier ou bien imprimé sur papier) il est nécessaire d'utiliser des dispositifs adaptés. La caméra a semblé le dispositif le plus adapté dans le cas du bureau numérique (Wellner, 1993).

Ensuite, pour surmonter le problème de la mobilité, il y a eu l'invention du stylo scanner doté d'une mini caméra permettant de capter, en image, l'écriture imprimée sur les documents. D'autre part, des dispositifs spécifiques étaient utilisés pour la capture des signes (les codes barres) ou bien du geste. Le lecteur de conducteur utilisé dans le projet Paper ++ en est un exemple particulier qui permet de capter un signal et le transformer en information.

Par ailleurs, les stylets informatiques ont été utilisés depuis les années 60 pour la capture de l'écriture manuscrite sur les tablettes PC et les dispositifs mobiles (PDA). Ces outils détectent la position exacte sur l'écran et enregistrent le temps de l'écriture et la position du stylet. Ces informations sont intéressantes dans le cas de l'analyse d'une trace manuscrite. Cependant, ces informations n'étaient pas récupérables dans le cas de la capture de l'écriture manuscrite sur papier.

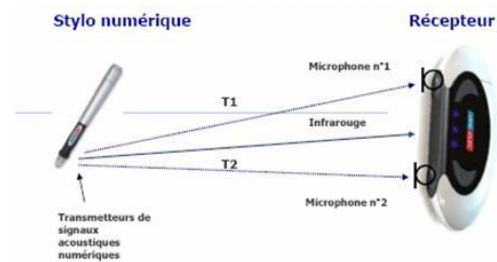
En 1995, le suédois Christer Fahraeus a inventé le C-PEN (Science & Vie, 2005), un stylo muni d'une caméra numérique infrarouge et d'un micro-processeur permettant de capturer l'écriture manuscrite sur papier. Le dispositif permettait la capture de l'écriture manuscrite sur un papier ordinaire mais ne permettait pas de connaître la position (x,y) du stylo sur le papier.

Le stylo a été commercialisé par Anoto, la société créée par l'inventeur de l'instrument en 2000.

Un autre dispositif de détection de l'écriture sur papier était commercialisé sous le nom de InkLink par Seiko Instrument dans les années 2000. Ce dispositif, attaché à un papier ordinaire (ou autre surface d'écriture comme le tableau blanc ou les transparents), calcule la position du stylo par rapport à deux points de références.



(a) InkLink⁵



(b) Z-Pen⁶

Figure 2 : Autres exemples de dispositif de capture de l'écriture sur papier.

Le stylo Z-Pen, actuellement commercialisé, applique le même principe (voir Figure 2 b). Cependant, cette technologie ne peut référencer qu'un document à page unique, puisqu'il faut calibrer le système à chaque changement de page.

Pour la détection à la fois de l'écriture manuscrite et de la position du stylo sur le papier, la société Anoto a développé la trame Anoto, une grille de points minuscules imprimés sur le papier. Par un algorithme spécifique, la trame permet au stylo la détection de sa position exacte sur le papier.

⁵ <http://www.amazon.com/Seiko-Instruments-LINK-IR-P-InkLink-Handwriting/dp/B00006B7UH>

⁶ <http://www.danedigital.com/6-Zpen/>

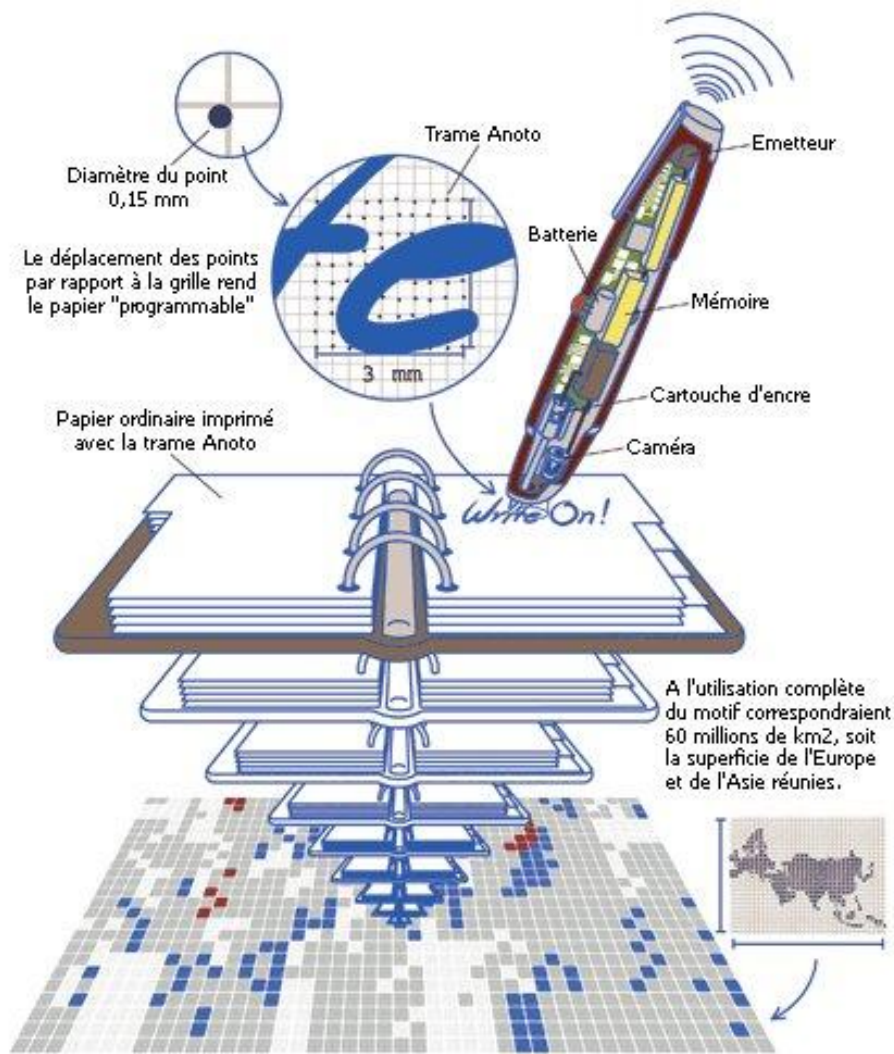


Figure 3 : La technologie Anoto du papier et stylo numériques⁷.

En plus de la capture de la position du stylo par rapport au temps, la technologie Anoto permet l'enregistrement des informations concernant les caractéristiques physiques des dispositifs utilisés ainsi que des détails destinés à des applications appropriées comme les systèmes de reconnaissance de l'écriture manuscrite ou les systèmes d'authentification et de traitement de formulaire pré rempli.

1.2.2. L'encre numérique

L'encre numérique est une information binaire collectée à l'aide des dispositifs de capture de l'écriture sur les interfaces à interaction stylo. Ce type de trace est intéressant du fait qu'il contient l'information telle que l'utilisateur l'a introduite y compris du texte, du tracé ainsi

⁷ <http://www.anotogroup.com>

que l'image et le temps de l'écriture.

La difficulté de transmission, de traitement et de présentation de l'écriture manuscrite issue de différents dispositifs de capture, a abouti à l'apparition du standard W3C⁸ InkML, un langage de balisage de « l'encre numérique ».

Il s'agit d'un format de données en XML permettant de représenter sous forme informatique l'écriture manuscrite produite par un stylo sur interface numérique ou un stylo et papier numériques. Il a pour fonction de faciliter l'échange de fichiers d'écriture entre les applications. Il stocke les informations relatives aux caractéristiques de l'outil effectuant l'enregistrement ainsi que les données des déplacements du stylo, ce qui permet son utilisation par les applications orientées reconnaissance d'écriture.

1.2.3. *Systèmes de reconnaissance*

La reconnaissance de l'écriture manuscrite, ou imprimée, est un traitement informatique qui permet de traduire un texte écrit, ou imprimé, en un texte codé numériquement. Il existe deux types de reconnaissance :

- la reconnaissance hors-ligne qui généralement travaille sur une image instantanée de l'encre et permet d'extraire des formes à partir de cette image. C'est le type de reconnaissance adéquat pour le traitement asynchrone de données telles que les documents imprimés ou pour le tri postal.
- La reconnaissance en-ligne qui travaille sur un échantillon d'encre constitué d'un ensemble de coordonnées enregistrées dans le temps. Ce type de reconnaissance permet de suivre le tracé et de connaître les moments de posés et levés du stylo ainsi que son inclinaison et sa vitesse.

Aujourd'hui, les systèmes de reconnaissance sont beaucoup plus efficaces. En effet, la plupart des interfaces à interaction stylo sont équipées de système de reconnaissance. Les recherches

⁸ <http://www.w3.org/TR/InkML/>

se sont donc focalisées de plus en plus sur d'autres problématiques tels que le traitement et la visualisation de l'encre numérique.

1.2.4. Traitement et visualisation de l'encre

Le processus de traitement de l'encre numérique change selon l'usage et la complexité des applications auxquelles ces données sont destinées. Par exemple, si ces données représentent des gestes ou des commandes informatiques, elles seront acheminées vers un traitement d'exécution. D'autres données vont être calculées et représentées. Dans d'autres contextes d'usage, les données textuelles par exemple ont besoin d'être visualisées sur une interface de traitement de texte.

Les applications de traitement et de visualisation de l'encre numérique sous format InkML sont donc propres à leur contexte d'utilisation. Ainsi, ce langage de balisage de l'encre numérique permet de présenter l'écriture sous différentes formes. La représentation de la Figure 4 est un simple exemple d'écriture du mot « hello » extrait de la documentation officielle du langage InkML :

```
<ink>
<trace> 10 0, 9 14, 8 28, 7 42, 6 56, 6 70, 8 84, 8 98, 8 112, 9 126, 10 140, 13 154, 14 168,
17 182, 18 188, 23 174, 30 160, 38 147, 49 135, 58 124, 72 121, 77 135, 80 149, 82 163, 84
177, 87 191, 93 205 </trace>
<trace> 130 155, 144 159, 158 160, 170 154, 179 143, 179 129, 152 128, 140 136, 131 149, 126
163, 124 177, 128 190, 150 208, 163 210, 178 208, 192 201, 205 192, 214 180 </trace>
<trace> 227 50, 226 64, 225 78, 227 92, 228 106, 228 120,
229
230 148, 234 162, 235 176, 238 190, 241 204 </trace>
<trace> 282 45, 281 59, 284 73, 285 87, 287 101, 288 115,
290
291 143, 294 157, 294 171, 294 185, 296 199, 300 213
</trace>
<trace> 366 130, 359 143, 354 157, 349 171, 352 185, 359
197, 371 204, 385 205, 398 202, 408 191, 413 177, 413 163,
405 150, 392 143, 378 141, 365 150 </trace>
</ink>
```



Figure 4 : Un extrait de données encre numérique représenté en InkML⁹.

Le contenu d'un document InkML est délimité par la balise <ink>. Chaque lettre, représentée dans une balise <trace>, est décomposée en plusieurs points représentés par leurs coordonnées x et y.

Il est également possible de spécifier comment représenter chaque point d'un élément trace grâce à la balise <traceFormat> qui permet de lister des informations spécifiques comme

⁹ <http://www.w3.org/TR/InkML/>

l'orientation du crayon, la pression exercée, ou également l'accélération pour pouvoir reproduire le tracé. De même, il est possible de spécifier des informations qui concernent le contexte de la capture par le biais de balises comme <inkSource> pour spécifier l'outil utilisé, <brush> pour obtenir des informations sur le pinceau utilisé (couleur, taille, écrire ou effacer...), <timestamp> pour avoir des informations temporelles de l'écriture. Enfin, la balise <traceGroup> permet de regrouper un ensemble de tracés pour désigner une sémantique comme par exemple un schéma explicatif, un tableau, un paragraphe ou bien une note.

En conséquence, InkML est utilisé pour différentes applications telles que la reconnaissance du geste, l'écriture mathématique ou encore les applications de traitement de formulaires pré imprimés mais aussi pour les applications de visualisation de l'écriture.

1.2.5. La technologie Anoto et son fonctionnement

L'originalité de cette technologie repose sur la technique de traitement automatique de formulaire pré imprimé. Cette technique permet aux utilisateurs de délimiter des zones d'écriture dans le document sur lequel la trame Anoto va être imprimée et de définir en plus des métadonnées relatives à chaque zone. C'est pour cette raison que nous avons opté pour utiliser cette technologie dans notre étude.

Sa trame repose sur un algorithme mathématique qui positionne les points de trame avec un décalage particulier par rapport à une matrice bidimensionnelle virtuelle ayant un pas de 0,3 mm (voir Figure 5). Chaque point peut avoir 4 positions différentes par rapport au nœud de la matrice virtuelle qui sont : au dessus, au dessous, à droite, à gauche. Les 4 positions sont codées sur 2 bits par point. La trame ainsi disposée permet de connaître la position exacte du stylo de façon unique, et ce sur une surface équivalente à près de 60 millions de km², sur le document.

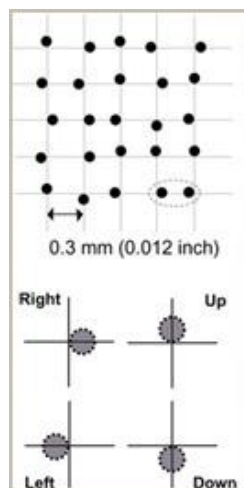


Figure 5 : Représentation d'une pointe d'un stylo dans la matrice de la trame d'Anoto¹⁰.

La trame peut être imprimée sur un papier, ou bien sur une zone spécifique d'un document, avec une résolution de 1000 dpi. L'impression nécessite une demande d'attribution d'un « *pattern* » en fonction des besoins, auprès d'Anoto, qui gère l'ensemble des modèles.

Lorsque le stylo se déplace sur le papier imprimé avec la trame d'Anoto, des prises de vue sont enregistrées avec la caméra infrarouge placée à la mine du stylo. La trame est imprimée avec du noir de carbone, qui a la propriété d'absorber le rayonnement infrarouge de la caméra. Elle permet d'identifier la position des points de trame lorsque le stylo écrit sur le papier. L'encre du stylo ne doit pas contenir du noir de carbone pour ne pas interférer avec la trame.

Ensuite, les informations captées sont transmises via un émetteur Bluetooth ou bien via un port USB vers un téléphone portable ou bien un ordinateur. Ces informations manuscrites enregistrées comportent pour chaque petit trait : les coordonnées (X, Y) de ses points de début et de fin, les données d'inclinaison du stylo et les données temps. Ces données sont compatibles avec le langage de description de l'encre numérique InkML, et donc sont exploitables par différentes applications web ou bien de bureautique.

La modélisation d'un formulaire est réalisée à partir du logiciel de développement proposé par Anoto : FDK (Forms Development Kit), qui est une extension (plug-in) d'Adobe® Acrobat Professional 6. L'outil de conception de formulaires (FDK) est utilisé pour concevoir les formulaires et les documents à utiliser avec des fonctionnalités Anoto. Utilisées par un

¹⁰ <http://www.anoto.com/the-paper-3.aspx>

graphiste ou bien un développeur, les fonctionnalités de base de cet outil sont :

- Convertir un fichier de présentation d'un formulaire en PDF
- Ouvrir le fichier *PDF dans Adobe® Acrobat Professional* ;
- Activer le formulaire en utilisant une licence pattern Anoto ;
- Ajouter les zones d'items et leurs propriétés dans le formulaire ;
- Générer le fichier de définition de la structure du formulaire *Paper Application Definition (PAD)* ;
- Générer des images de fond du formulaire ;
- Générer un fichier *ZIP* incluant les fichiers ci-dessus (*PAD, PNG et PDF*) ;
- Générer un fichier *postscript* prêt pour imprimer le document avec la trame.

L'outil *Paper SDK* permet de personnaliser la conception de l'interaction sur le papier par le paramétrage des zones du formulaire. Le plug-in dans Adobe ® Acrobat Professional peut donc être amélioré pour assigner une quantité illimitée de fonctions et de méta données à tout endroit de la feuille de papier, selon des spécifications affectées par les programmeurs.

Pour utiliser la technologie Anoto, la société met à disposition de ses clients, commerciaux ou bien universitaires, un ensemble d'outils informatiques (*Paper SDK, Forms Development Kit plug-in, et le Development SDK*). Ces outils permettent la conception des documents avec la disposition de la trame Anoto, de l'imprimer et puis d'intégrer les données du stylo dans un système d'information. Ainsi, les utilisateurs de ces outils ont la possibilité de développer des applications adaptées à leurs domaines d'expertises.

Les utilisateurs devront donc développer des solutions pour la conception et l'intégration des données issues du stylo numérique. Ils pourront demander une prestation auprès d'un partenaire service informatique de Anoto pour réaliser cette tâche. Dans notre recherche, nous avons opté pour cette solution dans une collaboration avec l'entreprise Kayentis dans le projet Eneide.

1.2.6. Les outils Kayentis

L'entreprise Kayentis, spécialisée dans le traitement des formulaires pour les métiers de la santé, nous a prêté 8 stylos (et nous a vendu 2 autres stylos) *Logitech Io2* et un ensemble d'outils, basés sur le *SDK* de Anoto. Ces outils sont développés chez HP et sont commercialisés sous forme de suites de logiciels utilisables avec les imprimantes HP notamment pour les métiers de santé.



Figure 6 : Le stylo numérique Logitech Io2.

Nous avons utilisé ces outils, ré-habillés par Kayentis, pour faciliter le contrôle et le traitement des données de la phase d'instrumentation du formulaire jusqu'à la récupération d'un fichier XML contenant la trace d'écriture.

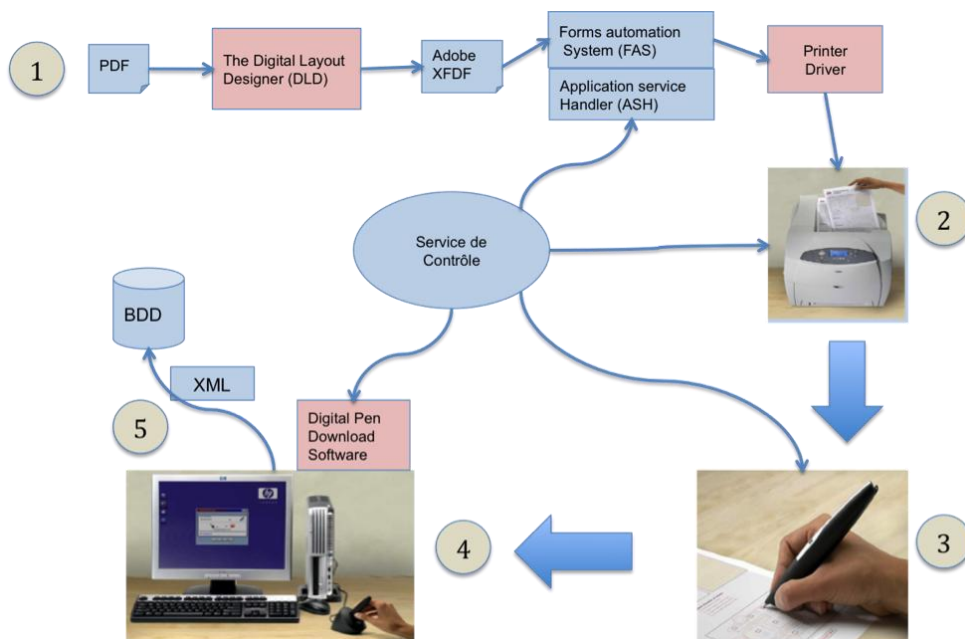


Figure 7 : Le fonctionnement de la technologie Anoto avec les services de Kayentis.

Pour effectuer une expérience de remplissage de formulaires avec stylo numérique en utilisant la technologie de Kayentis, il faut passer par 5 étapes (voir Figure 7), à savoir :

- La conception et l'instrumentation du formulaire sur le logiciel *Adobe® Acrobat Professional*,
- L'impression du formulaire sur une imprimante HP à l'aide d'un plugin adapté,
- Le remplissage du formulaire,
- Le transfert des données par « dock » du stylo à l'aide d'un logiciel dédié,
- La correction et la validation de la reconnaissance de l'écriture sur une plateforme sur Internet,
- Enfin l'export des données en XML à l'aide d'un Web service.

1.2.7. La plateforme de Kayentis Forms2.0

Forms2.0 est une plate-forme permettant la récupération, la vérification et l'exportation des données du « Papier numérique ». Elle permet de télécharger « l'encre numérique » provenant du stylo numérique Anoto. *Forms2.0* se compose d'une interface utilisateur permettant le filtrage des données, le traitement et la validation de l'écriture manuscrite, ainsi que la génération de ces données sous les formats SVG, PDF, Word, Excel, etc., pour intégration dans une application tiers.

Les 4 grandes étapes du traitement des données implémentées sous *Forms2.0* sont les suivantes :

- Import de données pour le pré remplissage des formulaires
- Décryptage de l'encre et la reconnaissance d'écriture
- Interface de correction et validation de la reconnaissance de l'écriture par l'utilisateur
- Export des données aux formats XML (via le Web service) et CSV (via le portail Web)

Chacun des utilisateurs aura un profil par projet qui lui permet d'avoir des droits pour accéder aux différentes fonctionnalités. En tant que gestionnaire pour nos projets sur la zone 'Eneide' notre profil nous permet de :

- Importer un formulaire dans Forms2.0. ;
- Télécharger les informations écrites avec le stylo numérique dans Forms2.0. ;
- Visualiser les formulaires : voir les images des exemplaires des formulaires téléchargées ;
 - Visualiser les données : voir l'écriture sur les images des exemplaires (la valeur du champ de saisie) ;
 - Correction de données : corriger les champs mal reconnus ou ne respectant pas une règle définie dans l'instrumentation ;
 - Administration des formulaires : avoir accès à l'onglet de fonctionnalités de paramétrage des formulaires pour nos projets. ;
 - Administration des utilisateurs : avoir accès aux onglets de fonctionnalités de paramétrage des utilisateurs et des profils ;
 - Export des données en formats XML et CSV : avoir accès à l'import/export des données via le portail et le Web Service.

Pour chaque expérimentation dans notre étude, nous avons suivi les phases suivantes :

- Concevoir un formulaire tenant compte des contraintes d'exploitation du stylo numérique utilisé et du contexte de l'expérience.
- Instrumenter le formulaire sur *Adobe Acrobat professionnel* à l'aide d'un plugin *Kayentis (Digital Layout Designer)*. Ce plugin permet de définir les trames et les zones de saisie. On déclare pour chaque zone le nom, la langue et le seuil de la reconnaissance, et éventuellement les valeurs correspondantes aux cases à cocher et aux boutons radio. Ensuite, on imprime le formulaire à l'aide d'une imprimante compatible avec un pilote spécial (*Digital Paper Printer Driver*) pour avoir la trame.
- Faire le remplissage du formulaire avec le stylo numérique soit par le concepteur du formulaire soit par un autre utilisateur. Ensuite, il s'agit de transférer les données depuis le stylo (grâce au pilote *Pen Software*) sur la plate-forme *Forms2.0*.
- Visualiser les données écrites sous forme d'images depuis *Forms2.0*. Il est possible d'interagir sur ces images pour vérifier ou valider la reconnaissance de l'écriture.

Ensuite on les récupère à l'aide d'un programme Web Service sous la forme d'un fichier XML. Ce fichier contient le nom de chaque zone et la valeur écrite, ainsi que les instants du début de l'écriture et ceux de la fin dans chaque item.

1.3. Contexte scientifique : usage du stylo et papier numériques

Les propriétés physiques du papier nous permettent de faire plusieurs tâches comme écrire, dessiner, marquer et raturer. Le papier continue à avoir un rôle majeur dans tout travail intellectuel. De plus, le papier est un instrument portable et mobile efficace aussi bien sur le terrain qu'en situation d'examen.

L'invention et le développement de la technologie du stylo et papier numériques a provoqué la curiosité des usagers professionnels et industriels (Belgnaoui, 2008), mais également a attiré l'attention des chercheurs et universitaires en différents domaines durant ces 10 dernières années.

Au début, les recherches s'intéressaient à l'exploitation du papier augmenté en tant qu'interface utilisateur (Liao, Guimbretière, & Hinckley, 2005), (Lee et al. , 2008). Dans ce cas, des fonctionnalités d'interaction se sont développées dans le but de faciliter la manipulation de données numériques à travers l'interface stylo et papier.

Ensuite, en éducation, des expériences ont été menées (Oviatt, Arthur, & Cohen, 2006), (Maldonado, Lee, & Klemmer, 2006), (Miura, Kunifuji, & Sakamoto, 2007), (Pietrzak, Malacria, & Lecolinet, 2010), concernant différents objectifs d'usages comme le travail collaboratif, la prise de notes en classe, l'évaluation ou le tutorat dans un contexte éducatif.

Enfin, d'autres travaux (Yeh et al., 2006), (Tabard, Mackay, & Eastmond, 2008) ont utilisé cette technologie pour faciliter la collecte de données en recherche scientifique et la synchronisation des données manuscrites avec d'autres données hétérogènes.

1.3.1. Interface utilisateur interactive

Avec l'apparition de la technologie Anoto, le problème de numérisation de l'écriture étant résolu, de nouveaux intérêts sont apparus dans l'usage du papier augmenté. En effet, parmi les

nouveaux problèmes la possibilité de rendre le papier plus interactif jusqu'au point qu'il devienne une interface utilisateur pour interagir avec un système numérique.

Les *Paper Augmented Digital Documents*, PADD (Guimbretière, 2003), sont des documents numériques qui peuvent être manipulés soit sur un écran d'ordinateur soit sur papier. Les PADD, qui utilisent la technologie Anoto, sont principalement des documents numériques qui peuvent être imprimés sur du papier tramé et annotés avec un stylo numérique pour être en mesure de numériser les données nouvellement écrites sur ces documents. En utilisant une application spécialement programmée, les nouvelles données sur le papier imprimé peuvent être fusionnées avec la version numérique originale du document. Une base de données va garder la trace de ces changements (voir Figure 8).

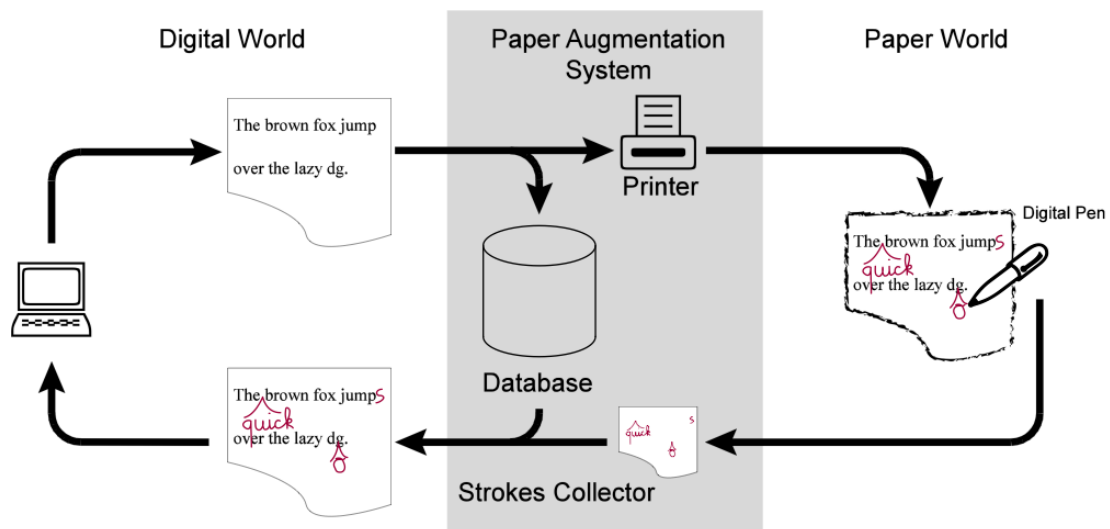


Figure 8 : L'infrastructure des PADD (Guimbretière, 2003).

L'auteur des PADD a collaboré ensuite avec d'autres chercheurs (Liao et al., 2005) pour travailler sur *PaperCraft*, un système de commande gestuelle à base de papier tramé interactif. Il utilise le papier comme une interface utilisateur pour interagir avec les documents numériques.

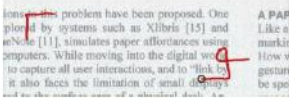

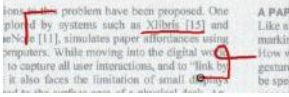

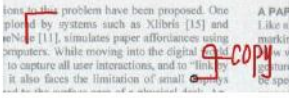

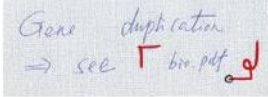
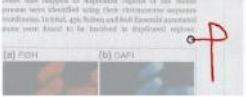
	Command on Page 1	Command on Page 2
Excerpt		
Excerpt with keyword		
Excerpt with command name		
Hyperlink		

Figure 9 : Extrait des commandes traitées par PaperCraft pour l'annotation de document (Liao et al., 2005).

Le système permet aux utilisateurs d'annoter un document imprimé ou dessiner une commande gestuelle pour déclencher des opérations telles que copier et coller une zone de document, ou créer un lien.

Ces recherches ont encouragé à l'usage du papier et stylo numériques dans différents domaines de recherche. D'ailleurs, le système *PaperCraft* était à la base de développement d'autres systèmes comme *ButterflyNet*¹¹ utilisé en recherche en biologie et en éducation.

1.3.2. En éducation

La technologie Anoto a facilité la capture de l'écriture manuscrite par le papier augmenté. Plusieurs recherches l'ont utilisé pour développer des systèmes qui combinent les données numériques et les données sur papier. En éducation, nous avons remarqué que l'usage de cette technologie a concerné le travail collaboratif en classe et le partage des notes manuscrites, d'une part, et le tutorat et l'évaluation en classe, d'autre part.

¹¹ <http://hci.stanford.edu/research/butterflynet/butterflynet/>

Maldonado et al. (2006) ont utilisé le stylo numérique dans le cadre du travail collaboratif pour l'échange des idées spontanément exprimées sur papier dans des cours de conception informatique. Ils ont montré que cette technologie hybride qui maintient la souplesse des cahiers et du papier pour communiquer les idées peut assurer la fluidité nécessaire entre les coéquipiers dans des projets de conception.

AirTransNote (Miura et al., 2007) est un système interactif qui recueille l'écriture et les dessins manuscrits des étudiants et les transmet au PC de l'enseignant, afin que l'enseignant puisse suivre de près le travail de ses élèves en classe de mathématique. *AirTransNote* permet également à l'enseignant de fournir une rétroaction sur les PDA des étudiants.

Une première expérience (Miura, Kunifuji, Shizuki, & Tanaka, 2005) avec ce système emploie les stylos numériques de *Seiko InkLink*. Les élèves qui l'ont utilisé saisissent des notes sur une feuille de papier ordinaire et connectent le capteur ultrason à une interface PDA pour le transfert des données.

Dans une seconde expérience (Sugihara, Miura, & Kunifuji, 2010) de séance en classe de mathématiques, les auteurs ont eu recours à intégrer la technologie Anoto pour réduire les tâches des étudiants pour transmettre l'écriture vers l'écran de l'enseignant. Ils ont couplé une solution commerciale, pour l'instrumentation et l'intégration des données écrites par les étudiants, avec le système *AirTransNote*. L'enseignant et les étudiants avaient la possibilité de visualiser les résultats automatiquement générés par le système *AirTransNote*.

Par ailleurs, des recherches s'intéressaient à la capture du déroulement d'une séance en classe (Brotherton & Abowd, 2004), par combinaison de toutes les données capturées en un flux unique pour le diffuser à travers une interface web ou des vidéo téléchargeables. Ainsi, *U-Note* (Pietrzak et al., 2010) est un système permettant de capturer puis de consulter le contexte d'une séance en classe (niveau collège et lycée, 11-21 ans) par rapport à la note prise par l'élève. En classe, l'élève prend ses notes sur son cahier augmenté avec un stylo numérique pendant que l'enseignant donne des explications orales ou bien écrites au tableau et diffuse des documents numériques. Lors des révisions à la maison, l'élève dispose d'une application de bureau qui complète le cahier. L'interface de cette application permet de visualiser à la fois les notes écrites par l'élève et les données fournies par l'enseignant. Ces données sont

synchronisées par rapport au moment de la prise de note faite par l'élève.

Dans l'enseignement des mathématiques, (Oviatt et al., 2006) ont comparé l'utilisation de cette technologie avec d'autres technologies (ordinateur avec clavier, tablette graphique avec stylo...) dans la résolution de problème par des élèves de collège. Les auteurs ont utilisé les temps recueillis avec le stylo numérique pour calculer la durée totale de résolution d'un problème. Leurs résultats montrent que les élèves prennent moins de temps pour résoudre les problèmes avec le stylo numérique qu'avec le clavier et la tablette graphique. Cette expérience représente le premier cas d'usage de cette technologie pour des tests de résolution de problèmes en éducation.

Utilisé pour le tutorat, *Nowten'Pen* (Lee et al., 2008) est un système de commande basé sur une interface papier et un stylo équipé d'un haut-parleur et d'un micro-processeur. Ce système utilise la technologie Anoto comme interface pour interagir avec un programme de tutorat sonore à travers le stylo. Le prototype instrumenté pour cette expérience représente un exercice de statique sur papier tramé destiné aux étudiants en mécanique. Un autre programme, installé dans la mémoire du stylo, déclenche le son synthétique du tuteur à chaque contact du stylo avec une zone spécifique du papier.

L'architecture de ce système est faite de façon à obliger les étudiants à écrire dans un ordre spécifique. Par exemple, lors de l'élaboration du diagramme du corps libre, l'étudiant doit d'abord dessiner un système de coordonnées, puis le corps, et enfin les forces. De même, lors de l'écriture d'une équation d'équilibre, l'étudiant doit d'abord spécifier le type de l'équation et puis signer la convention. Selon les auteurs, contraindre l'ordre de résolution de problèmes augmente la précision de la reconnaissance en réduisant les possibilités de confusion. En outre, l'approche est cohérente avec la façon dont de nombreux instructeurs enseignent. Cependant, un essai correct dans un ordre différent peut être considéré comme une erreur par le système. Finalement, cet outil de tutorat se caractérise par son composant de reconnaissance de forme, comme les flèches des forces d'un objet par exemple, et de formules mathématiques manuscrites.

Dans ce dernier cas, les chercheurs ont utilisé un système commercialisé avec le stylo pour la

plupart des tâches d'interactions nécessaires pour le bon fonctionnement du tutorat. De plus, ils ont personnalisé des fonctionnalités pour la reconnaissance de forme car la matière du tutorat nécessite ce traitement spécifique. De même pour le cas de l'outil *Air@TransNote* (Sugihara et al., 2010), où les chercheurs ont également couplé une solution commerciale, pour l'instrumentation et le recueil des données écrites par les étudiants, avec le système.

Enfin, selon l'objectif et le contexte d'usage de la technologie du papier et stylo numériques en éducation, les outils informatiques utilisés par les chercheurs varient. En effet, avec la technologie Anoto, pour résoudre le problème d'instrumentation des formulaires, généralement, les solutions commerciales étaient satisfaisantes pour exploiter le maximum d'interaction du papier tramé. En revanche, pour la reconnaissance, ces systèmes ne permettaient pas de reconnaître plus que l'écriture. Pour cela, certains chercheurs se sont efforcés de développer des composants spécifiques à la reconnaissance de forme ou du geste, pour satisfaire les objectifs de leurs recherches. Finalement pour le traitement et la visualisation des données, même si les outils commerciaux fournissent des solutions variées, chaque recherche est différente et a des objectifs spécifiques. Ce qui demande aux chercheurs de faire du retraitement personnalisé des données recueillies notamment pour pouvoir visualiser ces données sur des outils de visualisation interactifs adéquats.

1.3.3. *Prise de note pour l'observation scientifique*

Un autre domaine a utilisé la technologie du papier et stylo numérique, il s'agit de la prise de note pour l'observation. En effet, l'enregistrement instantané des données de type notes d'observation sur le terrain est le principal avantage d'usage de cette technologie. Ainsi, l'observateur prend des notes sans avoir besoin de chronomètre sur la moment, ni ensuite, de transcription de son écriture. Les chercheurs ont utilisé cette technologie essentiellement pour synchroniser l'écriture avec d'autres données hétérogènes comme l'enregistrement vidéo ou sonore de l'observation.

Pour soutenir la recherche en biologie sur le terrain, des chercheurs (Yeh et al., 2006) ont utilisé l'infrastructure des travaux *PapierCraft* pour développer le système *ButterflyNet*. Essentiellement utilisé pour la prise de note d'observation sur le terrain, *ButterflyNet* est un système d'accès aux données recueillies avec des interfaces mobiles (stylo numérique, caméra, appareil photo numérique, GPS, etc.) pour les scientifiques de terrain. Il permet aux

scientifiques de trouver rapidement leurs notes manuscrites, photographies, lectures de sonde, les journaux de trajet GPS, et le contenu d'autres recherches. Les données écrites dans le carnet sont fusionnées avec les photos numériques ou les enregistrements vidéo dans l'interface du système *ButterflyNet*.

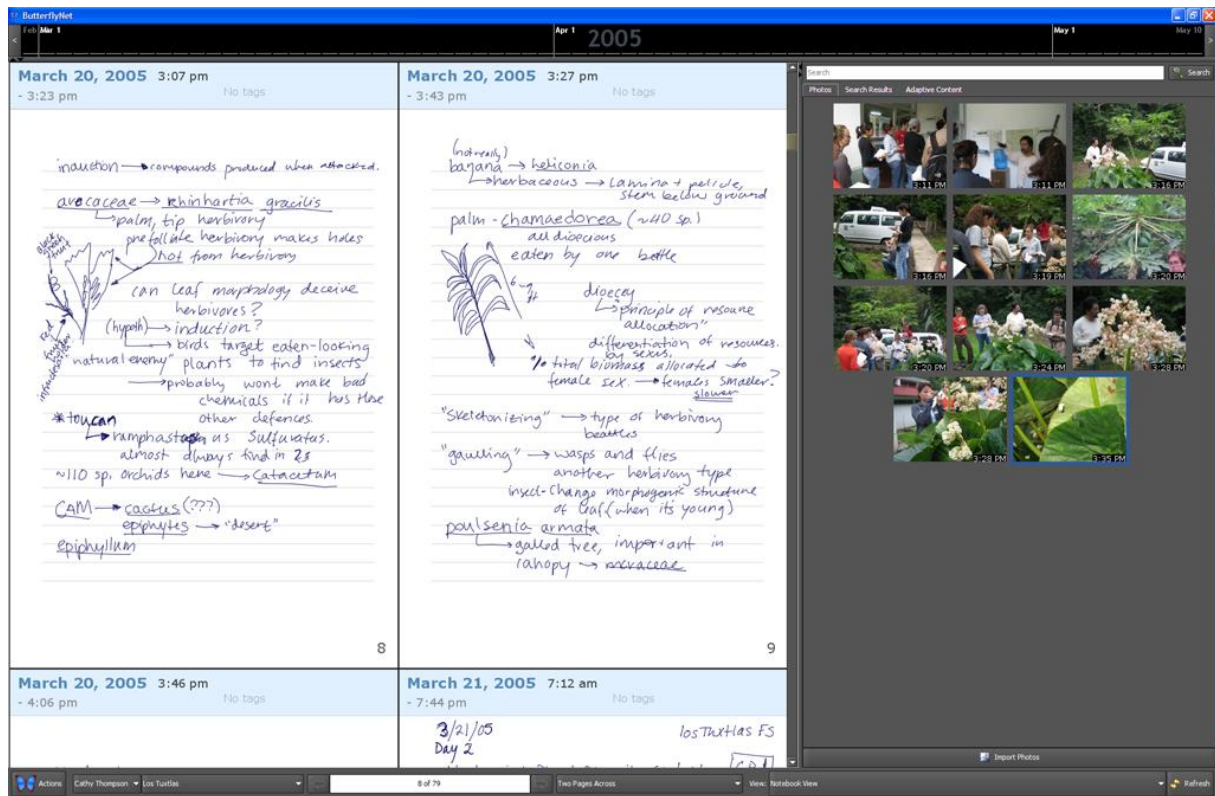


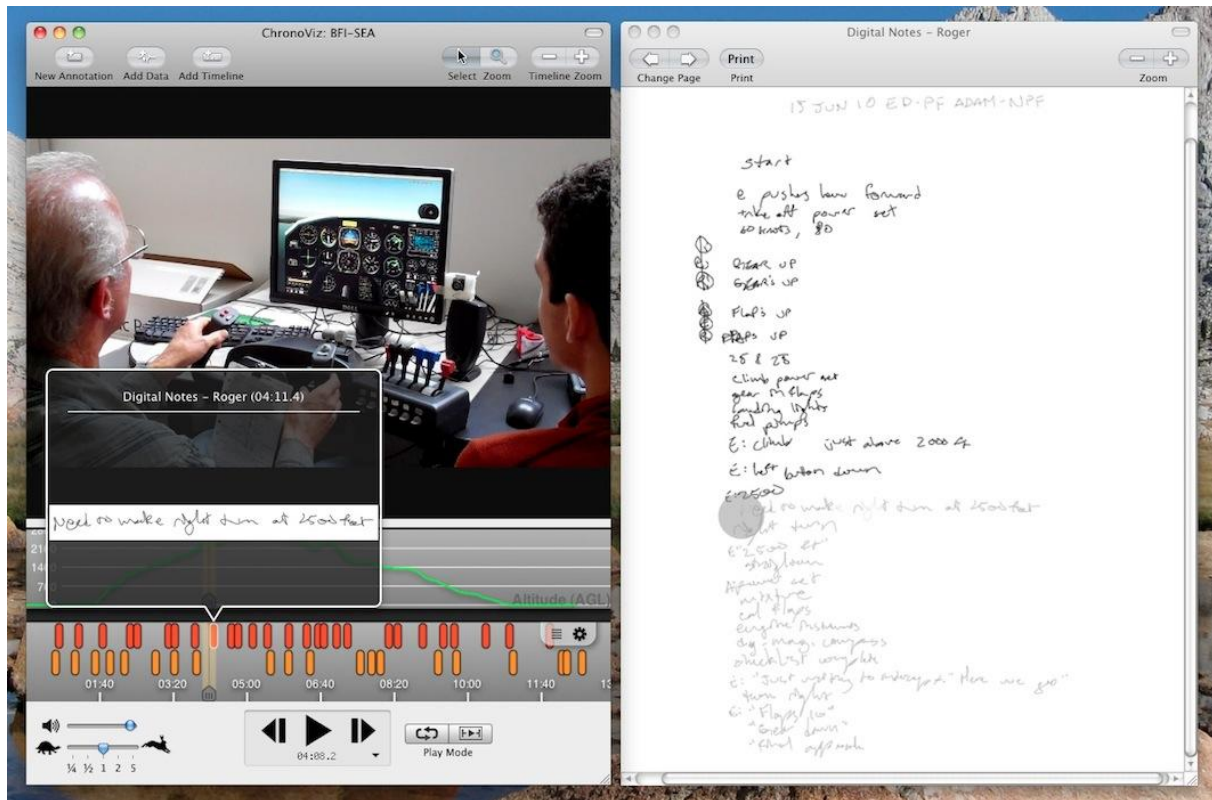
Figure 10 : L'interface du navigateur de l'outil ButterflyNet¹².

L'application se compose d'un navigateur chronologique qui présente plusieurs vues des données hétérogènes. Ces vues sont automatiquement liées par leur temps, leur emplacement, ou d'autres métadonnées. Il existe également une interface de type tableur qui permet la manipulation des données.

Plus récemment, des chercheurs en science cognitive et IHM (Weibel, Fouse, Hutchins, & Hollan, 2011) ont intégré la technologie du papier et stylo numériques pour faciliter le processus de collecte de données dans de la recherche observationnelle. Le stylo, permettant la capture automatique d'annotations, est utilisé dans les séances d'observation avec d'autres

¹² <http://hci.stanford.edu/research/butterflynet/butterflynet/>

moyens de collecte de données. Le système exploite les flux de prise de notes et les intègre dans un outil de synchronisation de données multimodales basées sur le temps. A la base, *ChronoViz* est un outil pour l'annotation de la vidéo en combinaison avec d'autres sources de données.



**Figure 11 : L'interface de ChronoViz :
l'annotation numérique s'affiche sur la frise temporelle en bas à gauche,
et la représentation numérique du papier à droite (Weibel et al., 2011).**

Le but de cette expérience est de diminuer le temps d'analyse des vidéos par l'intégration de l'annotation numérique prise au cours de l'observation. Cette annotation peut ensuite être enrichie par une annotation post observation sur la vidéo.

1.4. Synthèse

Outre l'exploitation des caractéristiques d'interactivité de la technologie du papier et stylo numérique pour son usage comme interface utilisateur, nous soulignons essentiellement deux cas d'usage principaux.

Le premier est l'usage en éducation. Ici les objectifs varient : travail collaboratif, prise de note

en classe, partage de la classe et tutorat. Ainsi, les outils d'instrumentation et de traitement des données sont différents.

Cependant, nous remarquons le manque d'intérêt de l'usage de cette technologie en éducation dans l'évaluation, malgré les solutions de pré-instrumentation et d'automatisation que la technologie peut apporter.

De plus, nous remarquons que le traitement des données récupérées n'a pas été approfondi pour pouvoir suivre, chronologiquement, la trace d'écriture et d'interaction de l'utilisateur avec cette technologie.

Le deuxième cas d'usage, qui est plus ou moins récent, est la prise de note d'observation scientifique. Dans ce cas, nous remarquons que la technologie a été utilisée avec un ensemble d'outils de recueil de données habituellement utilisé dans ce domaine. Outre les avantages de numérisation des notes manuscrites et la flexibilité de la technologie comme interface utilisateur, l'apport de ce nouvel instrument de recueil de données est la possibilité de synchroniser l'écriture avec d'autres média.

Cependant, même si des systèmes de visualisation de ces données synchronisées ont été mis en place, nous pensons que cette technologie peut être mieux exploitée dans ce domaine pour, par exemple, le codage des grilles d'observation. Ainsi, il est possible de coder ces grilles à travers des métadonnées prédéfinies à l'instrumentation.

1.5. Objectifs de la recherche

Notre travail de recherche aborde la question d'usage du papier et stylo numériques pour le recueil des données de recherche en sciences de l'éducation. Nous mettons en œuvre l'apport de l'analyse des données temps pour deux cas d'utilisation du stylo numérique : les tests d'évaluation et la prise des notes d'observation. De cette problématique découlent deux questions principales :

- Quelles stratégies de recueil avec l'instrumentation disponible fournissent des données pertinentes ?

- Quelles formes de traitement et de visualisation sont nécessaires, utiles et intéressantes ?

Notre objectif est d'étudier l'usage de la technologie du papier et stylo numériques par des chercheurs en sciences de l'éducation dans le but de concevoir un modèle générique de système d'aide à l'analyse des données temps issues d'un stylo numérique, ou d'autres outils de recueil, applicable à différents contextes expérimentaux de recherche en éducation.

Au niveau opérationnel, notre objectif est la conception et la mise en œuvre d'un prototype informatique, plateforme web, pour assister le chercheur à réaliser un ensemble de manipulations sur des données temps à travers une interface de visualisation dans un processus d'analyse de données de recherche.

1.6. Méthodologie et outil de la recherche

Il s'agit d'une étude exploratoire. La spécificité du système conçu, par rapport aux outils de visualisation précédemment cités (ChronoViz, ButterflyNet, U-Note), réside dans le fait qu'il permet de visualiser des données traitées et des représentations de ces données d'analyse scientifique, notamment les données temps. Ainsi, contrairement à un système d'analyse automatique d'interaction, il ne s'agit pas de produire un ensemble d'indicateurs spécifiques et des représentations destinées à l'utilisateur de l'interface. Dans notre cas, l'utilisateur de l'outil stylo numérique peut être un étudiant, un répondant, ou bien le chercheur lui-même. Tandis que l'utilisateur de notre système d'aide à l'analyse des données est le chercheur en science de l'éducation.

Habituellement, les spécifications d'un système sont déterminées à travers l'analyse de questionnaires destinés à l'utilisateur ou bien l'analyse de l'observation de l'utilisateur en situation d'utilisation du système, etc.

Dans notre cas d'étude exploratoire, l'utilisateur veut utiliser pour la première fois cette nouvelle méthode de recueil des données, et analyser pour la première fois ces données temps dans son contexte expérimental. De plus, il n'y a pas de système similaire existant.

Pour cela, nous avons suivi un processus de conception par prototypage exploratoire qui permet d'éclaircir le besoin de l'utilisateur et de préciser les spécifications du système. Nous

avons utilisé ce processus dans quatre expériences de cas d'usage de l'outil stylo numérique par des chercheurs en sciences de l'éducation.

Pour chaque cas d'usage, nous avons produit un prototype, conçu en participation avec l'utilisateur, et nous l'avons utilisé pour traiter les données recueillies au cours de l'expérience. L'analyse de chaque prototype, nous a permis de déterminer l'ensemble de spécifications en terme de traitements et de représentations utiles.

Concernant notre démarche de travail, nous avons commencé par une phase de découverte de la technologie du stylo numérique et de son usage dans la littérature. Nous avons également exploré la trace fournie par cet outil ainsi que le fonctionnement de la technologie Anoto et Kayentis. Parallèlement, nous avons effectué une petite expérience avec 3 étudiants en situation de test sur papier ordinaire. L'analyse chronologique des vidéos enregistrées durant le test nous a permis de mieux comprendre la situation de passation d'un test sur papier.

Ensuite, nous avons commencé l'organisation des expériences concrètes avec la technologie du stylo numérique en collaboration avec les chercheurs concernés par chacun des quatre cas d'usage. Au cours de chaque expérience, nous avons produit un prototype de système de traitement et de visualisation répondant aux objectifs et besoins définis dans le contexte de l'expérience. Chaque prototype se base sur l'amélioration du prototype précédent.

Enfin, une phase d'analyse de chaque cas d'usage et du prototype utilisé permet de déterminer un ensemble de bonnes pratiques et un ensemble des fonctionnalités et des représentations qui caractérisent le système conçu. Cette analyse permet d'abstraire un modèle générique de système de traitement et de visualisation des données temps.

Le schéma suivant présente le fonctionnement de notre méthode qui passe par quatre phases à savoir :

- une phase d'instrumentation
- une phase d'exploration et d'analyse du besoin

- une phase de prototypage
- et une phase d'analyse générale.

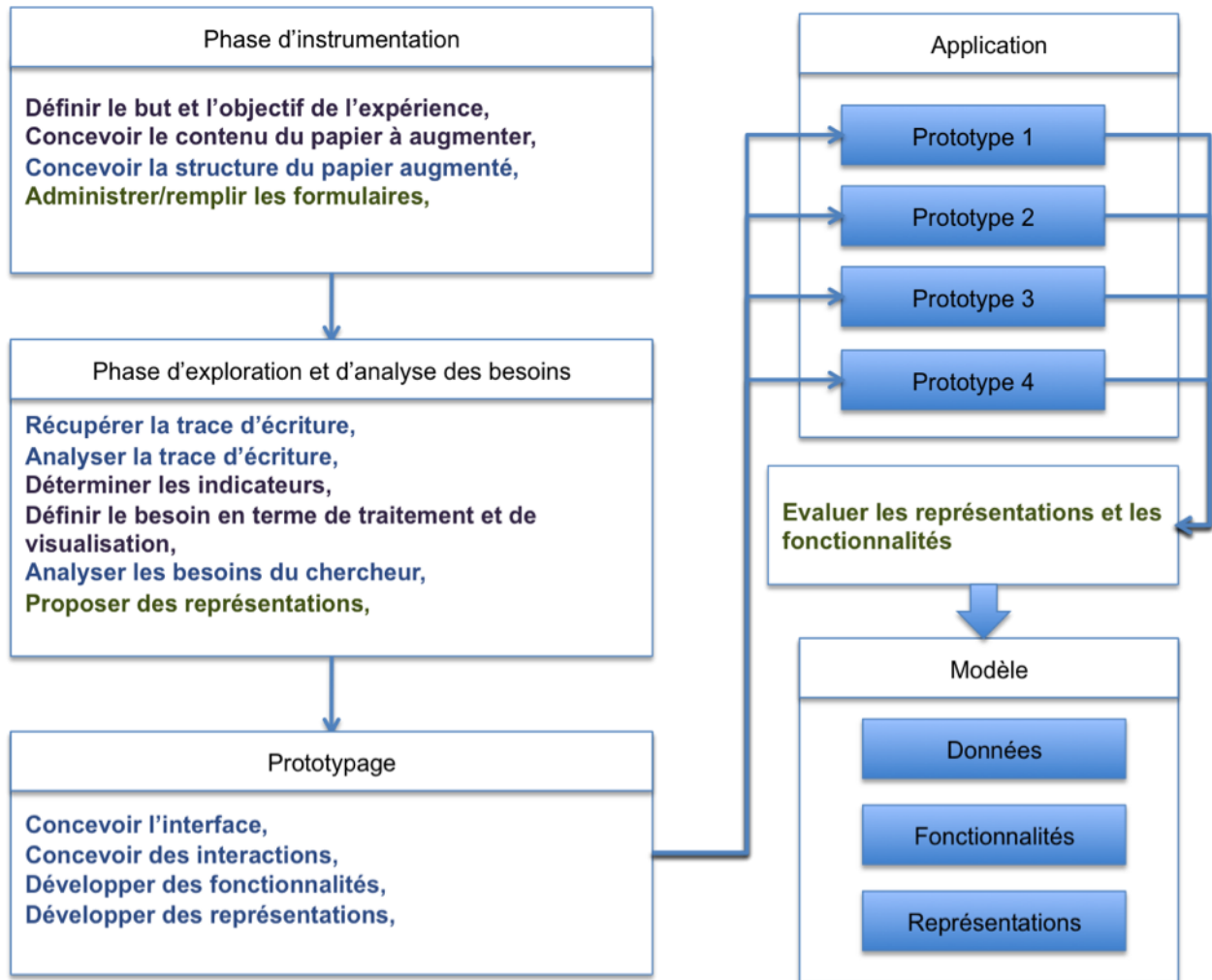


Figure 12 : Méthode de travail.

1.7. Organisation du mémoire

A la suite de ce chapitre d'introduction et de présentation des contextes, le mémoire est constitué comme suit : une première partie dans laquelle nous présentons les expériences effectuées, une deuxième partie dans laquelle nous décrivons une méthode générique ainsi que l'outil Segell, et enfin, un chapitre de conclusion et perspectives.

La première partie, dans laquelle nous présentons les expériences effectuées avec le stylo numérique, est composée de 3 chapitres. Dans chaque chapitre le contexte et l'objectif de

l'expérience sont exprimés. Ensuite, nous présentons les traitements et les représentations graphiques générées pour l'utilisateur. Ainsi, dans le chapitre deux, nous traitons le cas des tests d'évaluation ordonnés. Dans le chapitre trois nous présentons les tests d'évaluation non ordonnés. Et dans le chapitre quatre, nous présentons le cas des grilles de notes d'observation.

La deuxième partie est composée de deux chapitres (5 et 6). Dans le chapitre cinq, nous décrivons l'implémentation d'un prototype de l'outil SEGELL, notamment les spécifications fonctionnelles et techniques ainsi qu'une analyse d'utilisation de l'application. Le chapitre six décrit, d'une façon plus abstraite, un modèle générique de traitement et de visualisation de trace d'utilisation du stylo numérique. Ce modèle général sera établi à partir de la classification des fonctionnalités, des représentations et des interactions produites dans les prototypes.

Enfin, dans le chapitre sept, les conclusions et les perspectives du travail sont établies.

Chapitre 2. L'usage du stylo numérique pour un test d'évaluation ordonné

2.1. Introduction

Dans le cadre de notre étude exploratoire, nous avons participé à l'expérimentation d'usage de la technologie du stylo numérique pour la passation des tests d'évaluation de connaissances en tableur. Notre objectif est d'aider le chercheur en sciences de l'éducation à recueillir les données à l'aide de cette technologie et de proposer le traitement et la visualisation des données temps issues de la trace d'écriture des répondants.

A cette fin, nous avons organisé, en collaboration avec des collègues du laboratoire STEF, la passation de tests d'évaluation de connaissances sur le tableur auprès des étudiants de niveau licence.

Après l'analyse des besoins globaux de cette étude, nous avons produit un prototype permettant le traitement et la manipulation de différents types « d'indicateurs » à travers une interface interactive permettant de visualiser des représentations de ces données.

Enfin, nous avons établi une analyse de l'expérience basée sur l'évaluation des chercheurs et sur nos observations de la phase d'instrumentation ainsi que les fonctionnalités et les représentations offertes par ce prototype.

2.2. Analyse préalable

Dans ce cas d'usage du stylo numérique, l'utilisateur est un étudiant, ou bien un répondant, qui est dans une situation particulière de passation de test. Pour cela, nous avons précédé cette étude par cette une petite expérience pour mieux comprendre la situation de passation d'un test sur papier.

Cette expérience (voir annexe A) consiste à observer et filmer des étudiants au cours d'un test

de 15 minutes sur papier ordinaire. Nous avons demandé à 3 étudiants, de niveau licence 3, de faire ce test dans un cadre non académique (hors séance de cours). Le test était suivi d'un entretien semi directif avec les étudiants à propos de leur stratégie de réponse à un test.

L'analyse des observations et des entretiens avec ces étudiants nous permet de déterminer trois profils de répondants en terme de stratégie de réponse :

- (1) Lire la totalité du test puis commencer à répondre du début question par question et revenir sur les questions difficiles à la fin.
- (2) Lire la totalité du test puis commencer par les questions les plus faciles aux plus difficiles.
- (3) Lire et répondre au fur et à mesure question par question et revenir sur les questions difficiles à la fin.

Nous avons également effectué une analyse chronologique des enregistrements vidéo de ces étudiants au cours du test (voir annexe). Cette analyse révèle les constats suivants :

- Les tâches observables qui se manifestent sont notamment l'écriture et la lecture des questions. En outre, nous pouvons observé l'hésitation ou une réflexion après une écriture. Cependant, ces moments sont plus courts.
- Il existe un temps de lecture à considérer au début de chaque test qui précède l'instant de la première écriture de réponses.
- La tâche d'écriture est généralement précédée d'une tâche de lecture.
- La durée de lecture est plus importante que la durée d'écriture (l'écriture : allant de 1 seconde à 15 secondes ; la lecture : allant de 3 secondes à 2 minutes et 15 secondes).
- Il existe des temps de lecture passés sur des questions « difficiles » auxquelles l'étudiant n'a pas forcément répondu.

- Du chevauchement peut exister entre lecture et écriture des questions, par exemple, lecture d'une question, puis lecture et écriture dans une autre question, puis écriture de la première question.

Même si ces constats semblent être évidents, ils constituent une réflexion sur la façon de mener cette expérience avec le stylo numérique par rapport à la conception des tests et au traitement des données recueillies. Par ailleurs, quoique les données du stylo numérique auront une spécificité nécessitant une méthodologie autre que l'analyse des vidéos, les deux outils fournissent des données temps plus fines (instant de début et de fin d'écriture). Autant d'informations qui permettent d'envisager de répondre aux questions suivantes :

- Sur le processus de résolution des tests, quels sont les parcours individuels des répondants ?
- Y a-t-il des parcours particuliers ou des tendances de parcours ?
- Dans quel ordre chacun a-t-il répondu ?
- Quelle question a été traitée en premier ou en dernier ?
- Sur le temps passé sur chaque question, quelle est la durée d'écriture et la durée de lecture ou autres ?
- Il y a-t-il des relations entre les données temps (durée et ordre) et la difficulté des questions ?
- Il y a-t-il des relations entre les données temps (durée et ordre) et le résultat du test ?

La réponse à ces questions constitue, dans notre démarche, les besoins auxquels doit répondre notre solution à partir du traitement des données recueillies avec le stylo numérique.

2.3. Contexte et objectif de l'expérience

En éducation, nous avons remarqué (voir le chapitre 1) que l'usage de cette technologie a

concerné essentiellement le travail collaboratif (Maldonado et al., 2006) en classe et le partage des notes manuscrites (Miura et al., 2007), d'une part, et le tutorat et l'évaluation en classe (Lee et al., 2008; Oviatt et al., 2006), d'autre part.

Cependant, aucune étude, à notre connaissance, ne s'est intéressée à l'analyse de la trace d'écriture de l'élève enregistrée par le stylo numérique. Ces données contiennent, en plus de la note écrite, des données temps. L'analyse de ces données, comparable à l'analyse des traces numériques issues des environnements informatique d'apprentissage humain, permet d'observer et de construire des analyses intéressantes sur le comportement des élèves au cours d'un test d'évaluation.

Des travaux antérieurs réalisés au STEF, dans le cadre du projet DIDATAB¹³, avaient porté sur l'administration de tests pratiques sur le logiciel tableur, sur ordinateur. Ces travaux ont révélé que la façon dont les élèves procèdent lors de la résolution de tests sur le tableur est tout aussi intéressante que leur production. Afin d'étudier ces processus, des enregistrements vidéo des interactions des élèves avec le tableur ont été exploités (Tort, Blondel, & Bruillard, 2009). L'analyse qualitative de ces enregistrements révèle d'autres informations sur les processus de résolution des problèmes et les comportements des élèves au cours du test pratique, notamment les comportements consistant à explorer systématiquement l'interface du logiciel.

A la différence des tests sur ordinateur, les questionnaires papier présentent l'intérêt de pouvoir poser des questions sur les concepts et principes du logiciel, et de centrer les élèves sur ce qu'ils savent, sans qu'ils aient recours à une quelconque aide du logiciel. De plus, ils peuvent être administrés à une plus large population. Toutefois, ils ne donnent aucune indication sur le processus de réponse au questionnaire.

L'instrumentation de questionnaires sur papier tramé, en permettant l'enregistrement des temps d'écriture des réponses avec le stylo numérique, fournit un moyen de collecte des données sur le processus de résolution du test. Par cette étude, nous voulons exploiter ces

¹³ <http://www.stef.ens-cachan.fr/rech/didatab.htm>

données collectées dans le cas d'un test d'évaluation ordonné.

Pour avoir des cas concrets de tests d'évaluation, nous avons collaboré avec Françoise Tort, en tant que chercheur qui a participé au projet DIDATAB, d'une part, et en tant qu'enseignante à l'ENS Cachan qui fait des cours de formation sur le tableur destinés à des étudiants de L3 – économie et gestion, d'autre part.

De plus, Pascale Aoudé, qui était doctorante en Sciences de l'éducation à l'Université Paris Descartes, a été intéressée par l'utilisation de la technologie du papier et stylo numérique pour l'administration des tests d'évaluation dans le cadre de son projet de thèse intitulé « Les futurs enseignants du primaire face aux TIC: Questions de compétences et de formation : Le cas du tableur ».

Avec leurs participations, nous avons commencé par analyser les besoins méthodologiques du chercheur de la situation de collecte des données avec le stylo numérique jusqu'à l'exploitation et la visualisation des données recueillies. En outre, en tant qu'étude exploratoire, certains besoins ont été déclarés *a priori* par les chercheurs, alors que d'autres besoins, notamment ceux liés aux données temps, sont nés au fur et à mesure du traitement des données recueillies.

2.4. Organisation de l'expérience

Les tests d'évaluation de compétences en tableur ont été conçus dans le cadre méthodologique établi lors du projet DIDATAB visant à évaluer les acquis en connaissances sur le tableur des élèves de niveau collège et lycée. Une base de données de questions d'évaluation de compétences a été élaborée. Ces compétences étaient classés en cinq grandes catégories qui sont : Entités (Identification des entités, choix de leurs paramètres d'affichage et manipulation), Formules (Ecriture des formules), Graphiques (Traduction des données en graphiques), Modélisation (conception de tableaux de données pour résoudre un problème), et Tables (Gestion des données en tables).

Au cours de cette étude du cas d'usage de la technologie du papier et stylo numériques, nous avons instrumenté à l'aide de la technologie Anoto des tests conçus par les chercheurs et administrés auprès d'étudiants dans deux contextes concrets de tests d'évaluation en classe.

Le premier concerne trois tests destinés à des étudiants de Licence 3 en économie et gestion à l'ENS Cachan, et qui suivaient un cours de formation sur le tableur. Le deuxième contexte concerne deux tests destinés à des étudiants de niveau licence provenant de 3 institutions de formation en Ile-de-France (l'IUT de Rambouillet, l'université de Paris 5 et l'Université Paris 8-Saint Denis), et qui suivaient une formation en C2i.

La conception et l'instrumentation des formulaires étaient différentes pour chaque cas. Dans la suite, nous présentons la spécification de conception et d'instrumentation de chaque cas selon le contexte et le besoin du chercheur.

2.5. La conception et l'instrumentation des tests

Dans le cadre de notre étude, la même base de données du projet DIDATAB a été utilisée par les chercheurs pour la conception du contenu des tests ordonnés. Il s'agissait de tests comportant la description d'une situation initiale avec présentation de données sur une feuille de calcul. Puis, chaque test, disposé sur deux pages, proposait 10 questions ouvertes ou fermées successives mais indépendantes demandant de résoudre un problème simple correspondant à une compétence (écrire une formule, éditer un graphique, trier une table, etc.).

Concernant l'instrumentation des tests avec la technologie du papier et stylo numérique, nous avons conçu la structure des formulaires en projetant sur la lisibilité des données recueillies et dans la limite des exigences de la technologie. Pour cela, pour chaque formulaire, nous avons prévu une zone « identifiant » dans laquelle l'étudiant notera au début du test un identifiant composé de ses initiales (voir Figure 13). Cette valeur permet d'identifier les répondants au cours de la phase de traitement, mais aussi, le premier instant d'écriture dans cette zone représente le temps zéro à partir duquel nous allons calculer les données temps d'écriture dans les autres zones de réponses.

Identification :

QCM N°1
(Lundi 5 octobre 2009)

Le tableau ci-dessous présente une commande de produits. Sont détaillés sur chaque ligne de commande : le numéro de la ligne, le code du produit, le prix, la quantité commandée, le montant calculé, le montant net après remise. Le taux de la remise est saisi en haut à droite de la feuille.

	A	B	C	D	E	F
1						
2	Commande				Remise	10,00%
3						
4						

Figure 13 : Illustration de la zone « identifiant » depuis le premier test destiné aux étudiants de L3 Economie et Gestion de l'ENS Cachan.

Deux types de questions constituent le test à savoir ; les questions ouvertes et les questions fermées. Pour les questions ouvertes, qui concernent généralement l'écriture de formules, nous avons aménagé des zones de type texte (voir l'exemple Figure 14) pour que le répondant ait un espace suffisant à l'écriture de la formule attendue.

Question 4

La cellule E6 contient la formule =C6*D6. Si on la recopie ~~par curiosité~~ en C3, quelle formule obtiendrait-on en C3 ?

C3	=	
----	---	--

Figure 14 : un exemple de question ouverte depuis le premier test destiné aux étudiants de L3 Economie et Gestion de l'ENS Cachan.

Les questions fermées peuvent être soit à choix unique soit à choix multiple. La technologie d'instrumentation permet de reconnaître chaque type différemment. Les questions à choix unique sont représentées par l'objet bouton radio (voir Figure 15). Pendant la phase d'instrumentation, une valeur est attribuée à chaque bouton comme étant la réponse voulue. Le système de reconnaissance renvoie cette valeur une fois le répondant note une marque sur ce bouton avec le stylo.

Question 2

F2 affiche 10,00% avec deux décimales et le signe %. Pour obtenir cet affichage, il a suffi de sélectionner F2 puis :

- Saisir =10/100 tout simplement
- Saisir 0,1 puis cliquer sur le bouton % de la barre d'outil
- Saisir 0,1 puis choisir % et 2 décimales dans le menu Format/Cellule/Nombre
- Saisir 0,1 puis choisir % et 2 décimales dans le menu Affichage/Barres d'outils/Mise en forme

NB : une seule réponse possible

Figure 15 : exemple de question fermée à choix unique depuis le premier test destiné aux étudiants de L3 Economie et Gestion de l'ENS Cachan.

De la même façon, dans la phase d'instrumentation, une valeur réponse doit être attribuée à chaque choix pour une question fermée à choix multiple (voir Figure 16). Cependant, le système de reconnaissance va renvoyer chaque valeur où le répondant a noté une marque. Pour cela, nous allons avoir la valeur et les données temps de chaque réponse cochée.

Question 3

G16 contient une formule mais affiche #NOM? Que signifie cet affichage ? *NB : plusieurs réponses possibles*

- La formule comporte un nom de fonction mal orthographié.
- Un argument de type texte est passé à une fonction au lieu d'un numérique.
- La formule fait référence à une cellule qui a été supprimée
- La formule contient un nom de cellule qui n'existe pas
- Une plage de cellules est passée en argument d'une fonction au lieu d'une seule cellule

ENABLING *Amis* FUNCTIONALITY

Figure 16 : exemple de question fermée à choix multiple depuis le deuxième test destiné aux étudiants de L3 Economie et Gestion de l'ENS Cachan.

Concernant les tests destinés aux étudiants suivant une formation en C2i, un besoin spécifique au contexte de la recherche a nécessité d'ajouter devant chaque question une case pour manifester la sûreté du répondant de sa réponse. Si le répondant pense être sûr de sa réponse, il va mettre une marque dans cette case (voir Figure 17).

Nous avons représenté ce choix dans l'instrumentation par un groupe « sûreté » de cases à cocher, sous forme de question fermée à choix multiple où chaque case correspond à une question. Ainsi, le système de reconnaissance, de la technologie Kayentis, renvoie le numéro de la question et le temps d'écriture, si une case de sûreté a été cochée par le répondant.

Question 6 :

E7 affiche la somme des échecs sur les 4 années. Quelle formule de calcul contient-elle ?

E7	=
----	---

Question 7 :

B7 affiche #####. Que signifie cet affichage ?

- La cellule contient une formule erronée.
- Les données sont masquées pour la confidentialité.
- La valeur numérique contenue dans la cellule est trop grande pour la largeur de la cellule.

Figure 17 : illustration depuis le premier test de C2i des cases à cochés pour marquer la sûreté à propos des réponses aux questions.

A l'issue d'une phase de vérification et de validation du contenu et de l'instrumentation des tests avec les chercheurs, nous importons la structure du formulaire dans la plateforme de Kayentis et nous imprimons des exemplaires pour la préparation à la phase d'administration des tests.

2.5.1. L'administration des tests

Pendant la phase d'administration des tests, nous disposions d'une dizaine de stylos numériques disponibles. Pour cela, une partie des tests était imprimée sur papier ordinaire. Pour les chercheurs, ils ont recueilli les réponses de l'ensemble des étudiants, sans avoir les données temps de tous.

Pour les tests qui étaient destinés à des étudiants de L3 économie et gestion à l'ENS Cachan, chacun des tests portait sur des connaissances vues au cours des séances de TP sur le logiciel Excel. Trois tests ont été administrés au cours de l'année de formation (toutes les 4 séances de TP). 37 élèves ont effectué les tests, 9 d'entre eux ont utilisé les stylos numériques pour répondre. Chaque test a duré entre 10 et 15 minutes.

La note obtenue à ces tests était prise en compte dans l'évaluation finale des élèves pour l'obtention de leur diplôme, mais pour une part inférieure à l'examen final. La notation appliquée n'était pas à point négatif : 1 pour une bonne réponse, 0 pour une absence de réponse ou une mauvaise réponse. Les étudiants étaient donc plutôt incités à participer

correctement au test et à répondre aux questions posées.

Pour les tests passés à des formés en C2i, les répondants proviennent de 3 institutions : 41 % sont en première année de Gestion Administrative et Commerciale à l'IUT de Rambouillet, 45 % sont en L3 de sciences de l'éducation à Paris 5, et 14 % des répondants préparent une L1 en Sciences de l'Éducation à l'Université Paris 8- Saint Denis.

Deux tests ont été administrés auprès de ces étudiants. Le test 1 a été passé deux séances après le début de l'explication du module « tableur » et le test 2 a été passé deux séances plus tard. Il y a eu 58 répondants avec stylo numérique (sur 128 répondants) au Test 1 et 59 répondants avec stylo numérique (sur 98 répondants) au Test 2.

Une fois les tests administrés, une phase de vérification et de validation de la reconnaissance de l'écriture sur la plateforme de Kayentis était nécessaire. En effet, dans les questions ouvertes l'écriture des formules n'a pas été facile à valider par le système de reconnaissance de l'écriture. Même pour les questions fermées, il arrive que le répondant change d'avis sur un choix unique par exemple lorsqu'il rature avec le stylo une case et coche une autre, dans ce cas le système exige de ne valider qu'un seul choix. Pour cela, les chercheurs avaient comme tâches supplémentaires la vérification et la validation de certaines productions de leurs étudiants sur la plateforme de Kayentis.

Après la validation, nous effectuons l'exportation d'un fichier XML contenant les réponses et les données temps de tous les répondants correspondant à chaque test. C'est ce fichier qui sera traité par notre application afin de fournir les informations souhaitables en format lisible (sous forme de tableau xls ou bien en graphique) pour le chercheur.

2.6. Le traitement des données

Pour ce cas d'usage du stylo numérique, nous avons voulu approfondir le traitement des données recueillies afin de pouvoir mettre en évidence l'apport de l'analyse des données temps par rapport aux autres données relatifs à un test d'évaluation. De plus, ces traitements permettent d'analyser le parcours d'écriture des répondants à un test.

Notre démarche consiste à exploiter ces données temps pour en fournir des « indicateurs

temps » permettant au chercheur d'enrichir ses analyses.

D'autre part, le chercheur, dans le cadre de son étude, a besoin de définir un certain nombre « d'indicateurs » qui sont en relation avec les données relatives aux questions et aux réponses aux questions. Ce sont des données qui qualifieront les réponses obtenues aux tests, et permettront une analyse qualitative des réponses.

Nous distinguons donc deux types de traitement : le calcul automatique des « indicateurs temps », relatifs à la trace d'écriture avec le stylo numérique en général, et les « indicateurs » personnalisés relatifs au contexte d'une étude, et que le chercheur peut ajouter à son corpus de données à partir d'une interface utilisateur.

2.6.1. *Les indicateurs temps*

En tenant compte de l'expérience décrite dans le paragraphe 2 plus haut (Analyse préalable), nous rappelons que le processus de résolution d'un test, constitué essentiellement de deux tâches qui sont l'écriture et la lecture sur des durées variées, est différent pour chaque répondant. Nous voulons savoir l'apport de l'analyse de ces processus et de ces données temps dans le domaine d'évaluation en éducation pour un test ordonné. Pour cela, nous avons exploré quelques recherches, qui s'intéressent à l'analyse des données temps, notamment dans l'analyse des questionnaires et en psychologie cognitive.

2.6.2. *Exploration théorique*

Des travaux portant sur l'analyse des questionnaires d'enquêtes par téléphone ont étudié le facteur temps de réponse, appelé « latence », pour comprendre les comportements et les attitudes politiques des répondants (Bassili & Fletcher, 1991). Plus récemment, (Mulligan, Grant, Mockabee, & Monson, 2003) ont couplé les enquêtes par téléphone à une technique informatisée de recueil des temps. L'analyse de ces temps de latence a été utilisée pour tester l'efficacité et la bonne formulation des questions (Bassili & Scott, 1996). Elle a été exploitée comme un indicateur d'erreurs pour (Draisma & Dijkstra, 2004). Enfin, dans l'analyse d'une enquête web effectuée auprès de demandeurs d'emploi d'une part et d'employés d'autre part, (Callegaro, Yang, Bholá, & Dillman, 2004) ont utilisé la mesure de la latence comme indicateur de la motivation du répondant. La latence est donc considérée comme un indicateur

important pour connaître l'efficacité et la difficulté des questions, mais aussi l'exactitude des réponses, l'attitude et la motivation du répondant.

Des travaux, en psychologie cognitive, s'intéressant principalement à la tâche d'écriture, ont étudié particulièrement les durées des pauses (Foulin, 1995). Ils ont montré que les durées de pauses étaient des indicateurs d'un traitement cognitif. Leur durée est révélatrice d'une tâche cognitive riche : recherche des savoirs, structuration syntaxique et grammaticale de la phrase, conception de la phrase (Olive, Alves, & Castro, 2008).

Dans le cas de questionnaires sur le web ou sur papier où le répondant doit écrire lui-même la réponse, la mesure de la latence commence à partir du moment de lecture de la question. Le processus cognitif du répondant comprend la formulation syntaxique et logique de la réponse. Pour cela, le temps de pause avant l'écriture d'une réponse fait partie de la durée totale de traitement d'une question.

Nous souhaitons instrumenter le calcul de durée d'écriture et de pause pour aider à analyser finement le comportement d'un répondant à un test sur papier dans le cadre de notre étude. Ainsi, la durée d'écriture, qui correspond à la tâche effective durant laquelle le répondant donne la réponse brute, et la durée de pause, qui correspond à une tâche cognitive durant laquelle le répondant lit et réfléchit sur la réponse, ces durées constituent la durée de traitement de la tâche.

2.6.3. Analyse de la trace

En pratique, le stylo numérique enregistre tous les instants d'écriture d'un tracé dans chaque zone. Cependant, le fichier de trace XML fourni par la société Kayentis, ne restitue que les instants de début d'écriture (appelés FirstStroke dans le fichier XML) et les instants de fin d'écriture (appelés LastStroke dans le fichier XML), de chaque question pour chaque répondant. Nous n'avons pas le détail de ce qui s'est passé entre ces deux moments. Nous avons donc utilisé ces données pour calculer les durées d'écriture des réponses et les durées de non écriture.

Rappelons que dans une situation de test sur papier, sans observation directe et sans enregistrement vidéo du répondant, nous ne pouvons pas avoir exactement les moments de

lecture d'une question. Notamment lorsque le répondant fait du chevauchement au niveau de lecture et écriture entre les questions. Ainsi, nous supposant que le temps qui précède l'instant d'écriture d'une réponse à une question c'est un temps de pause, ou de non écriture, qui contient un temps de lecture ou de réflexion sur, le plus probable, la question répondu, ou bien sur une autre question.

Dans le cas le plus simple (voir Figure 18 -a), où le répondant écrit dans chaque champ de réponse successivement et ne revient pas les modifier, la suite chronologique des temps collectés forme une suite des couples de FirstStroke et LastStroke de chaque champ de réponse. Ainsi, si le temps d'écriture pour une question (k) va correspondre à l'intervalle entre son « firststroke » et son « laststroke », le temps de latence pour la question (k) correspond à l'intervalle entre le « laststroke » d'une question (j) et le « firststroke » de la question (k).

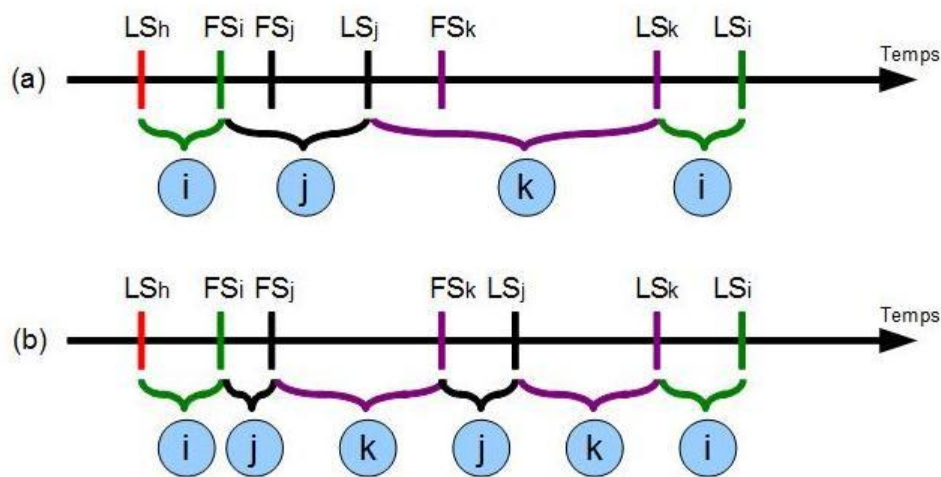


Figure 18 : La succession des FS et des LS dans le temps.

Dans le cas du chevauchement de lecture et d'écriture entre les questions, les instants de temps concernant une question ne se suivent plus dans la chronologie des temps collectés (voir Figure 18 -b). Dans ce cas, un retraitement des temps collectés est nécessaire pour déterminer les temps passés à écrire chaque réponse ainsi que les temps de latence passé ou de non écriture.

Pour toutes les situations, nous avons choisi d'attribuer la durée passée entre un premier instant correspondant à un champ de réponse (i-1) et un autre qui le suit correspondant à un champ de réponse (i) à ce dernier champ de réponse (i).

$$D_i = T_i - T_{i-1}$$

Si T_{i-1} correspond à un LastStroke, dans ce cas nous supposons que D_i est une durée d'écriture De_i . Sinon, nous supposons que D_i est une durée de « non écriture » (ou latence) Dne_i .

Nous recueillons ainsi pour chaque champ de réponse deux valeurs de durées (De_i et Dne_i). La durée totale consacrée à ce champ de réponse (i) est donc la somme des deux durées ainsi estimées.

$$DT_i = Dne_i + De_i$$

Compte tenu de ces calculs proposés, ces durées demeurent des approximations.

2.6.4. Les données traitées

Au moment de l'importation de la trace, l'application permet d'automatiser le traitement de certains paramètres, notamment relatifs aux données temps comme le moment de début de l'écriture, le moment de fin de l'écriture et les durées calculées comme décrit précédemment.

Pour le rang de traitement des questions, nous avons attribué par défaut le rang de traitement d'une question au rang du premier instant d'écriture dans le champ de réponse de la question.

Ces traitements ont constitué les transformations automatisées pour avoir un ensemble « d'indicateurs » pour chaque champ de réponse et chaque répondant :

- Le début d'écriture (FirstStroke)
- La fin de la dernière écriture (LastStroke)
- La durée d'écriture
- La durée de non écriture – ou latence
- La durée totale (écriture + latence)

- Le rang (dans l'ordre du traitement du test)
- Le type du champ de réponse (bouton radio, case à cocher, zone de texte)

Une autre donnée classique que le chercheur avait besoin de considérer dans son analyse est la notation de la réponse. Cette donnée pourrait être calculé pour les questions fermées mais pour les questions ouvertes son calcul dépend de la reconnaissance de l'écriture. Pour cela, le chercheur doit saisir pour chaque répondant et chaque réponse la note comme suit :

- La notation de la réponse (1 si c'est une bonne réponse et 0 si c'est une mauvaise réponse).

Ces traitements, déterminés en collaboration avec le chercheur, seront les mêmes pour tout test ou questionnaire instrumenté. Cependant, le chercheur a besoin d'autres « indicateurs » qui correspondent au contexte de sa recherche.

2.6.5. Les « indicateurs » personnalisés

Contrairement aux traitements automatisés que notre application permet de faire à l'importation de la trace, les traitements personnalisés nécessitent une interface utilisateur à travers laquelle le chercheur peut effectuer des manipulations.

Nous commençons par présenter les « indicateurs » personnalisés qui correspondent au contexte de l'expérience de l'usage du stylo numérique pour les tests d'évaluation ordonnés. Ensuite, nous présentons les fonctionnalités proposées au niveau de l'interface de notre application qui permettent de définir ces « indicateurs ».

Pour le cas des tests d'évaluation, le chercheur a besoin de spécifier trois « indicateurs » propres à ses analyses. Le premier est un paramètre lié à l'origine de l'expérience à savoir le projet DIDATAB. Appelé « compétence », cet indicateur spécifie la catégorie de compétence, que le chercheur veut évaluer, à laquelle correspond une question. Rappelons que chaque question du test relevait de l'une de ces familles de compétences.

De plus, pour répondre au questionnement à propos de la relation entre les données temps

(durée et ordre) et la difficulté des questions, le chercheur a défini l'indicateur « difficulté ». Les questions sont classées a posteriori en trois niveaux de difficulté d'après le taux de réussite des répondants (« difficile » : moins de 1/3 du nombre totale des répondants, « moyenne »: entre 1/3 et 2/3, « difficile » : plus que de 2/3 des répondants).

Finalement, un « indicateur » qui correspond au type des questions a été défini dans le but de pouvoir analyser les durées de réponses par rapport au type de la question. Pour cet « indicateur » « type de question » nous avons défini deux catégories : « fermée » pour les questions fermées et « ouverte » pour les questions ouvertes.

Ainsi, au sein d'une étude exploratoire, le chercheur commence par se fixer des besoins par rapport à l'ensemble de traitement qu'il a prévu faire. Après, d'autres besoins vont apparaître au fur et à mesure. En effet, le chercheur ne se contente pas des traitements automatisés mais il a besoin d'apporter d'autres manipulations des données. Dans ce cas, deux solutions sont envisageables ; soit de fournir au chercheur les données sous forme exploitable (en format csv ou xls par exemple) pour qu'il puisse faire des manipulations sur d'autres logiciels, soit de fournir une interface souple et paramétrable permettant d'apporter les manipulations souhaitées.

Dans notre application, en plus de la possibilité d'exporter les données traitées automatiquement en format xls, nous avons proposé une interface utilisateur qui permet au chercheur de définir des « indicateurs » personnalisés. A travers cette interface, le chercheur peut définir un nouvel « indicateur » et déclarer les différentes catégories possibles. Ensuite, il peut affecter les catégories aux champs de réponse constituant le test.

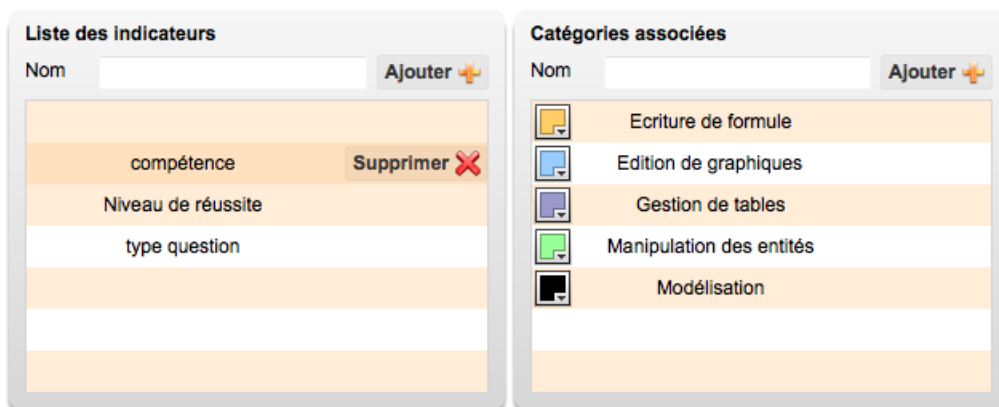


Figure 19 : Ecran de gestion des indicateurs illustrant l'indicateur « compétence » et les catégories y associées pour le premier test tableur – L3 économie et gestion de l'ENS Cachan.

Par exemple, pour le premier test de compétences sur le tableur administré auprès des étudiants de L3 économie et gestion de l'ENS Cachan, (voir Figure 19) le chercheur a défini des catégories « compétence » suivantes : « manipulation des entités », « écriture de formule », « édition de graphique », « gestion de tables », etc.). Cet « indicateur » va être considéré par la suite dans la phase de génération de représentation tabulaire et graphique et sera confronté aux données temps calculées.

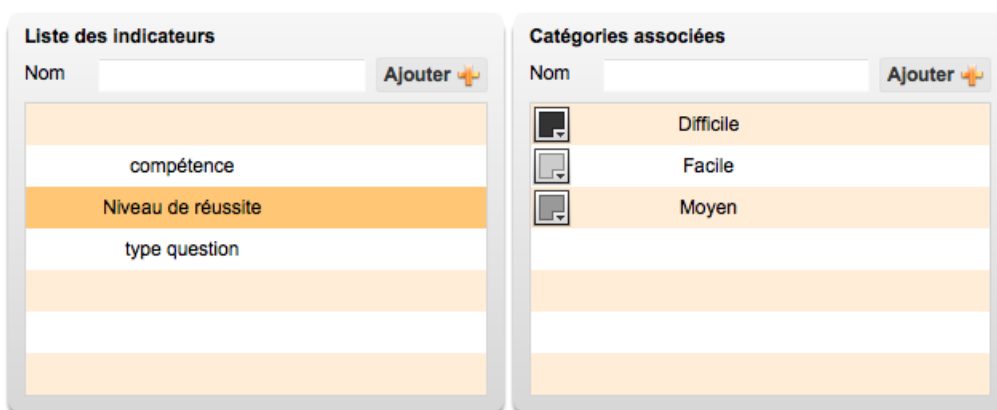


Figure 20 : Ecran de gestion des indicateurs illustrant l'indicateur « difficulté » et les catégories y associées pour le premier test tableur – L3 économie et gestion de l'ENS Cachan.

La Figure 20 illustre les catégories qui correspondent à « l'indicateur de réussite », et la Figure 21 présente les catégories « type de question ».

Figure 21 : Ecran de gestion des indicateurs illustrant l'indicateur « type de question » et les catégories y associées pour le premier test tableur – L3 économie et gestion de l'ENS Cachan.

Après avoir défini les indicateurs spécifiques au contexte de l'analyse dans l'application, arrive la phase d'affectation des champs de réponses à ces « indicateurs » définis. Dans l'illustration suivante (voir Figure 22) nous présentons un extrait de la classification des champs de réponses par « indicateurs personnalisés » utilisés dans le premier test tableur administré auprès d'étudiants L3 économie et gestion de l'ENS Cachan.

Nom du champ	Nom de l'indicateur	Nom de la categorie	Supprimer
question1	compétence	Manipulation des entités	Supprimer ✖
question1	Niveau de réussite	Moyen	Supprimer ✖
question1	type question	fermée	Supprimer ✖
question2	compétence	Manipulation des entités	Supprimer ✖
question2	Niveau de réussite	Facile	Supprimer ✖
question2	type question	fermée	Supprimer ✖

Filtrer par:
 Champ: Tous | Indicateur: Tous | Categorie: Toutes

Figure 22 : Ecran de classification des champs de réponses par indicateurs pour le premier test tableur – L3 économie et gestion de l'ENS Cachan.

Enfin, à la suite d'une démonstration de l'interface au chercheur, un nouveau besoin est né concerne la possibilité de personnaliser la nomenclature des champs de réponses ainsi que la possibilité de personnaliser la valeur « ordre » d'apparition des champs dans les représentations.

Donner un libelle ou un ordre à un champ de formulaire

Champ	Libelle	Ordre
Q3F6	Question 3	3
question1	Question 1	1
question2	Question 2	2
Q4C3	Question 4	4
Q6E11	Question 6	6
Q8A12	Question 8	8
Q10F6	Question 10	10
question5	Question 5	5
question7	Question 7	7
question9	Question 9	9


Enregistrer 

Figure 23 : Ecran de personnalisation de la valeur « libellé » de champ de réponse et la valeur « ordre » pour le premier test tableur – L3 économie et gestion de l'ENS Cachan.

Pour cela, nous avons ajouté dans l'interface une nouvelle fonctionnalité de gestion des champs. Ces manipulations permettent essentiellement de personnaliser la présentation des données (voir Figure 23).

2.7. La visualisation des données

La phase de représentation des données consiste à concevoir une interface qui permet au chercheur de visualiser l'ensemble de ses données avec des fonctionnalités permettant de filtrer et d'éditer les données temps calculées, en fonction des « indicateurs » définis, sous forme de représentations graphiques ou tabulaires.

Dans le contexte d'un test d'évaluation, rappelons que l'objectif de cette analyse est de pouvoir visualiser le processus de résolution d'un test par un répondant, d'une part, et par l'ensemble des répondants, d'autre part, afin de vérifier l'existence de stratégies générales ou bien particulières de résolution. L'analyse consiste donc à observer la production et les données qui correspondent à chaque individu et à l'ensemble des répondants.

Pour cela, nous avons distingué deux types de représentations ; individuelles et collectives. Pour chaque type nous avons proposé des représentations tabulaires et des représentations graphiques. Le chercheur a la possibilité d'exporter ces représentations sous ces deux formes.

La conception de ces représentations repose principalement sur les besoins exprimés par les

chercheurs. En effet, nous avons commencé par exporter des tableaux de données en format xls. Ces fichiers contiennent les données recueillies avec le stylo numérique (voir Figure 24). A partir de ces données, le chercheur a pu déterminer un ensemble de besoins en terme de traitement et de représentation.

Nous avons donc analysé ces besoins et ces représentations pour les transformer en fonctionnalités permettant de générer à travers une interface des représentations pertinentes pour le chercheur. Dans cette interface nous avons apporté des interactions.

	A	B	C	D	E
1	InstanceID	Name	Value	LastStroke	FirstStroke
2	96.529.1.670	identifiant	THGA	05/10/09 13:08	05/10/09 13:08
3	96.529.1.670	Q3F6	(1-\$FS2)*E6	05/10/09 13:18	05/10/09 13:11
4	96.529.1.670	question1	saisirunaun	05/10/09 13:10	05/10/09 13:10
5	96.529.1.670	question2	format/cellule/nbre	05/10/09 13:10	05/10/09 13:10
6	96.529.1.670	Q10F6	IF(E6>35000;(1-\$FS2*E6;E6)	05/10/09 13:18	05/10/09 13:16
7	96.529.1.670	Q4C3	36474,75	05/10/09 13:13	05/10/09 13:12
8	96.529.1.670	Q6E11	SOMME(E6:E10)	05/10/09 13:14	05/10/09 13:14
9	96.529.1.670	Q8A12	"le montant à payer est." & F11	05/10/09 13:16	05/10/09 13:15
10	96.529.1.670	question5	largeurcellule	05/10/09 13:13	05/10/09 13:13
11	96.529.1.670	question7	texte	05/10/09 13:15	05/10/09 13:15
12	96.529.1.670	question9	1fois	05/10/09 13:16	05/10/09 13:16
13	96.529.1.660	identifiant	SEBE	05/10/09 15:20	05/10/09 15:20
14	96.529.1.660	Q3F6	SE6x(-\$FS2+1)	05/10/09 15:23	05/10/09 15:22
15	96.529.1.660	question1	selectA6A7copie	05/10/09 15:21	05/10/09 15:21
16	96.529.1.660	question2	format/cellule/nbre	05/10/09 15:21	05/10/09 15:21
17	96.529.1.660	Q10F6	si (sup (35000));SE6x(1-\$FS2))	05/10/09 15:30	05/10/09 15:29
18	96.529.1.660	Q4C3	A3xB3	05/10/09 15:24	05/10/09 15:24
19	96.529.1.660	Q6E11	somme(ES6:ES10)	05/10/09 15:25	05/10/09 15:25
20	96.529.1.660	Q8A12	ECRI("le montant à payer est: F11")	05/10/09 15:31	05/10/09 15:30
21	96.529.1.660	question5	largeurcellule	05/10/09 15:24	05/10/09 15:24
22	96.529.1.660	question7	decimal	05/10/09 15:27	05/10/09 15:27
23	96.529.1.660	question9	1fois	05/10/09 15:26	05/10/09 15:26

Figure 24 : Un extrait d'un fichier xls contenant les données du stylo numérique de deux répondants au premier test tableau – L3 économie et gestion de l'ENS Cachan.

2.7.1. Les représentations individuelles

Pour visualiser les données correspondant à chaque répondant, nous avons deux formes classiques de représentation : une représentation tabulaire et une autre graphique.

Concernant la représentation tabulaire des données, à partir des besoins, nous avons conçu l'interface d'une représentation d'un tableau formaté des données de production d'un répondant à un test.

Le chercheur choisit un test et un répondant puis clique sur un bouton pour obtenir une fenêtre contenant un tableau avec les données suivantes : la liste des questions du test, la valeur réponse écrite par le répondant et capturé par le stylo numérique, l'instant de début d'écriture d'une réponse, l'instant de fin d'écriture d'une réponse, la durée totale passé sur une question,

la durée d'écriture de la réponse, la durée de non écriture et la note obtenue (voir Figure 25).

Tableau réponses par individu							
Répondant: BADU				Formulaire: QCM1 tableur			
Champ	Réponse	Début d'écriture	Fin d'écriture	Durée totale	Durée E	Durée NE	Note
Question 1	formuleA6plus1	2009-10-05 15:21:46	2009-10-05 15:21:48	00:00:02	00:00:02	00:00:00	0
Question 2	format/cellule/nbre	2009-10-05 15:22:27	2009-10-05 15:22:29	00:00:02	00:00:02	00:00:39	1
Question 3	E6x(1-F\$2)	2009-10-05 15:23:36	2009-10-05 15:24:30	00:02:01	00:00:54	00:01:07	1
Question 4	A3xB3	2009-10-05 15:25:32	2009-10-05 15:25:51	00:01:21	00:00:19	00:01:02	1
Question 5	largeurcellule	2009-10-05 15:26:26	2009-10-05 15:26:27	00:00:01	00:00:01	00:00:35	1
Question 6	SOMME(E6:E10)	2009-10-05 15:27:28	2009-10-05 15:27:52	00:01:25	00:00:24	00:01:01	1
Question 7	fraction	2009-10-05 15:29:09	2009-10-05 15:29:10	00:00:01	00:00:01	00:01:17	0
Question 8	"le montant À payer est:"8	2009-10-05 15:29:47	2009-10-05 15:33:06	00:01:47	00:01:10	00:00:37	1
Question 9	1fois	2009-10-05 15:30:36	2009-10-05 15:30:41	00:00:05	00:00:05	00:00:49	1
Question 10	SI(E6>35000;E6x(1-F\$2);E	2009-10-05 15:31:00	2009-10-05 15:31:56	00:01:15	00:00:56	00:00:19	1

Figure 25 : Tableau de réponse pour chaque individu extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.

À partir de cette interface, le chercheur peut affecter les notes (1 si c'est une bonne réponse et 0 si c'est une mauvaise réponse) pour chaque réponse. Il est possible également d'exporter ce tableau en format xls pour pouvoir le traiter sur un autre logiciel tableur.

Concernant la représentation graphique, nous avons commencé par la génération d'un graphique chronologique illustrant le parcours de réponse du répondant. Pour chaque champ de réponse, numéroté de 1 à 10 dans la figure en bas (voir Figure 26), sont affichés les instants de début et de fin d'écriture. Ce qui permet de distinguer l'ordre d'écriture dans les champs de réponse.

Et pour illustrer la durée de non écriture pour un champ de réponse, nous avons représenté l'instant de fin d'écriture du champ précédent comme étant le début du temps de latence pour ce champ. Ainsi, pour chaque champ, 3 temps lui sont affectés à savoir le LastStroke du champ précédent comme étant l'instant de début de la durée de non écriture, le FirstStroke du champ lui-même comme étant l'instant de début de la durée d'écriture et le LastStroke du champ lui-même comme étant l'instant de la fin de la durée d'écriture.

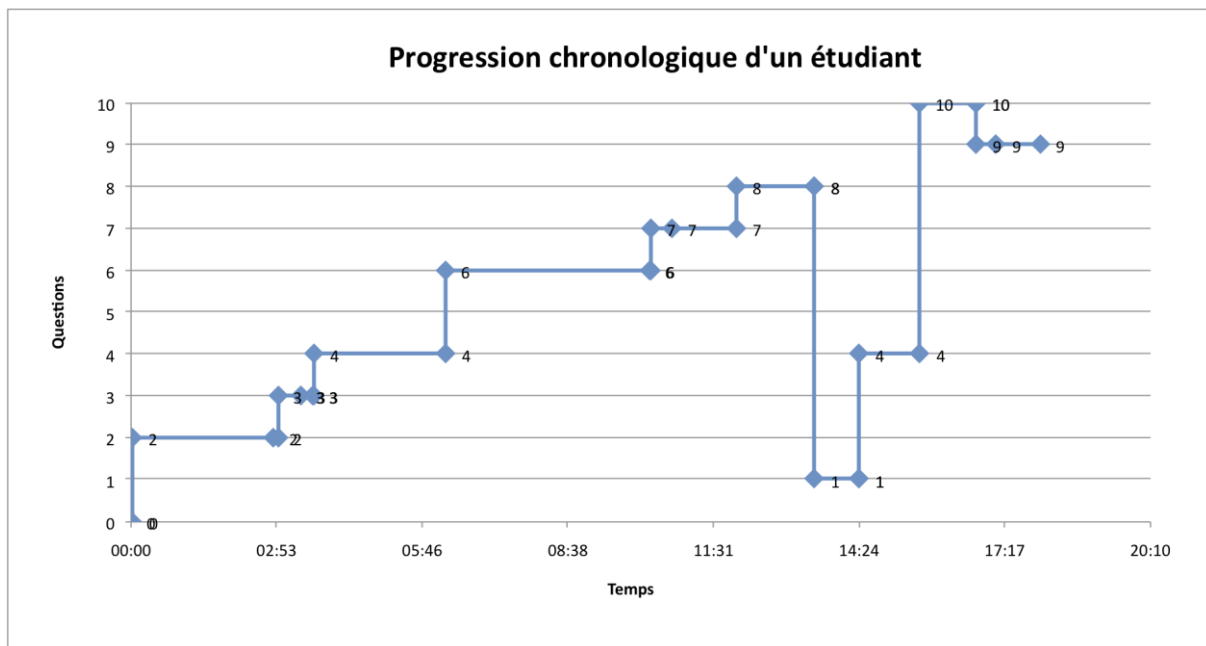


Figure 26 : Conception du graphique chronologique du parcours de réponse d'un répondant. Les numéros des questions qui sont en ordonnée permettent de distinguer l'ordre de traitement du test par le répondant.

Le graphique illustre donc, en plus des instants de début et de fin d'écriture dans un champ de réponse, la durée totale de réponse et l'ordre de traitement des questions. Cependant, le chercheur avait besoin de savoir pour quelles questions le répondant a réussi ou pas. Ce qui lui permet de comparer la réussite d'une question par rapport à sa durée ou son rang de traitement. De plus, le chercheur voulait comparer la durée passée sur une question par rapport à la durée moyenne passée par tous les répondants sur cette question. Autant d'informations que le chercheur a besoin de visualiser sur une même représentation pour pouvoir les comparer et les lier les unes aux autres.

Pour cela, nous avons proposé la représentation personnalisée illustrée dans la Figure 27 sous forme de graphique multiple dans lequel nous regroupons plusieurs dimensions et paramètres à visualiser (Barhoumi, 2010). En effet, nous avons proposé de représenter le résultat de la production par un signe coloré (si réussi en bleu sinon en rouge) sur le même graphique de progression chronologique. Nous avons proposé de colorer le fond du graphique pour chaque question selon son niveau de difficulté (difficile, moyenne, facile). Ainsi, cet indicateur permet de comparer le niveau du répondant par rapport à l'ensemble des répondants.

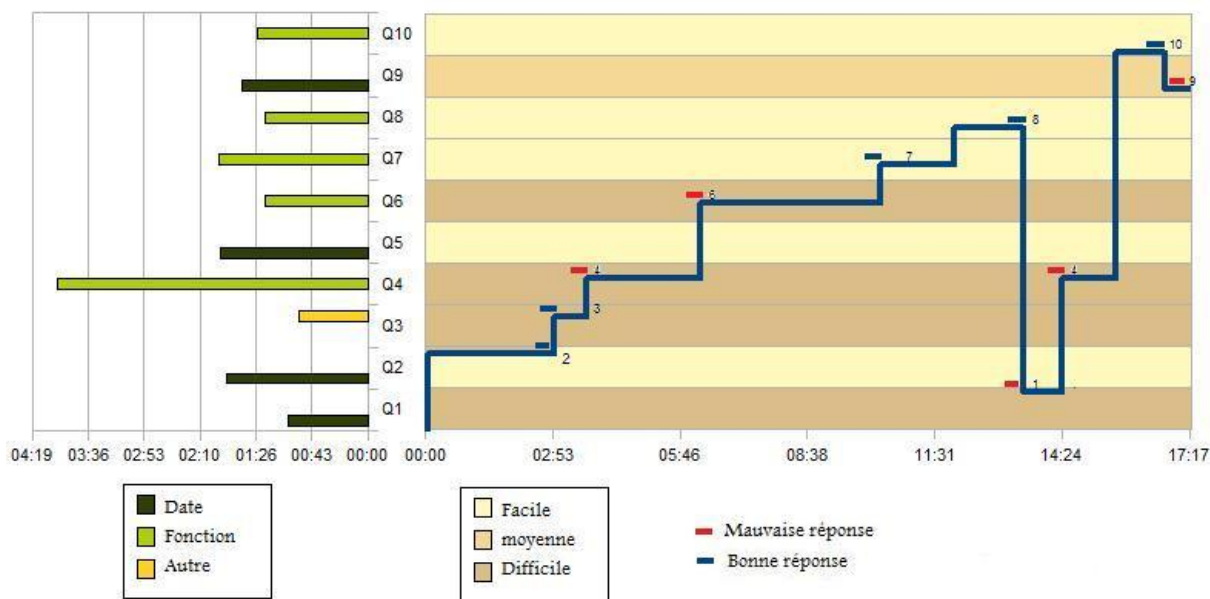


Figure 27 : Représentation multiple illustrant : (1) la progression chronologique d'un répondant, (2) le résultat de sa production, (3) le niveau de difficulté des questions, (4) la classification des questions selon les compétences, (5) le temps moyen passé sur chaque question.

La représentation de la progression chronologique est confrontée à un graphique en barre, à gauche, qui représente le temps moyen passé sur chaque question par tous les répondants. Les séries illustrées dans cette représentation constituent les catégories de compétences.

L'intérêt de cette représentation est de rendre visibles plusieurs paramètres en superposant des valeurs individuelles (progression d'un étudiant, résultat d'une question) sur des valeurs d'ensemble (la durée moyenne, le niveau de difficulté). Cependant, il s'avère aussi intéressant de rendre ce graphique dynamique et interactif.

Nous avons donc produit une interface utilisateur interactive permettant d'illustrer ce graphique (voir Figure 28). La représentation de la progression chronologique ainsi que le graphique en barre de la durée moyenne étant fixe, le chercheur peut modifier « l'indicateur » qui va être illustré à travers les couleurs du fond du graphique chronologique à droite et les couleurs des barres du graphique à gauche. Ces couleurs sont personnalisables à travers l'interface de gestion des « indicateurs » et des catégories. Le chercheur peut exporter chaque graphique en format jpg pour pouvoir l'importer dans d'autres logiciels.

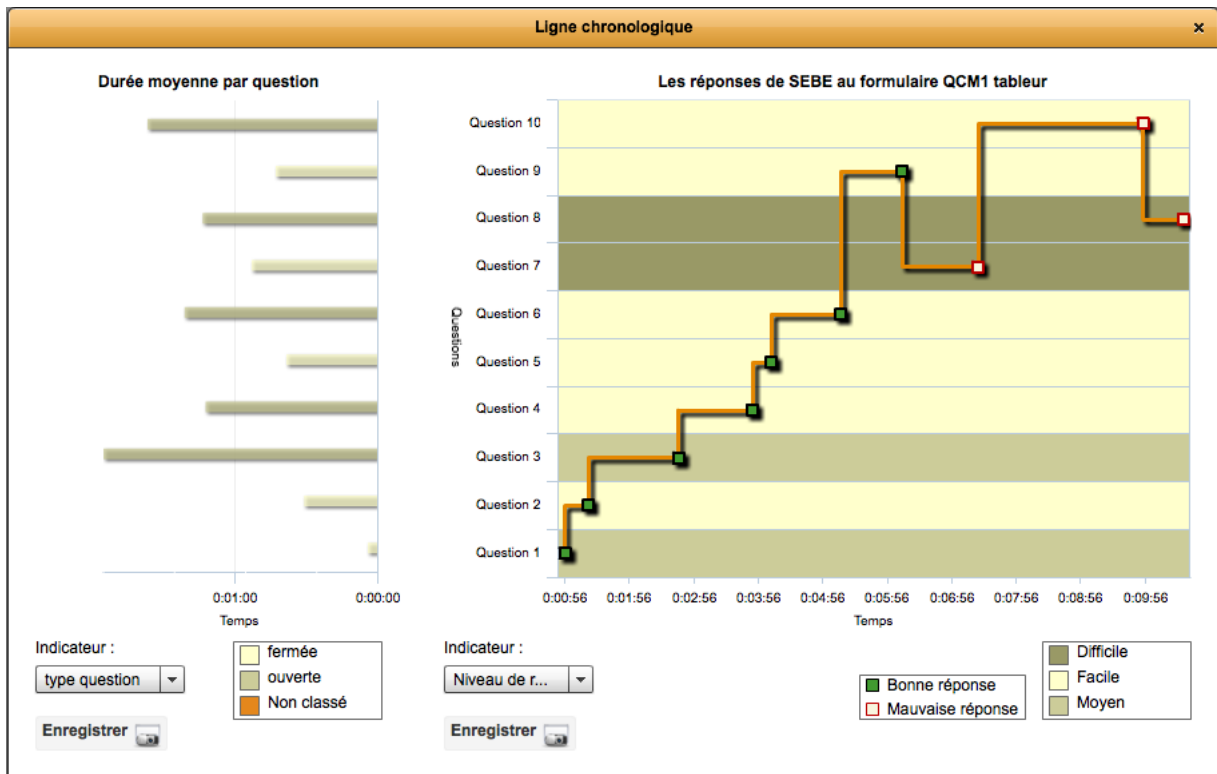


Figure 28 : Interface interactive qui illustre la représentation personnalisée du graphique multiple.

En somme, ce graphique se base sur la représentation des données temps qui sont les instants de début et de fin d'écriture, l'ordre de résolution des questions, la durée de traitement d'une question par un répondant, la durée moyenne de traitement d'une question par tous les répondants. Ce sont les données temps que nous voulons traiter à travers l'exploitation de la trace d'écriture avec un stylo numérique. Par ailleurs, cette représentation est enrichie par les « indicateurs personnalisés » par le chercheur qui sont le niveau de difficulté d'une question, le type d'une question et les catégories de compétences en tableur, mais qui peuvent changer selon le contexte de l'étude à analyser.

2.7.2. Les représentations collectives

La représentation des données de l'ensemble des répondants permet au chercheur d'avoir un aperçu global sur les tendances de stratégies de résolutions des tests et les processus particuliers. Comme pour les représentations individuelles, nous avons proposé une représentation tabulaire et des représentations graphiques.

	A	B	C	D	E	F	G
1							
2	Les instances par zone et par temps de LS du test eneide_tableur_test2C2I_v1						
3							
4	Q°	PATDT	BALCHA	ROXDAR	MAMSOP	RANJUL	BARSIP
5	Question1	00:00:23	00:02:06	00:00:28	00:00:10	00:00:16	00:00:23
6	Question10	00:03:19		00:06:07	00:07:46	00:09:34	00:05:37
7	Question2	00:00:39	00:03:45	00:01:28	00:01:18	00:11:42	00:02:28
8	Question3	00:00:52	00:02:34	00:01:47	00:00:14	00:02:47	00:00:27
9	Question4	00:02:03	00:06:35	00:07:06	00:04:42	00:11:29	00:07:12
10	Question5	00:02:10	00:07:11	00:03:58	00:03:18	00:05:26	00:01:25
11	Question6	00:02:24	00:10:23		00:08:59	00:10:05	
12	Question7	00:02:30	00:10:37	00:04:41	00:05:17	00:07:06	00:03:17
13	Question8	00:02:38	00:11:52	00:05:18	00:06:37	00:07:35	00:04:05
14	Question9	00:02:42	00:12:15	00:05:29	00:07:14	00:08:17	00:04:26
15	sure1	00:00:46		00:06:23	00:08:09	00:00:15	
16	sure2	00:03:20		00:06:34	00:05:18	00:08:19	
17							

Figure 29 : un extrait de fichier xls contenant les instants de fin d'écriture correspondant à tous les questions pour chaque répondant au deuxième test de C2i.

Pour la représentation tabulaire, nous avons produit des tableaux d'export xls selon le besoin du chercheur. En effet, il a eu besoin d'avoir des tableaux des données temps de tous les répondants pour chaque question. Les tableaux produits à cette étape contiennent des données qui correspondent à un indicateur temps (début et fin d'écriture, durée d'écriture, latence) pour tous les répondants à un test (voir Figure 29).

Pour cela, nous avons conçu une interface paramétrable qui permet de choisir un « indicateur » temps et de générer un tableau exportable en format xls. D'autres options permettent de filtrer ces données temps par rapport à un « indicateur personnalisé » (voir Figure 30).

Tableau récapitulatif

Formulaire: QCM1 tableur

Choisir une donnée à afficher

Début écriture

Fin d'écriture

Durée totale

Durée d'écriture

Durée non écriture

Note

Rang

Filter les champs

Par indicateurs

Niveau de r... Facile ✖

Ajouter +

Par pages

Toutes ▾

Filter les valeurs

Par réussite

Toutes les réponses

Bonnes réponses

Mauvaises réponses

Générer 🖨

X	Y	Champ	KEHE	NILI	LUCH	SEBE	MAGU	BADU	ROBE	CEBA
51	92	Q10F6	00:00:46	00:01:06	00:00:58	00:01:09	00:01:57	00:00:56	00:00:57	00:00:40
51	10	Q4C3	00:00:10	00:00:13	00:00:16	00:00:14	00:00:06	00:00:19	00:00:10	00:00:43
51	38	Q6E11	00:00:14	00:00:22		00:00:35	00:00:24	00:00:24	00:00:25	00:00:34
10	75	question2	00:00:04	00:00:01	00:00:01	00:00:00	00:00:02	00:00:02	00:00:01	00:00:00
10	22	question5	00:00:02	00:00:00	00:00:01	00:00:00	00:00:05	00:00:01	00:00:00	00:00:00
27	81	question9	00:00:02	00:00:01	00:00:01	00:00:00	00:00:02	00:00:05	00:00:02	00:00:00

Exporter

Figure 30 : interface permettant d'afficher les données temps et les filtrer selon un indicateur spécifique extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.

Concernant les représentations graphiques, nous nous sommes basé sur l'analyse du besoin et des prototypes de représentations graphiques proposés par le chercheur pour l'analyse de ses données qui se résument dans les points suivantes :

- Représenter en nuage de points les instants de début d'écriture pour chaque question.
- Représenter en nuage de points les instants de fin d'écriture pour chaque question.
- Représenter en ligne continue les instants de début d'écriture pour chaque question.
- Représenter en ligne continue les instants de fin d'écriture pour chaque question.
- Représenter en histogramme les durées moyennes pour chaque question.
- Représenter en pourcentage empilé le pourcentage du temps passé sur chaque question par chaque répondant.
- Représenter en histogramme le taux de réussite pour chaque question.

Nous avons proposé quatre interfaces permettant de générer ces représentations graphiques dans notre application.

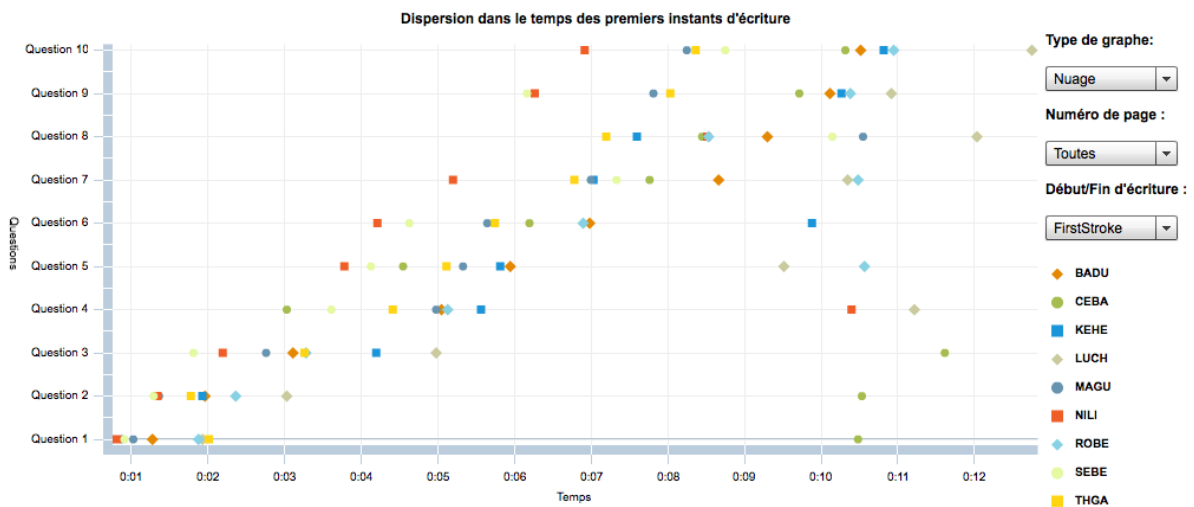


Figure 31 : représentation en nuage de points de la dispersion des données temps FS pour chaque question par tous les répondants extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.

(1) Une interface qui permet de générer un graphique de type chronologique qui représente la dispersion des réponses par FS (ou LS) dans le questionnaire par tous les répondants. Le chercheur a trois options d'affichage du graphique ; un choix par rapport au type de graphique soit en nuage de points soit en lignes, un choix par rapport aux données temps qui vont être représentées soit les instants de début de l'écriture dans les questions (FS) soit les instants de fin d'écriture, et un choix par rapport aux pages du test dans le cas où les données sont volumineux.

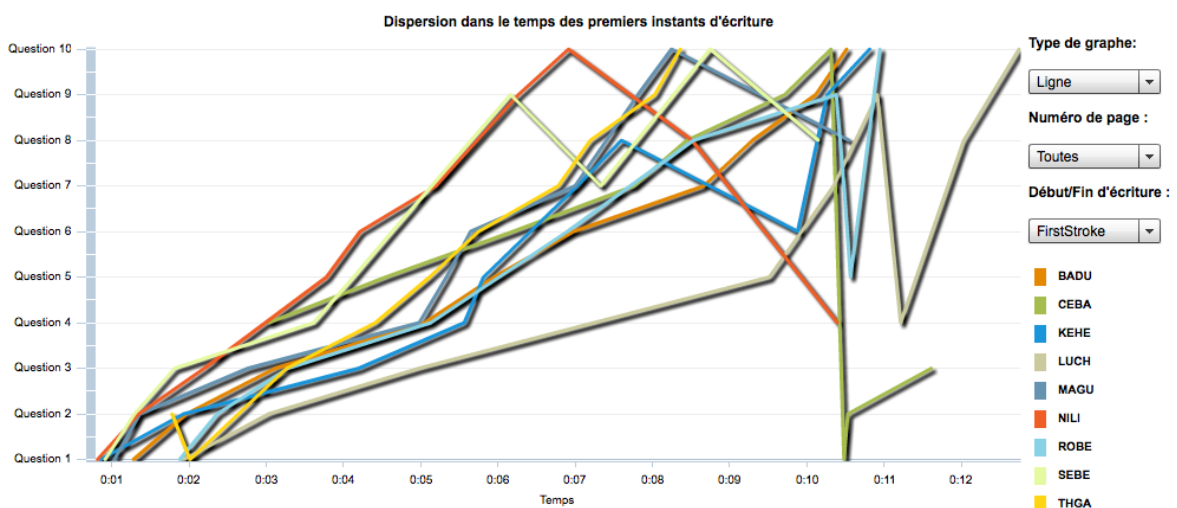


Figure 32 : représentation en lignes de la dispersion des données temps FS pour chaque question par tous les répondants extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.

Ces deux graphiques (Figure 31 et Figure 32) permettent de voir la tendance du groupe en terme de stratégie de traitement du test. On peut remarquer si les répondants traitent le test par ordre de présentation des questions et s'il existe des parcours particuliers. On peut également voir les questions qui ne sont pas traitées dans l'ordre.

(2) Une interface qui permet de générer un graphique de type histogramme qui représente des durées moyennes. Nous avons trois durées calculées automatiquement (durée totale, durée d'écriture, durée de non écriture) pour chaque question. Le chercheur peut choisir dans cette interface un indicateur personnalisé qui va être illustré en couleur des séries de la durée moyenne choisie.

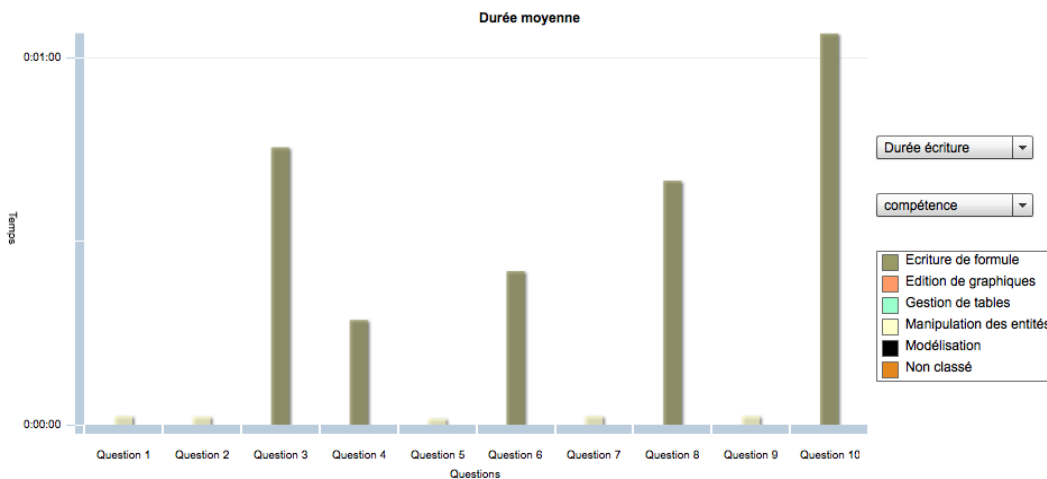


Figure 33 : représentation en histogramme de la durée moyenne d'écriture dans chaque question par rapport à l'indicateur « compétence » extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.

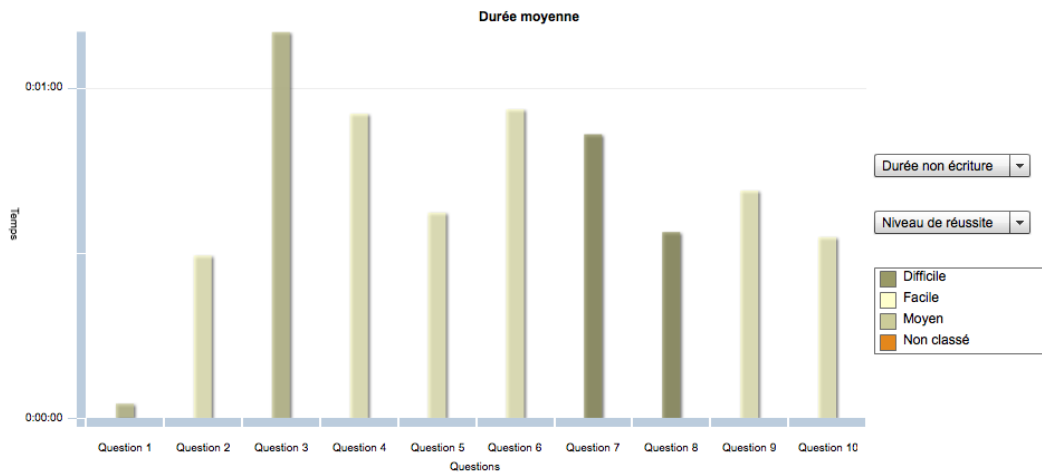


Figure 34 : représentation en histogramme de la durée moyenne de non écriture dans chaque question par rapport à l'indicateur « niveau de réussite » extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.

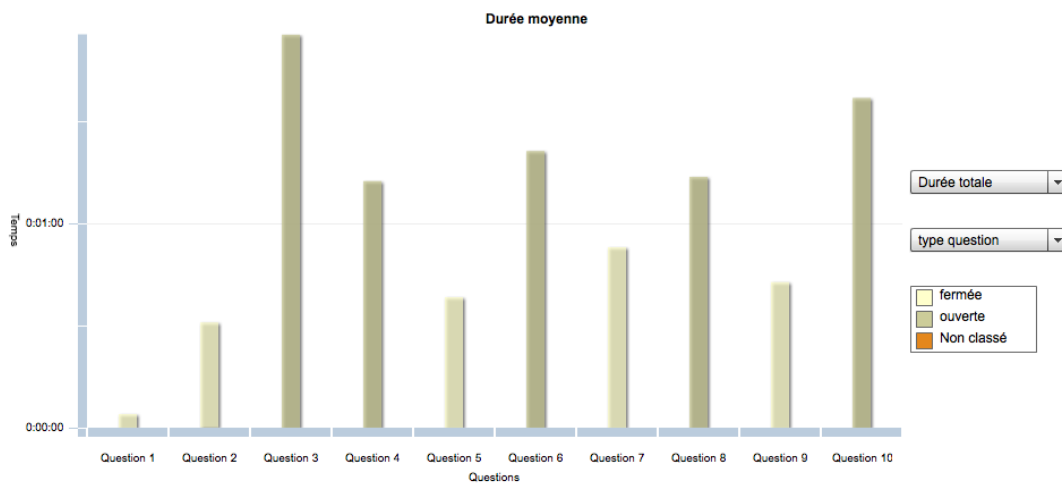


Figure 35 : représentation en histogramme de la durée moyenne totale (écriture + non écriture) dans chaque question par rapport à l'indicateur « type de question » extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.

(3) Un graphique en pourcentage empilé qui représente la distribution en pourcentage des durées totales passées sur chaque question par chaque répondant. Ce graphique montre les questions les plus consommatrices du temps (exemple : question 3 dans la Figure 36) et les questions qui ne sont pas.

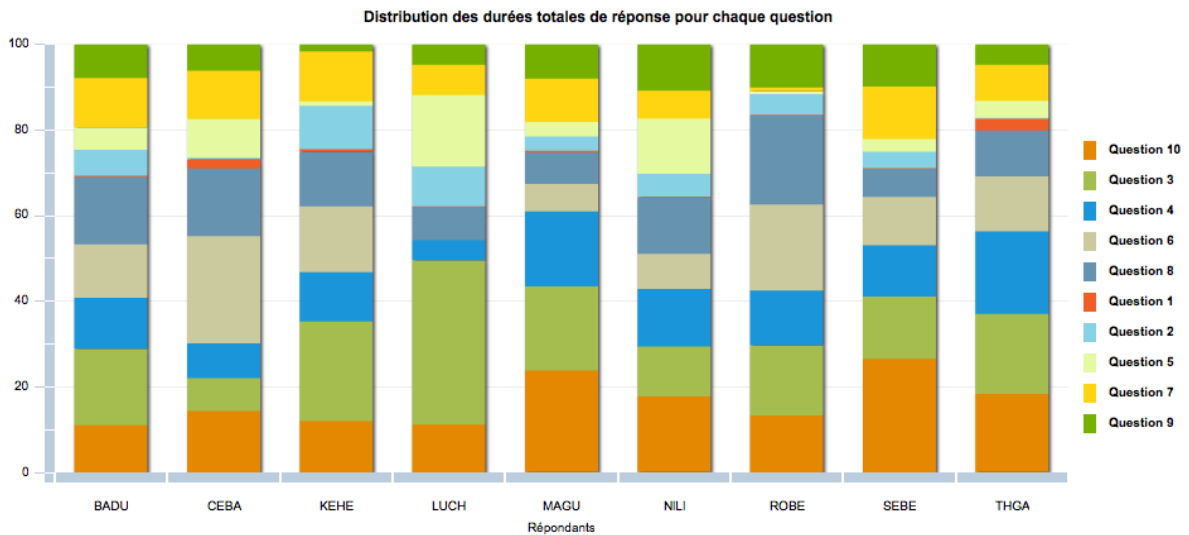


Figure 36 : représentation en pourcentage empilé de la durée totale passée dans chaque question par répondant, extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.

(4) Un graphique en histogramme qui représente le taux de réussite pour chaque question. Même si ce graphique ne représente pas des indicateurs temps, nous l'avons produit pour satisfaire le besoin du chercheur. En effet, le taux de réussite est un indicateur indispensable dans le cas d'analyse d'un test d'évaluation.

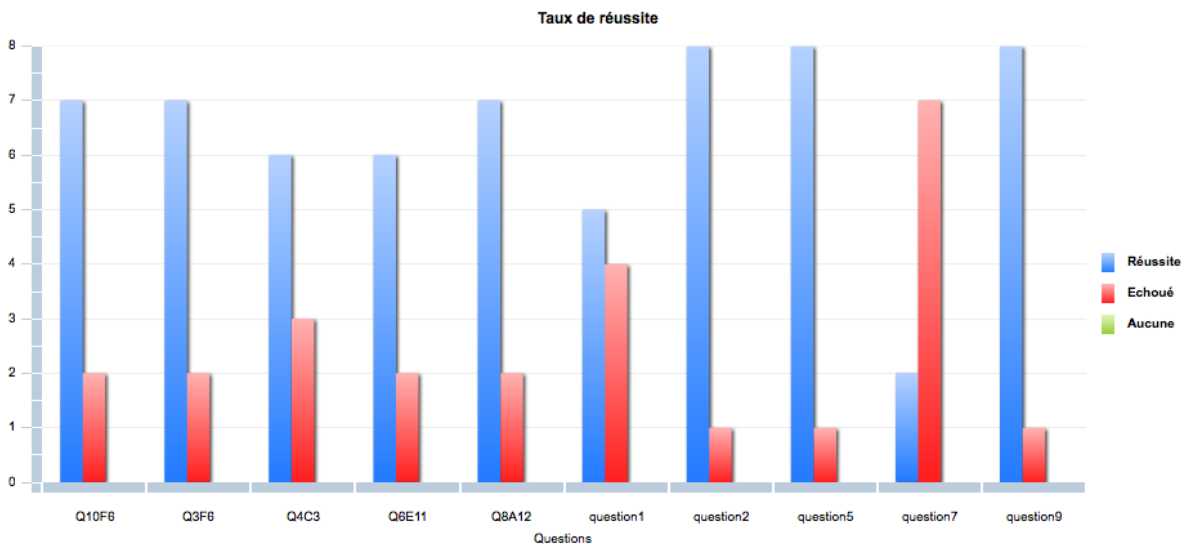


Figure 37 : représentation en histogramme du taux de réussite pour chaque question, extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.

Ces graphiques ont été conçus à la suite d'une première phase d'analyse des besoins par

rapport aux données recueillies et des indicateurs calculés. Une analyse avec ces représentations des trois tests tableur (Barhoumi & Tort, 2011), administrés auprès des étudiants de L3 économie et gestion de l'ENS Cachan, a permis d'avoir des remarques et des retours de la part du chercheur. Ces retours concernent l'interface et les traitements, d'une part, et concernent l'apport de ces traitements sur l'analyse des résultats de l'expérience, d'autre part.

Pour les deux tests tableur administrés auprès des étudiants formés en C2i, le chercheur s'est basé sur les représentations tabulaires que nous avons pu générer pour produire des représentations graphiques sur un logiciel tableur. Ses traitements tiennent compte des données temps et d'autres données extérieures nécessaires pour son étude et qui n'étaient pas recueillies avec le stylo numérique (par exemple des informations concernant les répondants). Toutefois, nous prenons en considération ses représentations et ses analyses pour l'amélioration de notre système.

2.7.3. Exemple d'analyse

Au niveau de l'apport des traitements, l'analyse des trois tests tableur administrés auprès d'étudiants L3 économie et gestion de l'ENS Cachan révèle des constats (Barhoumi & Tort, 2011).

La représentation des données FS et LS et les représentations chronologiques individuelles permettent de voir les stratégies de traitement du test par les répondants. Nous avons constaté que plus le test est difficile, plus les répondants changent de stratégie de réponse. En effet, les questions faciles sont traitées dans l'ordre normal (c'est à dire l'ordre de position de la question). Tandis que les questions difficiles sont celles les plus rapportées à la fin.

La représentation des durées moyennes montre que les questions fermées à choix unique ne sont pas consommatrices du temps généralement. En revanche, il est possible que la durée de latence de ces questions soit plus importante que les autres. De même, certaines questions considérées faciles ont consommé plus de temps que les questions difficiles. Pour cela, nous pouvons distinguer les questions qui consomment le plus de temps pour comprendre ce qui ne va pas avec ces questions. Il peut s'agir d'un problème de difficulté ou bien de formulation de la question. Par suite, cela peut être considéré dans l'évaluation des répondants.

Certes, les traitements effectués durant cette expérience ne sont pas sans ambiguïté, notamment la méthode de calcul des durées. En effet, les données recueillies avec le stylo numérique ne permettent pas d'avoir les durées de latence exactes. Pour cela, nous pouvons trouver des questions qui consomment plus de temps de latence alors que se sont des questions faciles. Il se trouve que ces questions sont précédées de questions difficiles sur lesquelles les répondants ont passé beaucoup de temps de lecture sans y répondre. Pourtant ce temps a été considéré avec le temps de latence de la question suivante. Dans tout les cas, ces données de durée ne devront être vérifiées sans que si le répondant ne soit observé ou bien filmé au moment de la passation du test.

Néanmoins, nous pensons que ces traitements peuvent être d'utilité dans le domaine des tests d'évaluation à grande échelle, comme le programme PISA de l'OCDE. D'ailleurs, en 2006, certains pays ont administré une évaluation assistée par ordinateur de l'alphabétisation scientifique des élèves dans le cadre du programme PISA de l'OCDE. Collecter « *de plus amples renseignements, tels que le temps passé par item, les frappes faites et le mouvement d'un élève à travers les questions*¹⁴ » a été parmi les avantages de la méthode citée dans (OECD, 2010). Malheureusement, ces données n'ont pas été considérées dans l'analyse de PISA 2006.

D'autre part, considérant les constats de l'analyse des tests, nous pensons qu'une analyse concentrée sur les questions qui composent un test pourrait être d'utilité dans le domaine d'analyse d'items. Cette technique docimologique, qui permet d'évaluer les questions d'un test, consiste en l'application d'un ensemble de procédés statistiques. Parmi ces procédés nous citons l'indice de difficulté d'un item, qui est comparable à l'indicateur de difficulté que nous avons considéré dans notre étude. Cet indice de difficulté représente en docimologie le pourcentage de réussite à la question (Bertrand, Blais, & Raïche, 2004).

Dans un contexte d'analyse d'item, nous pensons que la considération d'un indice temps de réponse et moyenne totale de réponse pour une question (ici un item) pourrait être enrichissante pour un modèle d'analyse d'items. D'ailleurs, un modèle émergent dans ce

¹⁴ Une traduction libre de « *more information, such as time taken per item, keystrokes made and a student's movement through the questions* » (OCDE, 2010).

domaine appelé le « le testing adaptatif » (Sodoké, Raïche, & Nkambou, 2007) considère le temps de réponse. Appliqué aux tests en ligne, ce modèle consiste à produire une suite de questions qui tiennent compte des paramètres associés au répondant (comme la culture et le niveau, etc.) et des données associées au déroulement du test (comme la maîtrise du sujet, le temps de réponse, l'état émotionnel, etc.). Le système choisit, en temps réel, la question suivante dans un test en ligne selon l'analyse de ces données.

Certes, ce dernier modèle ne peut pas être utilisé dans le cas des tests papiers, toutefois, l'idée de considérer les données temps dans l'évaluation des questions d'un test classique sur papier tramé reste à expérimenter dans ce domaine. De même, l'application d'une analyse tel que la notre, qui considère des indicateurs temps, sur des tests en ligne est une autre piste intéressante à expérimenter.

2.8. Feedback de l'expérience

A partir des analyses des tests d'évaluation en tableur instrumentés avec le stylo numérique, nous retenons les retours des chercheurs qui ont participé à cette étude. Ces retours nous permettent d'analyser notre production en terme d'utilité des traitements d'une part, et en terme de performance des interactions de l'interface et des représentations proposées d'autre part. Nous analysons également la méthode d'usage et d'instrumentation avec la technologie du papier et stylo numériques dans un contexte d'expériences en éducation pour le recueil des données.

2.8.1. Instrumentation des formulaires

La phase d'instrumentation consiste à spécifier la structure du formulaire du test. Cette phase est précédée par la conception du contenu du test, faite par le chercheur, et la conception de la disposition du contenu dans les pages du test.

Pendant cette phase, les contraintes technologiques exigent de n'utiliser que les types de questions QCM, c'est à dire des questions de type texte, case à cocher pour choix multiple et bouton radio pour choix unique. En effet, même si le stylo numérique capte tout tracé sur un papier tramé, le système de récupération et de reconnaissance de Kayentis ne permet de représenter que les données écrites dans les zones délimitées dans la phase d'instrumentation et

qui contiennent des champs d'écriture de type texte, case à coché ou bouton radio.

Ce qui est intéressant dans cette phase c'est la détermination a priori de la composition structurelle du formulaire. La spécification de la nomenclature et du type des champs d'écriture représente un prétraitement de ces données qui facilite par la suite le traitement de la trace d'écriture. Cependant, chaque test devrait avoir sa propre spécification et fait l'objet d'une structure de formulaire à part. Même toute modification du contenu d'un test nécessite une nouvelle version de formulaire composé d'une structure différente. Ce qui peut causer des problèmes au niveau de l'automatisation des traitements des données de la trace d'écriture du fait que nous utilisons les données de description de la structure d'un formulaire dans notre application. Pour cela, il est nécessaire de tester le questionnaire et de valider sa composition avant de l'avoir passer auprès des étudiants.

Au niveau du matériel utilisé, l'administration de test auprès d'étudiants dans une classe nécessite un nombre suffisant de stylos permettant à toute la classe de passer le test au même temps. Cependant, nous n'avions qu'une dizaine de stylos numériques à la disposition des étudiants. En conséquence, nous avons récupéré la trace d'écriture d'un nombre très limité d'étudiants.

Après la phase d'administration des tests, nous récupérons les données depuis les stylos dans la plateforme de Kayentis. Ensuite, avant de pouvoir récupérer la trace en format XML, il était nécessaire de vérifier et de valider la reconnaissance de certains écrits dans les exemplaires par les chercheurs.

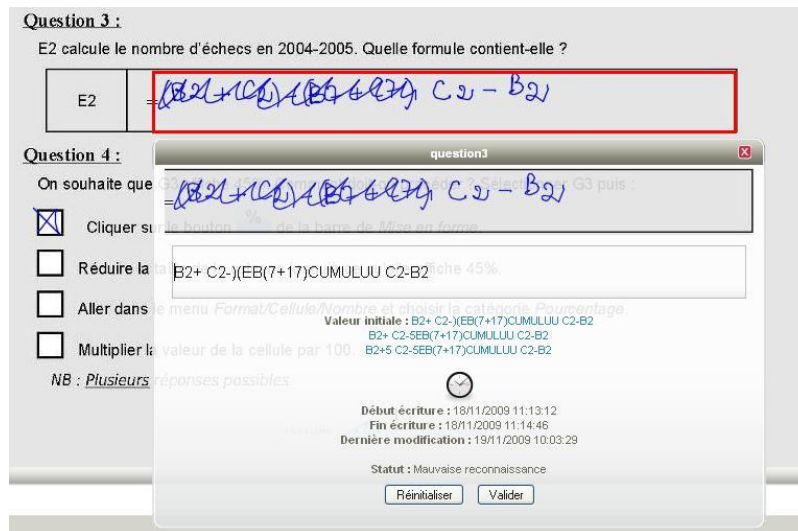


Figure 38 : Exemple de zone d'écriture qui n'est pas validé par le système de reconnaissance.

Par exemple, si le répondant écrit une réponse puis change d'avis et écrit une autre réponse, le système de reconnaissance laisse la zone d'écriture dans un encadré rouge pour vérifier et valider son contenu par l'utilisateur (voir l'exemple Figure 38). Malheureusement ce système ne garde pas le moment de modification de la réponse. On ne peut pas savoir si le répondant a modifié sa réponse après avoir répondu à une autre question ou pas. Pour cela, le calcul des données temps que nous avons utilisé dans notre application est approximatif.

2.8.2. L'interface d'utilisation

L'outil que nous avons développé dans le cadre de cette étude est un prototype d'une plateforme en ligne (SEGELLBETA). Dans ce prototype nous avons développé un ensemble de traitements et des interfaces utilisateur permettant de visualiser les données traitées à travers des tâches d'interaction. Toutefois, Plusieurs améliorations, notamment d'interaction, sont recommandées.

En effet, les traitements développés sont concentrés sur la sélection, le filtrage et le croisement des données. Ainsi, le paramétrage des interfaces qui permettent de générer des tableaux de données ou bien des graphiques était intéressant également.

Cependant, dans le but de pouvoir comparer les parcours par exemple, il était souhaitable de pouvoir choisir d'afficher les données qui correspondent à un seul répondant, ou bien à un groupe de répondants. De même, le choix d'afficher les données qui correspondent à une (ou

plusieurs) question permet de voir son rang de traitement par tous les répondants dans le même graphique.

Finalement, les chercheurs ont exprimé la nécessité de faire un graphique qui représente l'ordre de traitement des questions. Nous avons produit ce paramètre du fait que ça sera visible dans les représentations chronologiques des données. Cependant, il semble être nécessaire de quantifier le nombre de fois où une question a été traitée dans un ordre spécifique.

2.9. Conclusion

En collaboration avec des chercheurs en sciences de l'éducation, nous avons produit une solution qui aide à analyser ces données à travers un ensemble de traitements et de représentations. Nous avons un intérêt particulier aux données temps dans l'exploitation de la trace du stylo numérique dans le but d'enrichir le cercle d'analyse habituel dans le domaine de l'évaluation des compétences. Nous avons déterminé un ensemble d'indicateurs temps à partir du traitement des données recueillies, et nous avons représenté ces indicateurs, en représentation tabulaire ou bien graphique, par rapport à des indicateurs propres au contexte de l'analyse expérimenté.

Dans ce chapitre, nous avons présenté les travaux effectués dans le cas de traitement et d'analyse d'un test d'évaluation en classe. La spécificité de ce test, c'est qu'il est ordonné et nécessite au répondant d'avoir des connaissances vues en classe. Dans le chapitre suivant, nous traitons le cas de test d'évaluation non ordonné.

Chapitre 3. L'usage du stylo numérique pour des tests non ordonnés

3.1. Introduction

Dans le même contexte d'usage du stylo numérique pour la passation de tests d'évaluation, nous avons effectué d'autres tests personnalisés de formats originaux. Il s'agit de faire passer des questionnaires non ordonnés qui ne nécessitent pas de répondre selon un ordre imposé explicitement ou implicitement. Le but de ces expériences est d'explorer le traitement et l'analyse de différentes formes de tests. Nous avons personnalisé volontiers la forme et le contenu de ces tests pour faire émerger des traitements et des représentations nouvelles allant avec des objectifs prédéfinis.

Ainsi, nous avons organisé, en collaboration avec Françoise Tort, la passation de deux tests non ordonnés dans deux contextes différents.

Le premier a été administré auprès des chercheurs de la conférence EuSpRiG 2009. Ils étaient questionnés sur leur capacité à reconnaître et classer des erreurs effectuées dans l'interface d'un logiciel tableur (l'un des sujets principaux de la conférence).

Le second test, administré lors des Rencontres des jeunes chercheurs en EIAH 2010, portait sur des questions de culture générale. Il s'agissait de relier des photos de paysages aux pays d'où ils provenaient.

Dans ce chapitre, nous présentons le déroulement de chaque expérience depuis la définition du besoin jusqu'à la présentation de traitements et des représentations des données. Nous avons développé pour chaque expérience un prototype pour réaliser les traitements et les représentations. Enfin, nous présentons une évaluation des prototypes produits ainsi que la méthode d'instrumentation de chaque expérience.

3.2. L'expérience EUSPRIG

3.2.1. Contexte et objectif de l'expérience

À l'occasion de la 10^e édition de la conférence annuelle EUSPRIG (European Spreadsheet

Risks Interest Group), qui s'est déroulée à l'ENS Cachan, nous avons organisé, avec Françoise Tort, la passation d'un test de connaissance sur le thème de la conférence. Il s'agissait de reconnaître une série de cinq erreurs souvent effectuées sur tableur et de les classer selon la classification des erreurs définie par Rajalingham et al¹⁵ en 2000 à EUSPRiG.

Étant destiné aux participants d'une conférence anglophone, le questionnaire doit être en anglais et adapté à leur domaine d'expertise. Le thème du questionnaire, choisi par Françoise Tort, est à la fois général par rapport à ce domaine, puisque ça évoque les erreurs sur tableur, mais aussi spécifique à une classification peu ou moyennement connue par la communauté. Par suite, les participants peu familiarisés avec la classification de Rajalingham vont fournir un effort cognitif considérable soit pour se rappeler soit pour bien réfléchir sur une classification des erreurs plus logique.

Pour cela, nous avons voulu savoir le degré de familiarité des participants avec la classification de Rajalingham. Nous avons établi l'hypothèse suivante : Si les participants sont bien familiarisés avec cette classification alors ils vont répondre plus rapidement et/ou dans l'ordre de position des erreurs dans l'énoncé que ceux qui connaissent peu ou pas de tout la classification.

Par le traitement et l'analyse des données temps recueillies à l'issue de ce test, notre objectif est de savoir si des parcours de réponses particuliers peuvent émerger. Nous mettrons en avant la durée que chaque répondant prendra pour répondre au questionnaire, l'ordre de traitement des questions ainsi que le degré de familiarité des participants avec la classification de Rajalingham.


3.2.2. *Conception et instrumentation*

Le questionnaire doit être court et clairement lisible pour que les participants comprennent très vite l'objectif du test. En phase de conception, nous avons essayé de bien disposer les questions et les réponses dans les pages.


¹⁵ Rajalingham, K., Chadwick, D. R., et Knight, B., (2000), *Classification of Spreadsheet Errors*, Eusprig 2000, <http://arxiv.org/ftp/arxiv/papers/0805/0805.4224.pdf> (consulté le 27 juillet 2009).

Le questionnaire est présenté sur deux pages en paysage. Au début, nous avons fait deux pages disposées l'une en portrait et l'autre en paysage. Mais nous avons rencontré des problèmes techniques au moment de l'instrumentation sur Acrobat Professionnel. Donc nous avons remis les deux pages en paysage.

Le contenu de la première page comporte l'énoncé de l'exercice avec la description des cinq erreurs (voir Figure 39).



Classifying students' spreadsheet errors
A test using a digital pen



Below, we describe five errors made by French students in a spreadsheet computer test. You have to determine the category of each error according to the Rajalingham *et al.* (2000) classification, as revised by Purser and Chadwick (2006).

Please, characterize your familiarity with the Rajalingham classification: None Fair Good Excellent

A. Problem submitted to students:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Cells O5, P5 and Q5 automatically display the sum, the average and the median of the values entered in the cells A5 to N5.																
2																	
3	Write the formula in O5, in P5 and in Q5.																
4															Sum	Average	Median
5	22	15	63	84	25	31	76	40	92	49	13	54	65	10			
6																	

Expected values: Sum = 639, Average = 45.6, Median = 44.5

B. Description of the five observed errors:

Error 1
Value in O5 (Sum): 639
Formula in O5:
=A5+B5+C5+D5+E5+F5+G5+H5+I5+J5+K5+L5+M5+N5
Error description: *use of the + operator instead of the SUM function*

Error 2
Value in O5 (Sum): 614
Formula in O5:
=A5+B5+C5+D5+F5+G5+H5+I5+J5+K5+L5+M5+N5
Error description: *omission of E5*

Error 3
Value in P5 (Average): 45.6
Formula in P5: =O5/14
Error description: *hard-coding the number of values by a constant*

Error 4
Value in P5 (Average): 85.2
Formula in P5: =AVERAGE(A5:O5)
Error description: *use of an incorrect cell range (A5:O5)*

Error 5
Value in Q5 (Median): 7
Formula in Q5: =COUNT(A5:N5)/2
Error description: *use of an incorrect function for calculating the median*

Rajalingham, K., Chadwick, D. R., and Knight, B., (2000) "Classification of Spreadsheet Errors", EuSpRIG 2000
Purser, M., Chadwick, D. (2006) , "Does an awareness of differing types of spreadsheet errors aid end-users in identifying spreadsheet errors?", EuSpRIG 2006

Figure 39 : Le contenu de la première page qui contient l'énoncé et la position du problème.

Au préalable, le participant doit indiquer son niveau de familiarité avec la classification Rajalingham. Pour cela, dans la première page en haut, nous avons mis une question fermée à choix unique qui permet aux participants d'indiquer leur niveau de familiarité avec cette classification. Un choix parmi quatre suppositions : none, fair, good et excellent (voir Figure 40).

Please, characterize your familiarity with the Rajalingham classification: None Fair Good Excellent

Figure 40 : une zone pour notifier le niveau de familiarité avec la classification Rajamingham.

La réponse à cette question constitue un top de départ pour le temps de passation du test.

Nous avons également ajouté un champ de saisie pour un nombre de deux chiffres qui nous a servi comme identifiant anonyme de chaque participant.

La deuxième page contient l'arbre de classification et les champs de réponses aux questions. La disposition horizontale de cet arbre tel qu'elle était présentée par ses auteurs (voir Figure 41) nous semble difficile à aborder. Pour cela, nous avons orienté l'arbre verticalement (voir Figure 42).

Pour les cinq erreurs décrites dans la première page, le participant doit inscrire le numéro de l'erreur (de 1 à 5) dans un cercle au-dessous d'une classe d'erreur parmi dix. Une erreur peut être classée sous plusieurs classes.

Le choix de cette forme de réponse est contraint par le fait que dans un champ de saisie on ne peut pas avoir les instants de chaque modification apportée. Les données temps récupérées ne contiennent que les instants du début et de la fin de l'écriture dans un champ de saisie. Et comme on peut mettre plus qu'une erreur pour une classe, alors la nécessité de faire plusieurs champs de saisie au dessous de chaque classe (jusqu'à 5 le nombre d'erreurs).

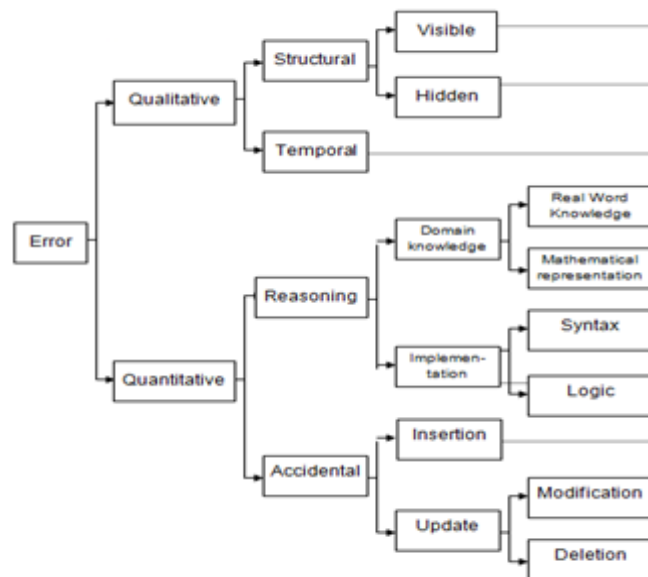


Figure 41 : L'arbre de classification selon Rajalingham et al (2000) révisé par Purser (2006)
<http://www.eusprig.org/2006/purser-awareness-types-spreadsheet-errors.pdf>.

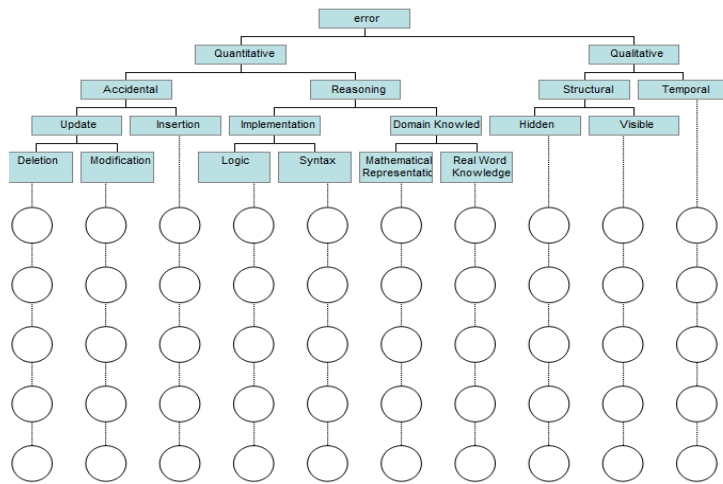


Figure 42 : La nouvelle disposition de l'arbre de classification en vertical.

Après la conception de la forme du formulaire, nous passons à la phase d'instrumentation sur le logiciel Acrobat professionnel d'Adobe qui consiste à définir les zones tramées à l'aide du plugin de Keyentis.

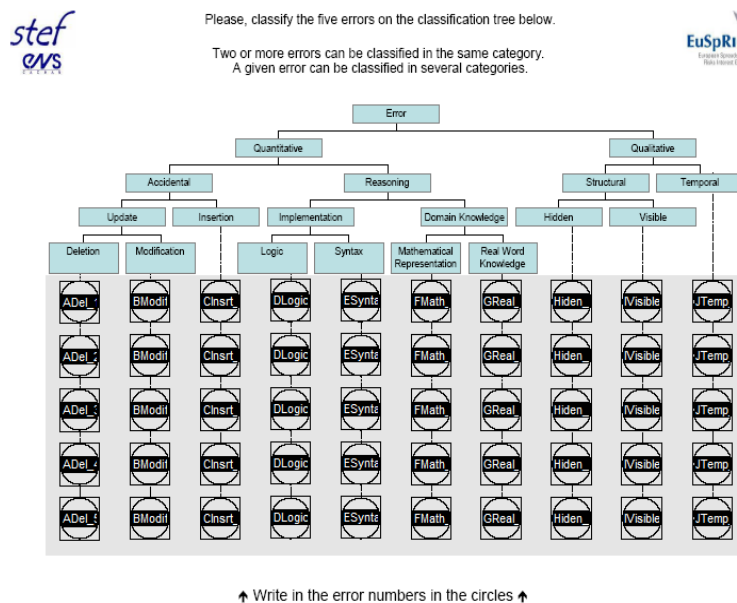


Figure 43 : Les zones de saisie sur chaque cercle au-dessous de la classe qui lui correspond.

Ces zones de saisie vont être reconnues par le stylo au moment de l'écriture. Pour cela, tout champ de saisie (voir Figure 43), bouton radio ou case à cocher, doit être tracé et nommé avec un nom unique et explicatif de la valeur qu'il contiendra. Les paramètres pour la reconnaissance comme la langue et le seuil de confiance sont ensuite définis pour chaque champ.

3.2.3. Administration du test

Nous avons utilisé 3 stylos numériques pour la passation de ce test. Le temps prévu de remplissage du formulaire du questionnaire dure entre 7 et 10mn. La passation s'est déroulée sur deux demies journées, la veille de la conférence et le matin du premier jour de la conférence. En effet, nous avons profité de la présence de certains participants à la conférence au cours d'un exposé pour leur demander de répondre au questionnaire la veille de la conférence. Cela nous a permis de prévoir une organisation meilleure pour le lendemain :

- expliquer aux participants le but du test, l'énoncé et où ils doivent répondre,
- demander de commencer par cocher la case concernant leur familiarité avec la classification,
- surveiller les participants au cours de la passation : certains participants ont eu le réflexe de mettre des croix dans les cercles au lieu de mettre des chiffres (le numéro de l'erreur),
- prévoir un formulaire rempli et le présenter comme modèle,
- prévoir un emplacement adapté (des box) pour être en situation d'examen pour réfléchir sur les réponses.

Généralement, la passation s'est bien déroulée. Nous avons récupéré 33 formulaires remplis dont deux inexploitable pour des raisons techniques. Nous avons rassemblé tous les exemplaires et les stylos pour récupérer les données et commencé la phase de vérification, de validation sur la plateforme de Kayentis et de traitement. Nous avons traité les données recueillies le soir de la passation avec une première version de l'outil SEGELL et nous avons présenté des maquettes des représentations graphiques des traitements programmés.

3.2.4. *Traitement des données*

Le premier traitement effectué pour cette expérience est l'importation des données XML dans une base de données. Notre objectif est de voir des parcours chronologiques de réponse particuliers. Nous voulons avoir les durées de traitement des questions et les présenter par rapport à la valeur de familiarité avec la classification Rajalingham déclarée par les répondants.

Le prototype que nous avons développé pour traiter les données de cette expérience permet d'importer uniquement les données de la trace d'écriture des répondants. Cette trace contient

des informations concernant le formulaire, le répondant et les données temps relatives à chaque écriture dans les champs de réponses. Nous avons donc calculé la durée totale passée dans chaque question (on a 5 questions) à l'aide des instants d'écriture dans les champs de réponses. La valeur de familiarité est récupéré a posteriori pour chaque répondant depuis le champ « familiarité ». Et pour vérifier notre hypothèse sur le rapport entre la familiarité et la rapidité de réponse, nous avons calculé la durée moyenne passé par les répondants de chaque catégorie de familiarité sur l'ensemble du test.

Ainsi, sur les données recueillies nous avons effectué les traitements suivants :

- Le nombre de personnes pour chaque catégorie de familiarité (None, Fair, Good et Excellent).
- La durée moyenne passé par les répondants d'une catégorie de familiarité sur l'ensemble du test.
- La durée totale passée dans chaque question par tous les répondants.
- Le nombre de réponses proposées pour chaque question par tous les répondants.

3.2.5. Les représentations graphiques

A l'aide du premier prototype, nous avons établi des maquettes de représentations graphiques pour visualiser les traitements effectués. Nous présentons ci-après l'illustration de ces représentations graphiques :

- **Un graphique histogramme qui représente le nombre de personnes pour chaque catégorie de familiarité (familiarité / nombre de personnes)**

Cette représentation illustre la répartition des répondants selon leur réponse à la question de familiarité avec la classification (voir Figure 44). Ici, la majorité des répondants ont déclaré n'avoir aucune connaissance.

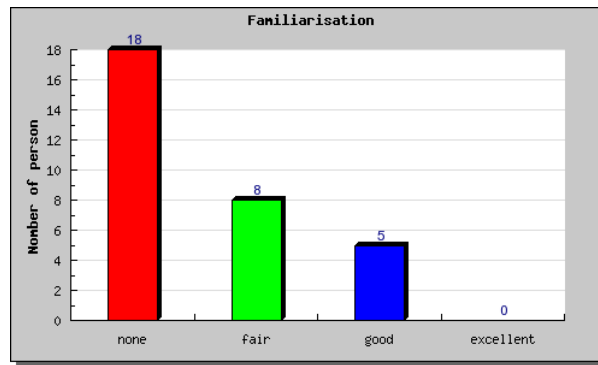


Figure 44 : Le nombre de personnes répartie selon leur familiarité avec la classification Rajalingham.

- Un histogramme qui représente la durée totale passée dans chaque question (numéro de question / durée totale en secondes)

La représentation illustre l'ensemble des durées passées en secondes dans chaque question par tous les répondants. La durée qui correspond à la première question (en rouge) semble plus longue que les autres du fait qu'elle comporte aussi le temps de découverte de l'énoncé du test. Une lecture et une réflexion globale sur l'ensemble du test étaient nécessaires.

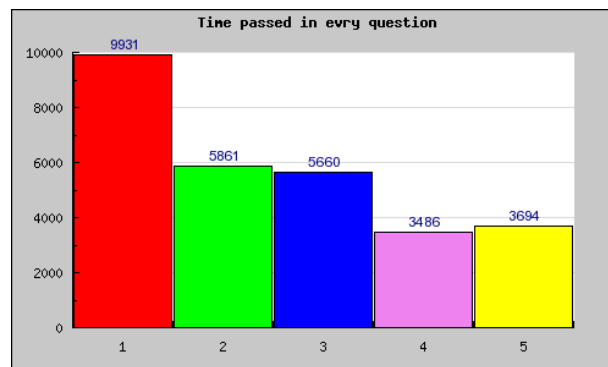
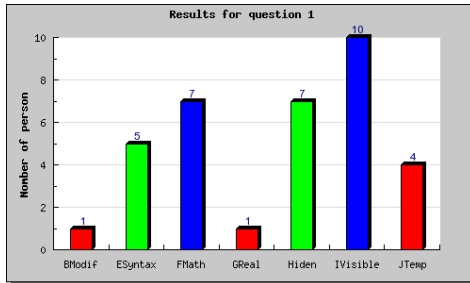


Figure 45 : les durées globales passés dans chaque question par tout les participants.

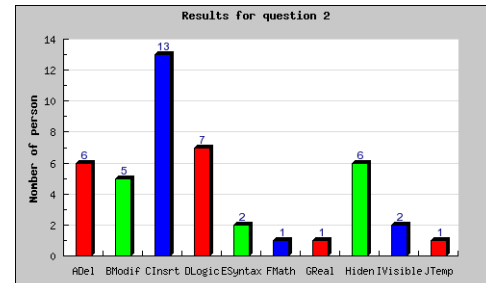
Ensuite, nous remarquons un temps moyen pour les deux questions qui suivent. Cependant, à la question 4 et 5 nous remarquons une chute de la durée passée. Cette chute peut s'expliquer, soit par le fait que les gens ont bien compris la problématique et deviennent plus rapides, ou au contraire, par le fait que ces deux questions étant plus difficiles, ils les abandonnent plus vite.

- Un histogramme pour représenter les réponses pour chaque question (les réponses proposées / nombre de personnes)

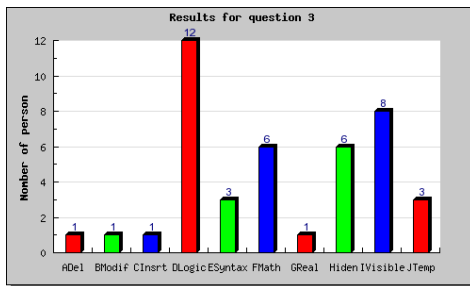
Cette forme présente les réponses aux questions en calculant le nombre de personnes par réponse proposé pour chaque question. Ces résultats sont illustrés dans les graphiques ci-après pour les 5 questions.



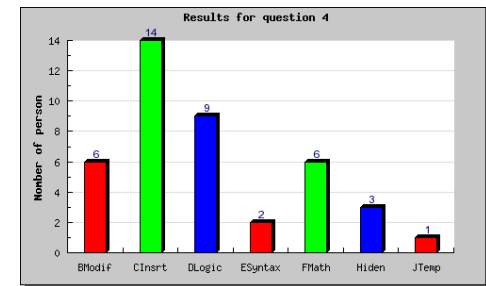
Graphique 1: le nombre de personnes par réponse proposé pour la question 1.



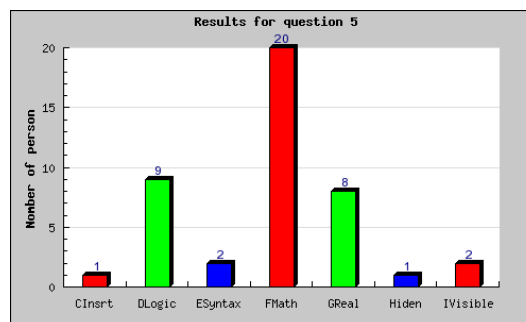
Graphique 2: le nombre de personnes par réponse proposé pour la question 2.



Graphique 3 : le nombre de personnes par réponse proposé pour la question 3.



Graphique 4 : le nombre de personnes par réponse proposé pour la question 4.



Graphique 5 : le nombre de personnes par réponse proposé pour la question 5.

- Une courbe avec marque pour illustrer le parcours des réponses pour chaque personne (chrono / les 5 réponses)

Nous avons voulu avoir le parcours de chaque personne (l'ordre des questions traitées) en fonction du temps. Pour cela, nous avons fait des graphiques courbes linéaires avec marque pour les 5 questions en ordonnées (on a ajouté le 6 pour la lisibilité du graphique), et les instants de la première écriture dans le champ correspondant à la question en abscisse.

Chaque point de cette courbe représente le couple (l'instant de début d'écriture, le numéro de la question). Chaque personne est représentée par une série de points d'une couleur aléatoirement choisie. Cela nous permet de voir l'ordre chronologique de traitement des questions de l'ensemble des répondants.

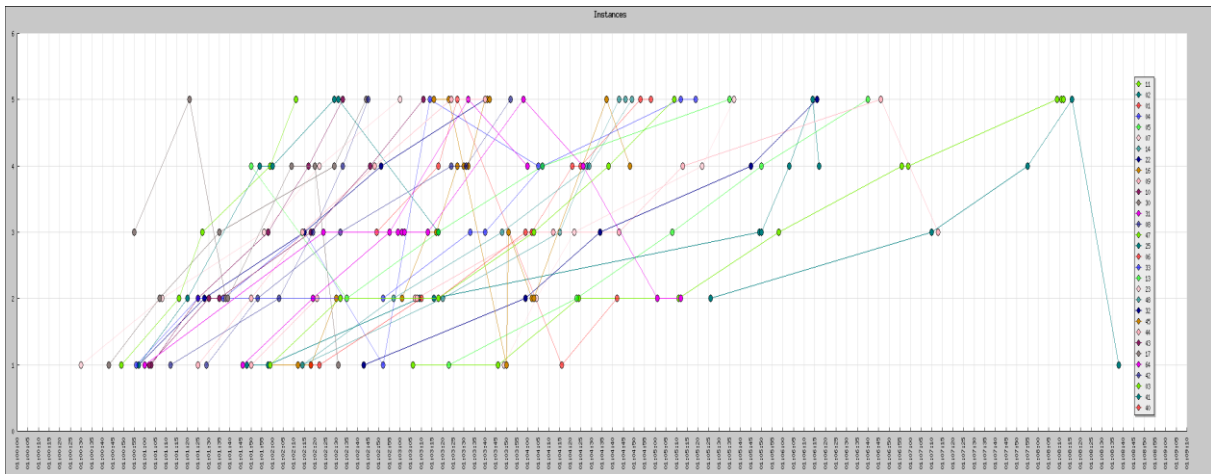


Figure 46 : les parcours chronologiques de réponses aux questions de 1 à 5 dans l'axe des ordonnées, en fonction du temps dans l'axe des abscisses, de tous les répondants.

Nous remarquons qu'il y a une tendance à faire ou à commencer à faire les questions dans un ordre "classique", mais après il y a eu des va-et-vient entre les questions selon plusieurs scénarios. Nous voyons également des parcours bien compliqués et d'autres marquant plus de temps pour passer d'une question à l'autre.

- **Une courbe avec marque pour représenter le parcours chronologique des réponses pour les personnes de la catégorie 'good familiarity'**

Dans ce graphique, nous avons regroupé les parcours des personnes déclarant une bonne familiarité avec la classification.

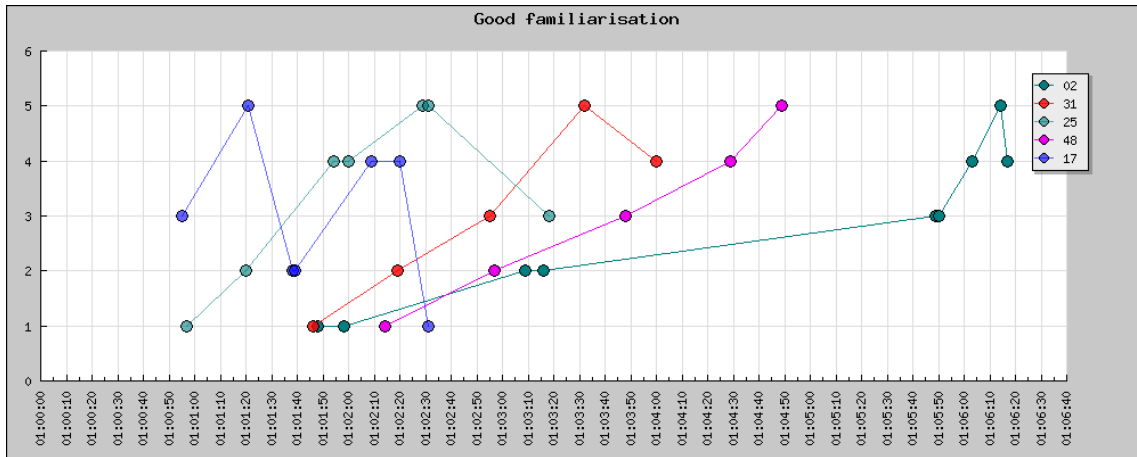


Figure : les parcours chronologiques des réponses aux questions, de 1 à 5 dans l'axe des ordonnées, en fonction du temps dans l'axe des abscisses, pour la catégorie 'good familiarity'.

Nous remarquons que les répondants de cette catégorie ont presque fait les questions dans l'ordre. Cependant, il reste à savoir s'ils ont été plus rapides que les autres répondants.

- Un histogramme pour représenter les durées moyennes passées par chaque catégorie de répondants pour finir le test (durée moyenne / catégorie de familiarité)

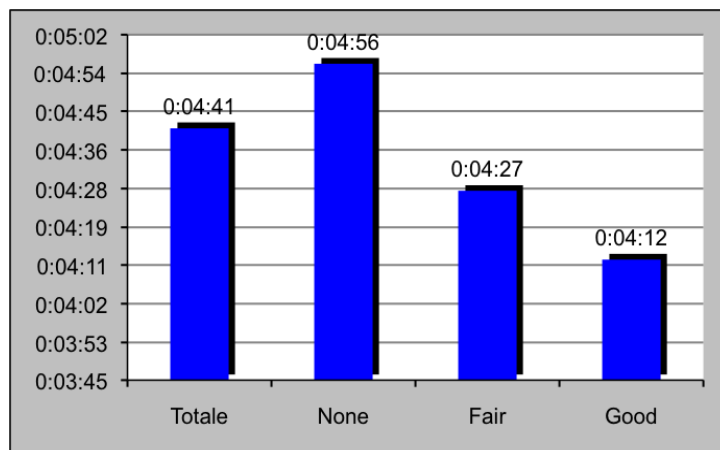


Figure 47 : les durées moyennes passées par chaque catégorie de répondants.

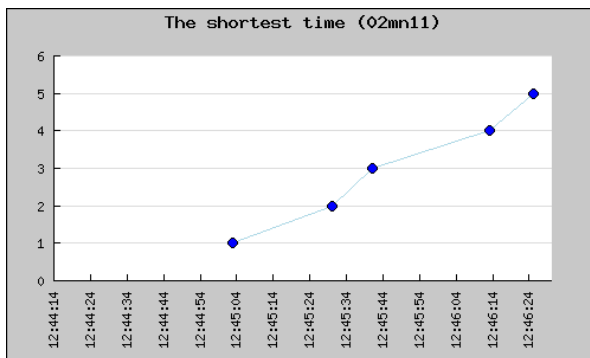
Pour vérifier si les répondants, qui ont une bonne familiarité avec la classification de Rajalingham, ont été plus rapides que les autres, nous avons calculé les durées moyennes passées sur le test par tous les répondants et par chaque catégorie de répondants (None, Fair et Good).

Le résultat montre clairement que l'hypothèse sur la familiarité et la rapidité se manifeste bien. En effet, les répondants qui ne connaissaient pas la classification étaient plus lents que

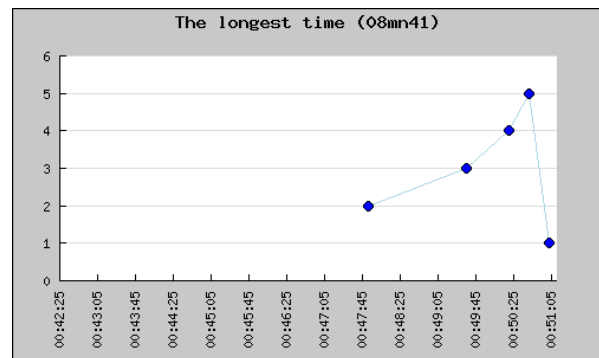
les répondants qui connaissaient un peu ou plutôt bien la classification. De plus, ces derniers ont répondu presque dans l'ordre de position des questions.

- **Une courbe avec marque pour représenter le parcours chronologique des réponses pour la personne la plus rapide et la personne la plus lente**

Nous avons représenté deux parcours extrêmes, celui du plus rapide et celui du plus lent. Nous remarquons que la personne la plus lente, et qui n'avait aucune connaissance préalable de la classification, a pris du temps pour commencer à répondre. En plus, elle a fait de longues durées entre chaque réponse. Cela explique une implication cognitive et une concentration sur la classification. Alors que la personne la plus rapide, et qui a une faible connaissance de la classification, a commencé rapidement et a fait les réponses successivement avec des durées courtes entre les réponses.



Graphique : le parcours de la plus rapide personne.



Graphique : le parcours de la plus lente personne.

- **Des courbes avec marques pour représenter les parcours chronologiques irréguliers**

Nous avons observé l'existence de certains parcours irréguliers, c'est à dire qui n'ont pas respecté l'ordre classique. Nous avons extrait ces parcours et nous les présentons dans le graphique ci-après (voir Figure 48). Les personnes qui correspondent à ces parcours ont déclaré une faible connaissance de la classification, certains n'avaient aucune. Nous remarquons plusieurs va et vient entre les questions.

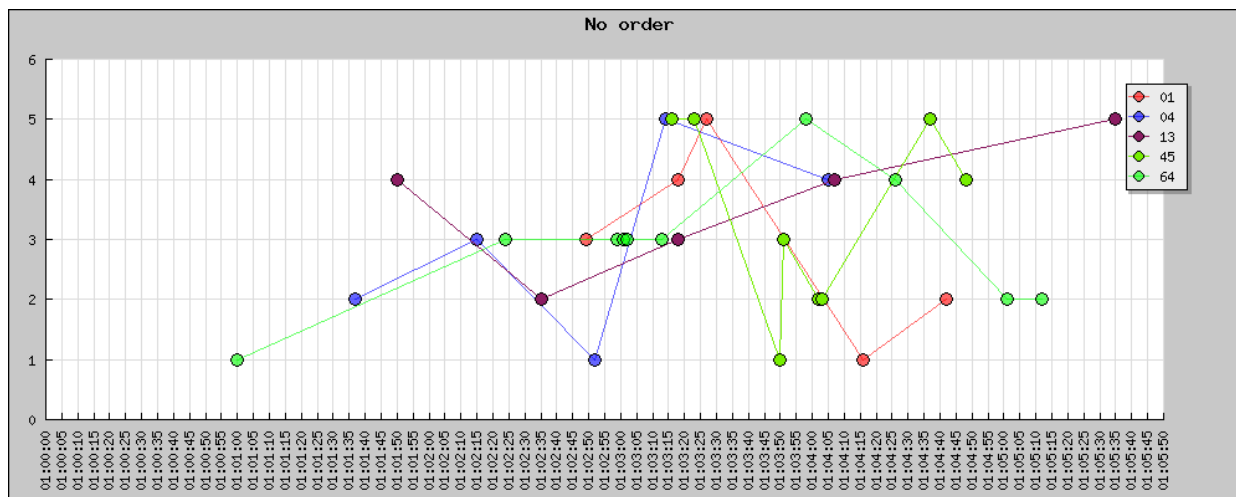


Figure 48 : des parcours variés qui n'ont pas respecté un ordre habituel.

3.2.6. Feedback de l'expérience

L'organisation de cette expérience constitue une première exploration de méthode de recueil et de traitement de données de trace d'écriture avec un stylo numérique. Nos objectifs sont l'instrumentation d'un test non ordonné, le traitement des données selon des besoins que nous avons spécifié et la représentation de ces données, notamment les données temps, pour faciliter l'analyse des résultats.

Au cours de cette expérience nous avons pu traiter et analyser une trentaine d'exemplaires de formulaire pour avoir des résultats dans les deux jours du colloque. Nous avons développé l'ensemble des traitements et des représentations à l'aide d'un premier prototype de notre application SEGELL programmé en PHP. Nous avons fait des maquettes de représentation des résultats du questionnaire par rapport aux données temporelles comme la durée de passation pour chaque question, la durée moyenne passé par les participants par catégorie de familiarité et les parcours chronologiques de réponse.

L'affichage des parcours chronologique selon les réponses : ce qui caractérise cette expérience c'est que nous avons affiché les courbes chronologiques par réponse (la valeur écrite par le répondant) et pas par champ de réponse (le nom de la zone de réponse défini à l'instrumentation).

La classification des répondants selon la familiarité : cette classification est déclaré par le répondant dans le champ « familiarity ». Nous voulons prendre compte de cette valeur pour classer les répondants. Pour ce faire, il faut avoir la possibilité de transformer la valeur d'un champ en catégorie de la population. Ensuite, il faut tenir en compte de cette valeur dans les graphiques comme indicateur.

Le filtrage des répondants pour les graphiques courbe de parcours chronologique : le filtrage comme moyen de sélection d'une catégorie de répondant garantit la lisibilité du graphique lorsque nous voulons nous focaliser sur certains parcours particuliers. Pour notre cas, nous avons effectué des filtres sur les séries pour voir les parcours des répondants d'une catégorie de familiarité ou de parcours particulier.

Pour les indicateurs temps, nous avons eu besoin de calculer les durées moyennes du temps passé sur le test ainsi que les durées globales passées sur chaque question par tous les répondants.

Les résultats de cette courte expérience sont intéressants du fait que nous avons pu sortir des graphiques en fonction du temps. Cependant, pour ce type de test, ces analyses devraient être enrichies et croisées avec d'autres données ou indicateurs. La détermination de ces croisements entre les données temps et les indicateurs dépend du contexte du questionnaire et de l'objectif de l'enquêteur. Ainsi, ces données peuvent être synchronisées avec d'autres types de média (son, vidéo, grille d'observation, ...).

3.3. L'expérience RJC

3.3.1. Contexte et objectif de l'expérience

Nous avons organisé la passation d'un test culturel au cours des Rencontres des Jeunes Chercheurs (RJC) en EIAH le 5 et 6 mai 2010 à Lyon afin d'expérimenter une forme différente de questionnaire. L'idée était de concevoir un test court, composé de deux questions sur deux pages, qui ne nécessite aucun savoir pré acquis. Le temps de remplissage estimé est de 5 minutes au maximum.

La première question présente des photographies de monuments ou de paysages des 4 pays ou régions des doctorants du STEF présents aux RJC : la Tunisie, la Syrie, la Normandie et le

Vietnam. Les répondants doivent attribuer le pays à chaque photographie.

L'objectif du test est de vérifier si la disposition et la taille des images influencent les ordres de réponses. Nous avons disposé les images selon quatre hypothèses :

- 1^e hypothèse : les répondants commenceront par les images en haut à gauche : 3 images ont été placées dans le coin haut gauche.
- 2^e hypothèse : les répondants commenceront par le centre, et ils suivront une disposition circulaire : 5 images ont été disposées au centre de la feuille et de façon circulaire.
- 3^e hypothèse : les répondants seront attirés par une disposition linéaire : En bas, nous avons disposé 5 images linéairement.
- 4^e hypothèse : les répondants commenceront par les grandes images faciles à reconnaître : 3 images de grande taille, ont été disposées à trois extrémités de la page, dont deux images supposées faciles reconnaître et une difficile.

La deuxième question consiste à localiser un lieu touristique sur une carte. Trois essais sont autorisés pour les répondants.

3.3.2. Conception et instrumentation du test

La première page présente l'ensemble des photographies pour la première question. La disposition de la feuille en paysage au lieu du portrait permet de mettre plus d'images et de faciliter l'accès visuel à tous les coins de la feuille. Nous avons disposé les images pour vérifier les hypothèses décrites précédemment. Nous avons opté pour ne mettre que des photographies correspondant à ces 4 pays afin de réduire la difficulté.

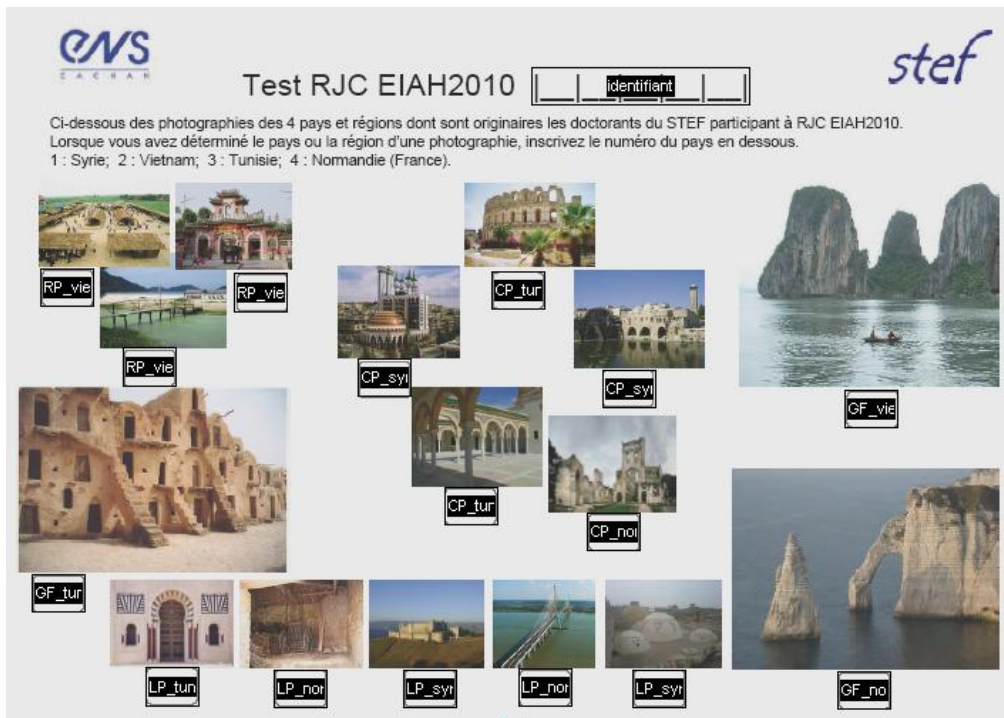


Figure 49 : Instrumentation de la première page du test RJC.

A l'instrumentation, nous avons mis des zones de textes comme champ de réponse. Nous avons paramétré la reconnaissance d'écriture à ces champs pour reconnaître des chiffres. La zone d'identification, qui nous permet d'identifier chaque instance dans le traitement, est composée d'un peigne de 5 caractères alphanumérique.

Nous avons défini une nomenclature spécifique pour les noms des champs groupés selon les hypothèses précédemment décrites. Cette nomenclature nous permet de faciliter l'identification des champs dans la phase de traitement et donc de pouvoir les classer rapidement.

L'idée de la deuxième question est de pouvoir localiser un endroit sur une carte. A la conception, la difficulté pour cet exercice est de pouvoir implicitement contraindre la réponse du répondant sur une zone spécifique. L'idée de placer des capitales dans la carte du monde ou des îles nous a paru très vague et difficile à contrôler depuis l'instrumentation avec les outils Kayentis. Pour cela, nous avons proposé de demander aux participants de localiser un endroit déjà vu dans le premier exercice. Il s'agit de localiser un endroit touristique au sud de la Tunisie.

Nous avons donc disposé une grille de cases à cocher ajustée à la forme de la carte. Cette grille est formée de 25 lignes composées d'un ensemble de cases à cocher chacune (soit une centaine de cases). Chaque case a pour nom la distance en kilomètre qui le sépare dans la carte de l'endroit à trouver, ainsi que la ville de proximité (exemple : 70KmRmeda, la case à cocher est loin du but de 70km et correspond à la ville de Rmeda).

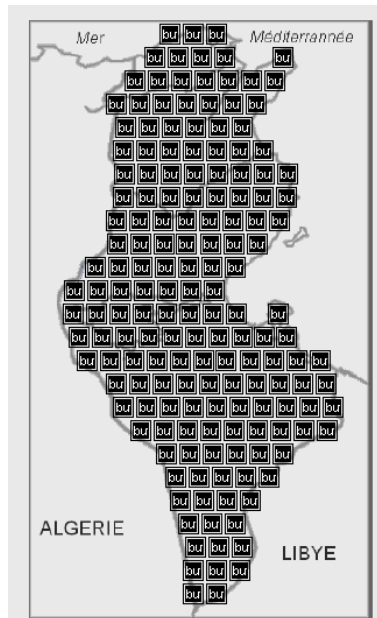


Figure 50 : Instrumentation de la deuxième page du test RJC ; les cases à cocher alignées sur la carte.

3.3.3. Administration du test

Au cours de ces rencontres, 66 personnes ont répondu au questionnaire. Le test, imprimé en couleur, contient deux pages attachées. Chacune correspond à un exercice.

Nous avons utilisé 10 stylos numériques confiés aux doctorants de STEF. Chacun devrait demander aux participants des RJC de répondre à notre questionnaire pendant les pauses. Chaque formulaire a été identifié par un code préservant l'anonymat du répondant.

Après avoir expliqué le but et l'objectif du test, les consignes à donner aux répondants sont les suivantes :

- Les répondants doivent inscrire le numéro du pays dans la case en dessous de chaque image qui lui correspond dans la première question.
- Ils ont le droit de ne pas répondre s'ils ne connaissent pas la réponse.

- Ils n'auront plus la possibilité de revenir sur le premier exercice une fois passé au suivant. En effet, une des images dans la première page, se trouve dans le sud de la Tunisie. La deuxième question consiste à la localiser dans la carte de la Tunisie par un point.
- Ils ont la possibilité de faire trois essais pour la deuxième question.

3.3.4. Traitement des données

Des traitements quantitatifs sont nécessaires à faire pour avoir le taux de réussite par image selon le pays auquel correspond l'image. Pour cela, nous avons introduit une notation sur le test de 1 point si une image est bien identifiée et 0 sinon. Ensuite, nous avons défini un ensemble de traitements par rapport aux objectifs et aux hypothèses posées.

La nomenclature des champs faite à l'instrumentation facilite la classification des réponses par pays et par catégorie. En effet le préfixe du nom de champ représente la catégorie selon les 4 hypothèses posées au début :

- RP : Rassemblement de Petites images en haut
- CP : Centrer de Petites images en forme circulaire au milieu
- GF : Grand Format d'images disposées dans les coins de la page
- LP : disposition Linéaire de Petites images en bas de la page

Le suffixe du nom du champ indique le pays à qui correspond l'image (exemple : RP_vietnam, CP_syrie, etc). Grâce à cette nomenclature, nous avons défini les classifications suivantes :

- « Pays » : avec les catégories Vietnam, Syrie, Tunisie et Normandie.
- Et « disposition » : avec les catégories RP, CP, GF et LP.

Ensuite, nous analysons ces classifications par rapport aux indicateurs temps calculé

automatiquement par l'outil SEGELL à savoir : les instants de la première écriture FS, les instants de la dernière écriture LS et l'ordre de traitement des champs. Ce dernier indicateur temps nous intéresse particulièrement pour cette expérience afin de savoir l'ordre des réponses par rapport à l'indicateur « disposition ».

3.3.5. *Les représentations graphiques*

Le traitement de cette expérience était fait l'après-midi du deuxième jour des RJC à l'aide d'un prototype de l'outil SEGELL qui permet de calculer automatiquement les indicateurs temps comme les durées et l'ordre de traitement. Toutefois, nous avons généré les maquettes des graphiques sur un logiciel tableur.

Nous avons proposé des formes classiques comme l'histogramme, bulle, courbe et barre empilé. Cependant, en utilisant les coordonnées spatiales des champs de réponses récupérées à l'aide du stylo numérique, nous avons pu représenter ces formes sur l'image de la page du test. Chaque valeur est donc représentée sur le champ qui lui correspond dans l'image de la feuille du test. Nous avons utilisé cette technique notamment pour représenter les indicateurs en fonctions du temps qui sont l'ordre de traitement et les parcours chronologiques des réponses. Nous avons gardé des représentations classiques pour l'indicateur de taux de réussite au test.

- **Graphique histogramme pour représenter la répartition des répondants par nombre de bonnes réponses.**

Sur les 16 images de la première question, les 66 répondants se sont globalement bien débrouillés avec ce test culturel. En effet, moins de 10 personnes n'ont pas pu identifier la moitié des images. Plus de 50 personnes ont réussi à identifier la moitié des images.

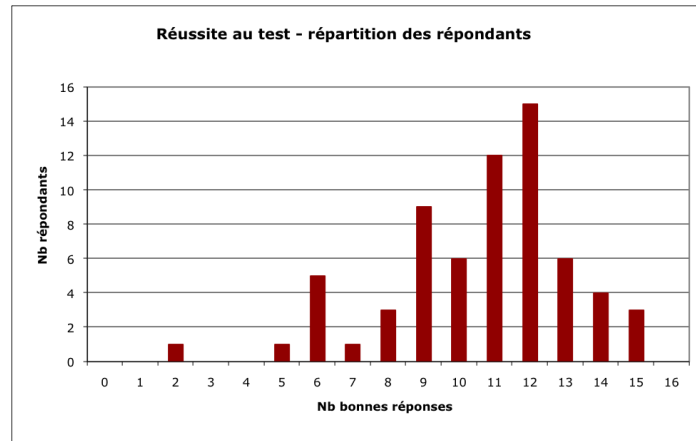


Figure 51 : répartition des répondants par taux de réussite.

- **Graphique en histogramme empilé pour représenter la répartition du taux de réussite par pays.**

La majorité des répondants ont correctement identifié les images correspondant à la Normandie. Le Vietnam, étant le seul pays asiatique, était également plus facile à identifier. Pour la Tunisie et la Syrie, les répondants ont confondu plusieurs images vue les caractéristiques culturelles communes à ces deux pays.

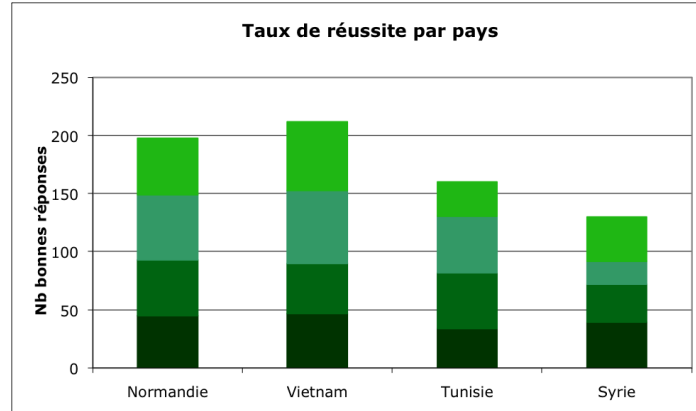


Figure 52 : répartition du taux de réussite par pays.

- **Graphique en bulle pour représenter l'ordre de traitement.**

Dans le but de vérifier les hypothèses posées, nous avons essayé de représenter l'ordre de traitement des images selon la classification « disposition ». Pour cela, nous avons eu besoin de calculer le nombre de fois qu'une image a été identifiée en premier ou en dernier.

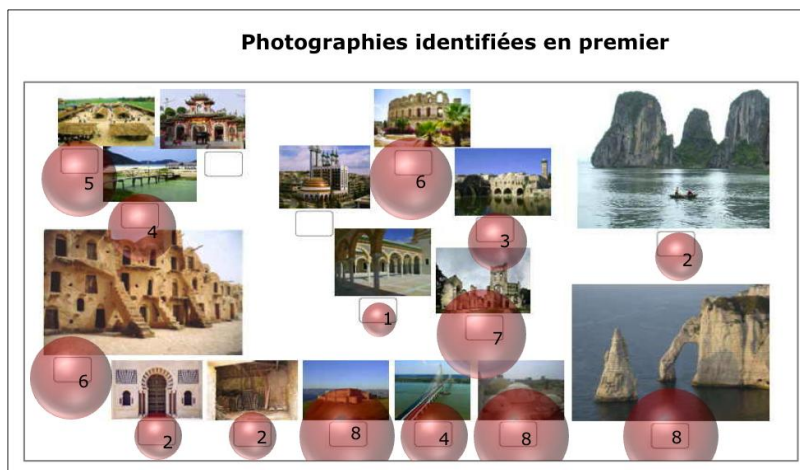


Figure 53 : Pour chaque image, le nombre de répondants l'ayant identifiée en premier.

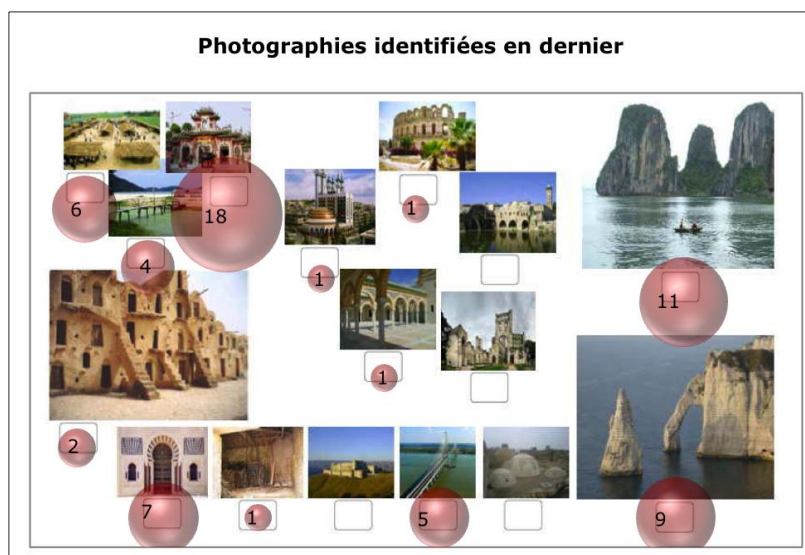


Figure 54 : Pour chaque image, le nombre de répondants l'ayant identifiée en dernier.

Par comparaison des deux graphiques précédents, nous remarquons :

- La première hypothèse n'a pas été tout à fait vérifiée puisque le nombre de fois que les images, rassemblées dans le coin haut de la page, ont été identifiées en premier est moins important que le nombre de fois où ces images ont été identifiées en dernier.
- La deuxième hypothèse a plutôt été vérifiée puisque le nombre de fois où les images, centrées au milieu de la page sous forme circulaire, ont été identifiées en premier est important.

- la 3^e hypothèse a été la plus vérifiée puisque les images, disposées d'une façon linéaire en bas, sont celles les plus identifiées en premier.
- La 4^e hypothèse n'a pas été vérifiée puisque les images en grand, supposées faciles à identifier, ont été identifiées en dernier par les répondants.

Les répondants ont tendance à renseigner les éléments centrés et linéaires malgré la difficulté de l'identification des images disposées ainsi. En effet, ces images marquent la confusion entre les pays Tunisie - Syrie et Syrie - Normandie.

- **Graphique en courbe pour représenter les parcours chronologiques individuels des répondants sur la page du test.**

Nous avons affiché, à l'aide des données temps et des coordonnées spatiales des champs de réponses sur la page, les parcours individuels de certains répondants. Nous distinguons deux parcours particuliers.

Dans le premier parcours, le répondant commence par répondre dans le centre de la page. Ensuite, il change de coin à peu près toutes les deux réponses. Ses réponses ne sont ordonnées ni par pays, ni par zone. Il semble qu'il procède plutôt en cherchant les images qu'il connaît pour les répondre en premier.

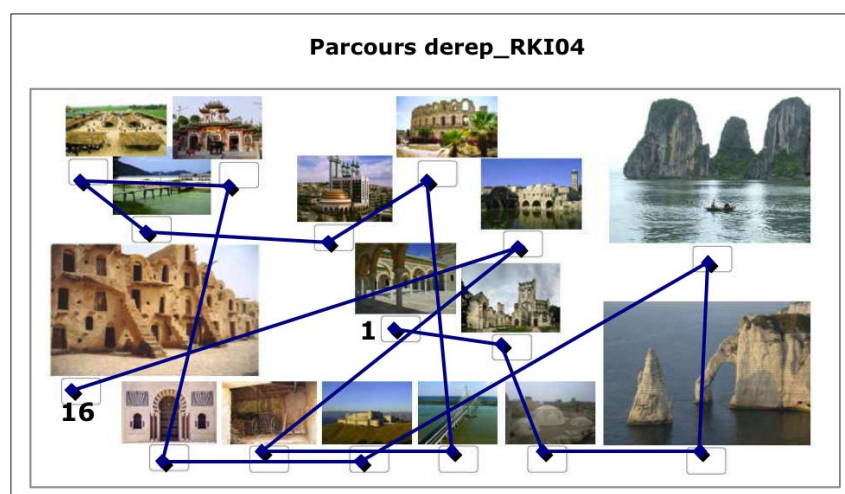


Figure 55 : Parcours du répondant RKI04.

Pour le deuxième parcours, le répondant a commencé par la zone des images linéaires.

Ensuite, il a suivi un parcours en spirale particulier allant des cotés de la page jusqu'à son milieu. On observe une forte influence de la disposition des images dans la feuille sur l'ordre des réponses de ce répondant.

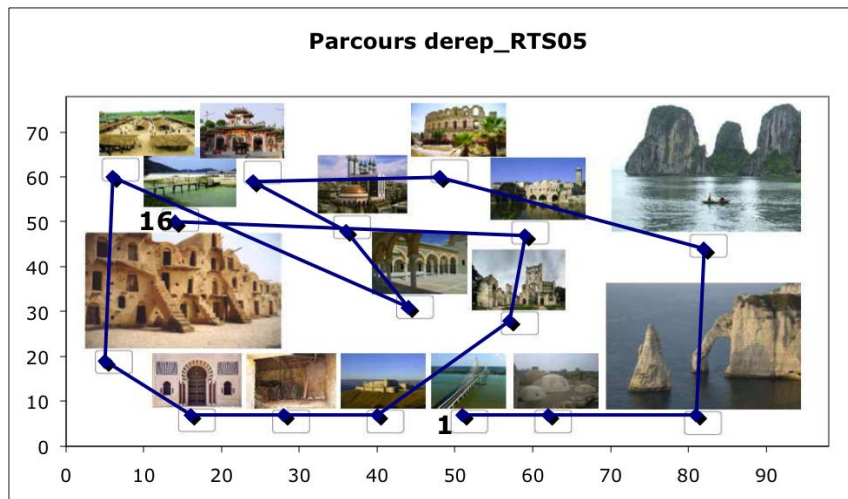
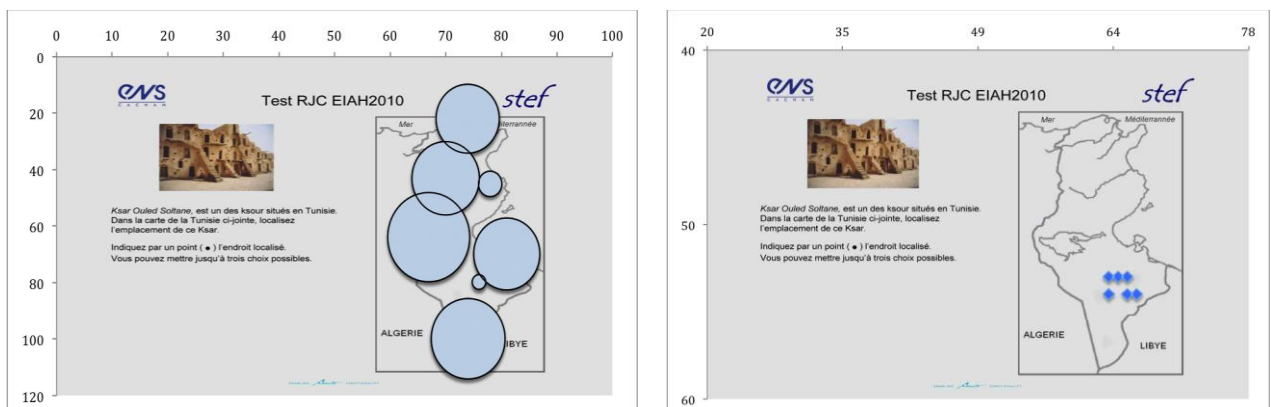


Figure 56 : Parcours du répondant RTS05.

- Un graphique en bulle pour représenter la répartition des réponses à la deuxième question.

Vu le nombre important des champs de réponses pour cet exercice, nous avons regroupé les réponses par zone géographique. Un seul répondant a trouvé la bonne réponse, 7 réponses se situent à moins de 30 km de la bonne réponse.



(a)

(b)

Figure 57 : (a) Répartition des réponses des répondants sur la carte, (b) les bonnes réponses.

3.3.6. Feedback de l'expérience

Le but de cette expérience était de tester un style différent de questionnaire de culture générale. La difficulté mise en place dans ce test consiste en l'identification des images de

monuments de 4 pays disposées dans la page de test selon 4 hypothèses. L'analyse des représentations des parcours chronologiques et de l'ordre de réponses des 66 répondants à ce test nous a permis de vérifier les 4 hypothèses proposées. Dans ce qui suit, nous présentons une analyse de l'expérience en terme d'instrumentation et méthodologie expérimentale, et en terme de traitement et de représentation des données recueillies.

En terme de méthodologie expérimentale, en plus des analyses quantitatives et qualitatives effectuées au cours de cette expérience, nous retenons l'importance de l'organisation de la passation du test. En effet, il est nécessaire de prévoir les consignes adéquates à donner aux répondants afin de recueillir les données souhaitées.

En terme d'instrumentation, le style de test original utilisé dans cette expérience nous a permis de personnaliser l'instrumentation du formulaire afin de faciliter le traitement des données récupérées.

Le traitement de cette expérience a été effectué grâce à un prototype de l'application SEGELL. Nous avons introduit les notes des réponses au test à travers l'interface de ce prototype. Nous avons également défini 2 indicateurs personnalisés qui sont la classification des champs par pays et par la disposition des images dans la page.

La nomenclature des champs définis à l'instrumentation nous a aidé à mieux les identifier en phase de traitement notamment pour les classer par pays et par disposition. Cependant, nous avons introduit à la main ces deux indicateurs à travers l'interface du prototype. Ce type de manipulation devient très vite pénible pour l'utilisateur si le nombre de répondants est important. Ce qui est le cas pour cette expérience.

L'indicateur ordre est calculé automatiquement par le prototype de SEGELL. Cependant, nous avons eu besoin de calculer pour chaque champ le nombre de fois qu'il a été renseigné en premier et le nombre de fois qu'un champ était renseigné en dernier. Ce traitement pourrait être automatisé par le système.

Nous avons exporté les données recueillies aux cours de ce test depuis SEGELL sur un logiciel tableur pour faire des maquettes de représentations graphiques.

Les maquettes des représentations illustrées sur l'image de la page de test nous ont paru intéressantes. En effet, l'analyste peut voir le feedback des résultats directement sur l'image du test. Ce qui favorise une meilleure analyse et une meilleure interprétation. L'idée de représenter les données sur l'image de la page de test ne pourrait être réalisée sans avoir les coordonnées des champs de réponses. Ces données sont collectées uniquement grâce à la technologie du papier et stylo numériques.

3.4. Conclusion

Les résultats de ces courtes expériences sont intéressants du fait que nous avons exploité les données temporelles du stylo numérique sous plusieurs formes comme la durée, l'ordre, la position etc. et nous avons pu représenter ces résultats en fonction des données temps. En effet, ces tests non ordonnés ont permis d'enrichir les fonctionnalités de SEGELL. Nous avons de nouvelles représentations et interactions.

Cependant, chaque expérience est spécifique avec ses objectifs, et donc nécessite un traitement spécifique de l'instrumentation jusqu'à la représentation des résultats. Les données temps recueillies avec le stylo numérique devraient être enrichies et croisées avec d'autres données qui représentent des indicateurs propres à chaque expérience. La détermination de ces données dépend du contexte du questionnaire et de l'objectif de l'enquêteur. Ainsi, ces données peuvent être synchronisées avec d'autres types de média (son, vidéo, grille d'observation, etc.).

Dans le chapitre suivant, nous présentons le cas d'usage pour la prise de note d'observation avec le stylo numérique. Ce cas est différent du fait que la trace recueillie représente l'écriture du chercheur. Les données traitées sont des actions observées et notées par le chercheur. L'analyse a nécessité l'importation d'autres données de transcription pour les synchroniser avec la trace recueillie.

Chapitre 4. L'usage du stylo numérique pour la prise de notes d'observation

4.1. Introduction

Après l'usage du stylo numérique en éducation, récemment, des recherches s'intéressent à l'usage de la technologie pour la prise de note d'observation (Weibel et al., 2011 ; Yeh et al., 2006). L'usage de cette technologie est intéressant pour compléter un processus de recueil de données et pour la synchronisation de ces données avec d'autres média.

En ce qui nous concerne, notre but est l'exploitation des données temps issues de ces enregistrements du stylo numérique pour faire une analyse plus fine par le chercheur. Pour cela, nous nous intéressons à l'usage du stylo numérique pour la prise de note d'observation en classe. Plus précisément, nous exploitons la technologie pour le codage de grilles d'observation en classe.

Dans ce chapitre, nous présentons un autre cas d'usage que nous avons expérimenté en collaboration avec notre collègue Ramia Etche Ogeli dans le cadre de son travail de thèse. Nous avons conçu des grilles d'observation tramées qu'elle a utilisé pour son étude. Les données enregistrées avec le stylo numérique sont ensuite traitées chronologiquement par rapport au codage de la grille et représentées par rapport au temps. Nous avons également traité et synchronisé des données issues de la transcription des enregistrements vidéo à travers des représentations qui permettent d'analyser ces données par rapport à des indicateurs spécifiques à cette étude.

4.2. L'objectif et le contexte de l'expérience

Notre objectif est d'exploiter l'usage du papier et stylo numériques dans le cas de la prise de note d'observation en classe. Nous voulons tester une nouvelle méthode de recueil de données à l'aide de cette nouvelle technologie qui permet, non seulement d'enregistrer des données « métadonnées » de l'écriture manuscrite, mais également de favoriser un traitement plus fin

et une représentation visuelle des données recueillies.

Dans cette perspective, nous avons travaillé sur un cas concret d'observation en classe. Il s'agit d'une étude menée par notre collègue Ramia Etche Ogeli dans le cadre de son projet de thèse intitulé « l'utilisation des TIC par les enfants autistes dans un cadre éducatif ». L'objectif de son étude étant d'étudier les modalités d'utilisation des ordinateurs par des élèves atteints d'autisme, elle vise observer et décrire les différentes formes de communication et d'interactions entre enseignant-enfant-ordinateur.

Plus précisément, son étude vise « à décrire : d'une part, les différentes formes de communication et les comportements répétitifs de l'autiste lors d'une séance informatique ; d'autre part, les actions de l'enseignant qui encourage l'interaction de l'autiste à travers l'ordinateur » (Etche Ogeli, 2010).

A travers des séances d'observation directe en classe dans différents lieux, elle cherche à trouver des indicateurs d'interaction (verbale ou non verbale) entre l'enseignant et l'élève. Les séances d'observations ont eu lieu dans trois institutions différentes : une classe d'intégration scolaire La Plaine à Cachan ; l'institut médico-éducatif de Luis le Guiland à Villejuif ; l'école spécialisée de la Fondation Vallée à Gentilly.

Dans cette expérience, deux instruments d'observation ont été utilisés. Ainsi, l'usage du logiciel Camtasia de TechSmith Corporation¹⁶ pour la capture d'écran permet de garder l'enregistrement de la communication verbale sous forme de son et de l'interaction de l'élève sur l'écran. Ces enregistrements ont été transcrits après à l'aide de l'outil d'annotation ELAN¹⁷.

Le stylo numérique a été utilisé pour plusieurs raisons. Tout d'abord, il n'était pas possible de filmer les enfants observés à l'aide d'une caméra dans certains cas. De plus, le stylo numérique permet d'enregistrer au moment de la prise de note le temps d'écriture. Ce qui permet de synchroniser ces données avec l'enregistrement de la capture d'écran. En outre, la technologie permet de coder la grille d'observation ce qui facilite le traitement automatique

¹⁶ <http://www.techsmith.com/>

¹⁷ <http://www.lat-mpi.eu/tools/elan>

des données et d'avoir des données en formats exploitables.

4.3. L'analyse des besoins

Pour l'observation en classe, le chercheur doit rapporter l'ensemble des actions faites par les élèves et le maître. Le but de ce type d'observation étant d'évaluer soit un élève par rapport à son entourage (l'enseignant et les autres élèves), soit l'enseignant par rapport à sa classe, ou encore d'évaluer un enseignement. Le chercheur pourra donc se focaliser sur un ou plusieurs individus. Pour cela, il sélectionne les éléments intéressants en fonction de son objectif final. Ainsi, on ne peut pas parler d'observation sans parler de grille d'observation.

Une grille d'observation est un outil technique conçu *a priori* ou bien *a posteriori* et qui sert à recueillir des faits selon une procédure de sélection et de codage des éléments à observer d'une façon précise qui répond à l'objectif du chercheur. La grille d'observation peut être, ou pas, l'instrument principal de recueil des données selon la technique d'observation utilisée.

Nous soulignons également l'importance de la considération du facteur temps dans un système d'observation. En effet, quelle que soit la méthode utilisée, le temps intervient dans la phase de recueil de données et dans la phase de traitement et d'analyse de l'enregistrement. Ainsi, la détermination des données temps (intervalle de temps, instant de début et de fin d'une action, durée, etc.) est capable d'enrichir l'analyse.

En terme de méthode ou technique d'observation, dans cette expérience le contexte exige l'usage de l'enregistrement de la capture d'écran de l'interaction de l'élève avec l'ordinateur complété par la prise de note d'observation des actions physiques et des comportements d'interaction faites par l'élève et l'enseignante au cours de la séance d'observation.

D'autre part, les données recueillies à l'aide du stylo numérique, qui sont en format XML, nécessitent un traitement qui facilite l'exploitation de ces données par le chercheur. Ce traitement, qui tient compte les données temps enregistrées, permet au chercheur d'avoir ces données (1) sous forme de tableau xls exploitable facilement sur d'autres outils tableur, et (2) sous formes de représentations graphiques selon des indicateurs spécifiques à la recherche.

Finalement, un autre intérêt de l'usage de la technologie du stylo numérique étant la

possibilité de synchroniser les notes écrites avec les notes de transcription de l'enregistrement vidéo. Pour cela, le chercheur était intéressé par la représentation des deux types de données synchronisés.

4.4. La conception et l'instrumentation des grilles d'observation

Concernant la composition de la grille, nous avons réfléchi à la conception de cette grille (1) en tant qu'interface utilisateur permettant de prendre des notes facilement et instantanément, et (2) en tant que structure numérique permettant l'indexation des notes selon les actions et les individus. Dans cette perspective, nous avons collaboré avec notre collègue Ramia Etche Ogeli pour la conception des grilles d'observation qui ont servi à son expérimentation.

Etant une recherche exploratoire, la conception des grilles est passée par deux étapes. Au cours de chaque étape, les grilles étaient testées et évaluées sur le terrain par le chercheur. Ensuite des améliorations ont été introduites en tenant compte du contexte de la classe et de l'affordance de la grille.

La première étape : elle a permis de déterminer, selon le besoin exprimé par l'observatrice, les acteurs principaux qui sont « l'enfant », « l'enseignante » et « l'ordinateur ». Les actions de chaque acteur vont être notées dans une colonne qui lui correspond (voir Figure 58).

De plus, Nous avons mis deux colonnes pour noter les actions de communication enseignante – enfant et les actions enfant – enseignante à droite de la grille. Nous avons également étalé la grille sur deux pages disposées en paysage puisque cette disposition permet de gagner plus d'espace pour ajouter des lignes.

Notes du 23 octobre 2009

identifiant

<i>Enfant</i>	<i>Enseignant</i>	<i>Machine</i>	<i>Enseignant : Enfant</i>	<i>Enfant : enseignant</i>
enfant0	enseignant0	machine0	ensenf0	enfens0
enfant1	enseignant1	machine1	ensenf1	enfens1
enfant2	enseignant2	machine2	ensenf2	enfens2
enfant3	enseignant3	machine3	ensenf3	enfens3
enfant4	enseignant4	machine4	ensenf4	enfens4
enfant5	enseignant5	machine5	ensenf5	enfens5
enfant6	enseignant6	machine6	ensenf6	enfens6

Figure 58 : un model instrumenté de conception de la grille d'observation.

Cette grille a été testée dans l'institut médico-éducatif de Luis le Guiland à Villejuif. Au cours de la séance, cette disposition était confortable et l'observatrice a pu noter plus vite. Cependant, elle a constaté que plusieurs actions observées se répètent. Ce qui lui a permis de définir les actions qui sont intéressantes pour sa recherche. De ce fait, elle a voulu une amélioration de la grille pour gagner plus de temps.

La deuxième étape : les acteurs et les actions étant précisés, nous avons donc divisé la feuille en trois zones, une pour l'enfant, une pour l'enseignante et une pour noter des commentaires (voir Figure 59).

Note-observatin / IME le Guillard												
Po-Ecra	Regard-M	Regarde-outil	Pointe-outil	Point_ecran	T-souris	Guide_S_avecE	Outil-concret	Commentaire				
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					

Figure 59 : Le model finale de conception de la grille d’observation.

Dans chaque zone, les actions sont classées en colonne. Les actions codées pour la zone enfant sont : pointer l’écran - regarder l’enseignante - regarder outil matériel - pointer outil matériel. Les actions codées pour l’enseignante sont : pointer l’écran - tenir la souris - guider la souris avec l’enfant - utiliser des outils matériels.

Dans la zone commentaire l’observatrice peut noter des remarques diverses ou bien des actions nouvellement observées.

Pour chaque action, deux cases à cochés sont prévues, une pour marquer le début de l’action et une pour marquer la fin de l’action. Il fallait le temps de cocher dans une case au moment de l’observation d’une action pour la noter. Ce qui facilite la tâche de l’observatrice et lui permet de se concentrer plus sur les actions faites par les acteurs, d’une part, et permet de gagner de l’espace dans la feuille, d’autre part. Ce choix de conception a permis de réduire la grille sur une seule page en paysage.

Les champs de la grille ont comme nomenclature la syntaxe suivante : l’acteur_l’action_Debut/Fin_numéro de la ligne. Par exemple : la mention « M_pointe_ecran_debut_1 » désigne le début de l’action « pointer l’écran » faite par l’enseignante dans la première ligne de la grille. Une grille peut contenir près de 340 champs référencés à 4 actions pour l’enfant, 4 actions pour l’enseignante et des commentaires.

Cette grille, qui permet de coder les actions plus facilement, a été utilisée notamment à l’institut médico-éducatif de Luis le Guillard à Villejuif. Cependant, des modifications ont été apportées sur les actions pour son usage dans la classe d’intégration scolaire La Plaine à

Cachan et à l'école spécialisée de la Fondation Vallée à Gentilly, tout en gardant la même conception de la grille.

Note-observatin_fondation								F1 - 27 novembre 2005	
Po-Ecran	Regard-M	Tappe_clavier	L_S	Point_ecran	T-souris	Point_clavier	Guide_S_ave c_E	Commentaire	
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	^>	
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Vidéo m... ms t ho	
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Il veut tou... tappe clavier	
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	acte S pour regarder le Request	
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Il utilise m... de pour s'aper cursor	

Figure 60 : Une grille d'observation remplie.

Pour cette dernière version de grille, les actions pour l'enfant sont : Pointer l'écran – regarder la maîtresse – taper sur le clavier – lâcher la souris, et les actions pour l'enseignante sont : Pointer l'écran – toucher la souris – pointer le clavier – guider la souris avec l'enfant. Ces actions sont adaptées à l'environnement informatique utilisé dans le contexte de ces classes.

Au cours d'une séance d'observation, l'observatrice peut utiliser une ou plusieurs feuilles de grille. A chaque nouvelle feuille elle note dans la zone « identifiant » (voir Figure 60) un code spécifique qui correspond à l'enfant observé et à la date de la séance. Ce codage est nécessaire pour le traitement des notes de chaque séance.

4.5. Le traitement des données

Dans la phase de traitement des données pour cette expérience, il n'y a pas eu usage d'indicateur temps comme dans les expériences précédentes. En effet, les actions d'interaction entre l'enfant et l'enseignante représentent les principaux indicateurs. Pour cela, le codage de la grille a été conçu pour faciliter l'obtention des actions relativement aux acteurs et dans un ordre chronologique. De plus, d'autres données quantitatives, comme le nombre d'occurrences d'une action, peuvent être calculées.

Pour le traitement des données, nous avons classé les champs par catégories d'actions tout en conservant les données temps relatives à chaque note écrite.

Dans le prototype de notre système utilisé pour le traitement des données de cette expérience, une feuille est traitée comme étant une instance de formulaire (une grille dans ce cas). Pour

cela, le chercheur doit spécifier pour le système les instances qui correspondent à une même séance d'observation.

Quoique des indicateurs temps comme la durée d'une action et la latence d'écriture n'étaient pas intéressants dans le contexte de cette expérience, mais les données temps exportées pour le chercheur étaient d'une utilité considérable. Par contre, le chercheur a rencontré certaines difficultés dans la représentation de ces données sur les logiciels tableur et dans leur synchronisation avec les données de transcription.

4.6. Les représentations graphiques

En premier lieu, le chercheur avait besoin des données en tableau qui illustrent les actions et les moments de notation de ces actions. Pour cela, nous avons fourni au chercheur des fichiers xls relatives à ses grilles. Le tableau suivant représente un extrait des données traitées et exportées en format xls pour le chercheur.

Observation : Ousmane 2	
Champ	Temps
M_pointe_ecran_debut_0	21/05/10 14:37:20
M_pointe_ecran_debut_1	21/05/10 14:37:29
M_pointe_ecran_debut_2	21/05/10 14:38:10
M_pointe_ecran_fin_2	21/05/10 14:38:12
E_regarde_M_debu_4	21/05/10 14:38:47
...	...
texte_4	21/05/10 14:39:14

Tableau 1 : un extrait des données traitées et exportées en format xls pour le chercheur.

Ces tableaux ont permis au chercheur de quantifier les actions faites par chaque acteur et de les comparer par rapport aux données de transcription de l'enregistrement vidéo de l'écran.

Nous avons utilisé un prototype de l'outil SEGELL pour la génération des graphiques voulus. En général, les graphiques sont de type points représentant les actions sur un axe de temps. Le temps de début à considérer est un temps relatif c'est à dire que le moment zéro est le moment de début d'écriture sur le papier. Il est possible de faire le défilement et le zoom sur l'axe du temps pour pouvoir visualiser en détail les zones denses. Il est également possible d'enregistrer le graphique généré en image .jpg.

Dans la suite, nous présentons les graphiques produits pour cette expérience. Nous avons pris comme exemple les données d'observation d'un élève de l'institut médico-éducatif de Luis le Guillaud à Villejuif :

- Un graphique qui représente le déroulement des actions dans le temps de chaque enfant durant une séance d'observation. Sur le même graphique, le chercheur voulait voir des données quantitatives qui représentent le nombre d'occurrences des actions notées.

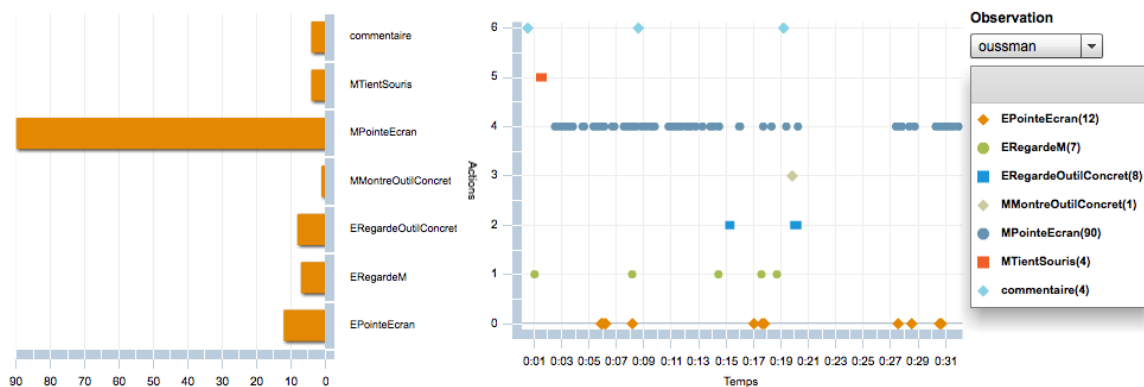


Figure 61 : illustration de la dispersion des actions observées dans le temps à droite, et à gauche le nombre d'occurrence pour chaque action.

Le graphique de droite représente les actions notées par l'observatrice durant une séance d'observation. Chaque action est représentée en point d'une forme et couleur différentes.

La liste de légende à droite illustre les actions notées ainsi que le nombre d'occurrences des actions entre parenthèses. Ce qui permet de savoir plus facilement l'action la plus fréquente durant une séance.

De plus, pour un aperçu visuel, nous avons ajouté une illustration du nombre d'occurrences des actions dans le graphique en barre à gauche.

- Le chercheur a voulu avoir un graphique qui représente le déroulement des actions dans le temps de chaque enfant pendant toutes les séances d'observation dans le but de pouvoir évaluer l'évolution de l'enfant dans chaque séance.

Le problème pour ce type de graphique c'est que les représentations chronologiques des actions de chaque séance seront éloignées (au moins une semaine) dans le temps. Donc la

comparaison des représentations sera difficile visuellement. Pour cela, nous avons proposé d'utiliser le premier graphique qui représente chaque séance à part et puis de les comparer.

Dans la même interface, une liste déroulante permet de choisir le nom de l'observation à illustrer. Nous présentons dans la suite trois représentations de séances d'observation notées successivement pour un seul enfant.

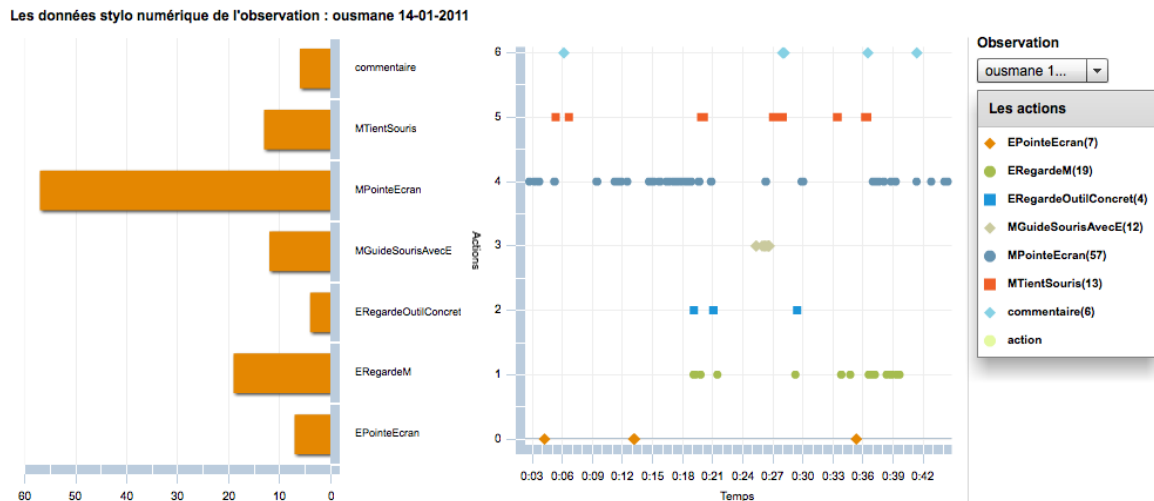


Figure 62 : Représentation de la séance d'observation Ousmane du 14-01-2011.

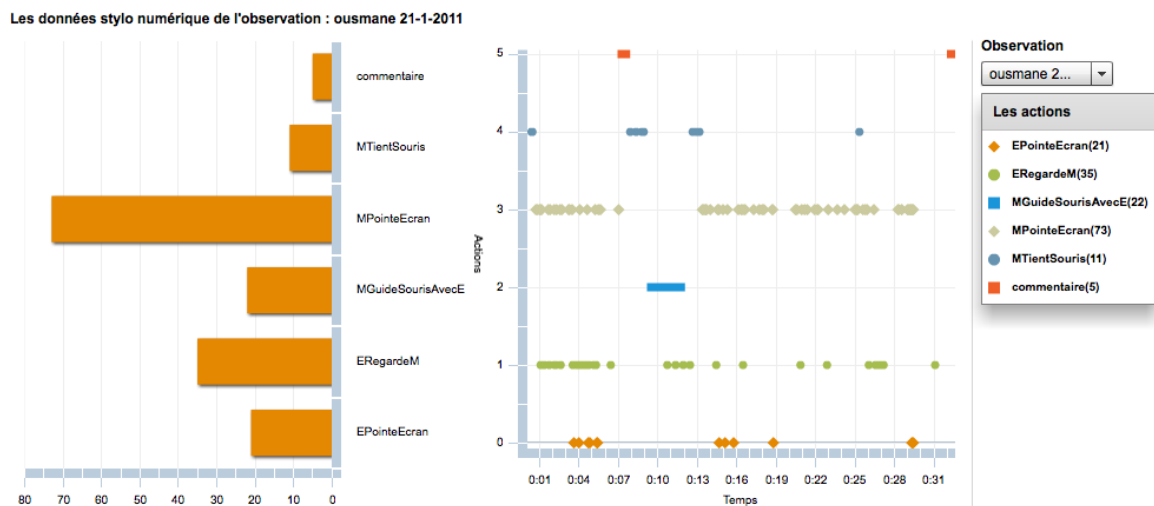


Figure 63 : Représentation de la séance d'observation Ousmane du 21-01-2011.

Les données stylo numérique de l'observation : ousmane 4-02-2011

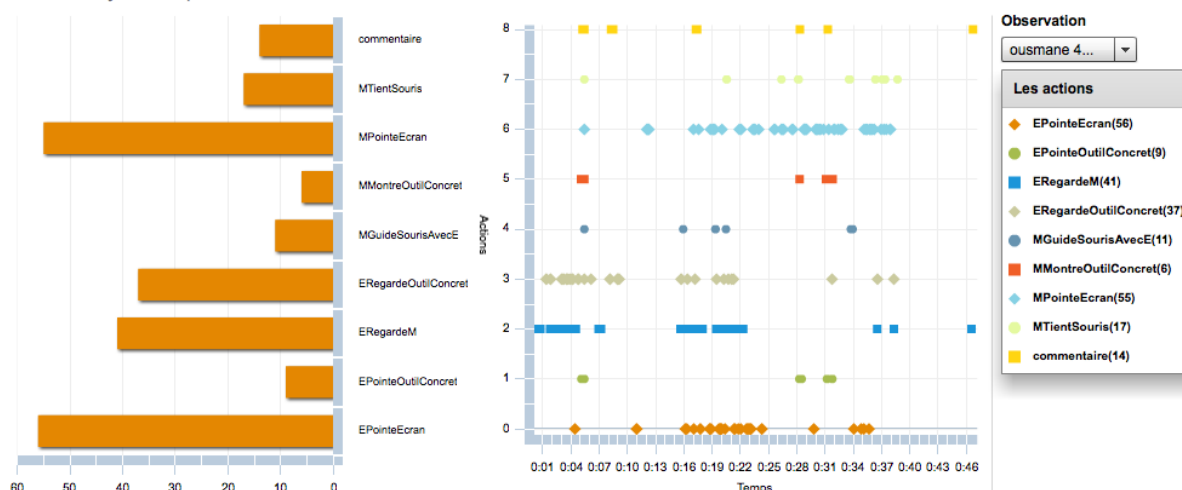


Figure 64 : Représentation de la séance d'observation Ousmane du 04-02-2011.

D'après ces graphiques, nous remarquons que les actions de l'enseignante sont plus nombreuses que celui de l'enfant notamment dans le premier graphique. Dans le deuxième graphique, nous remarquons l'apparition de nouvelles actions faites par l'enfant, mais l'enseignante a effectué plus d'actions que l'enfant. Cependant, dans le dernier graphique, nous remarquons que le nombre d'actions effectuées par l'enfant a dépassé celui des actions effectuées par l'enseignante.

Ces commentaires peuvent être significatifs d'une certaine évolution soit par rapport au niveau de l'enfant et sa familiarisation avec l'ordinateur, soit par rapport à l'interaction entre l'enseignante et l'enfant. Le type d'application utilisé par l'enfant au cours de chaque séance peut également influencer ces interactions. Pour cela, dans son analyse, le chercheur a couplé à ces données stylo numérique les données de transcription de l'enregistrement vidéo.

Ainsi, les besoins exprimés en terme de représentation graphique concernent non seulement les données stylo numérique, mais aussi les données de transcription des enregistrements vidéo. La synchronisation des deux types de notes permet au chercheur de mieux détecter et interpréter les interactions entre l'enfant et l'enseignante.

- Un graphique qui regroupe les deux données de stylo numérique et de transcription pour chaque enfant durant une séance d'observation.

Les données de transcription sont sous forme de fichier texte exporté depuis l'outil

d'annotation ELAN que le chercheur a utilisé. Nous avons importé dans la base de données de l'outil SEGELL les données contenu dans ce fichier qui sont les actions annotées, leur temps de début et de fin, la durée de chaque action et le commentaire noté par le chercheur.

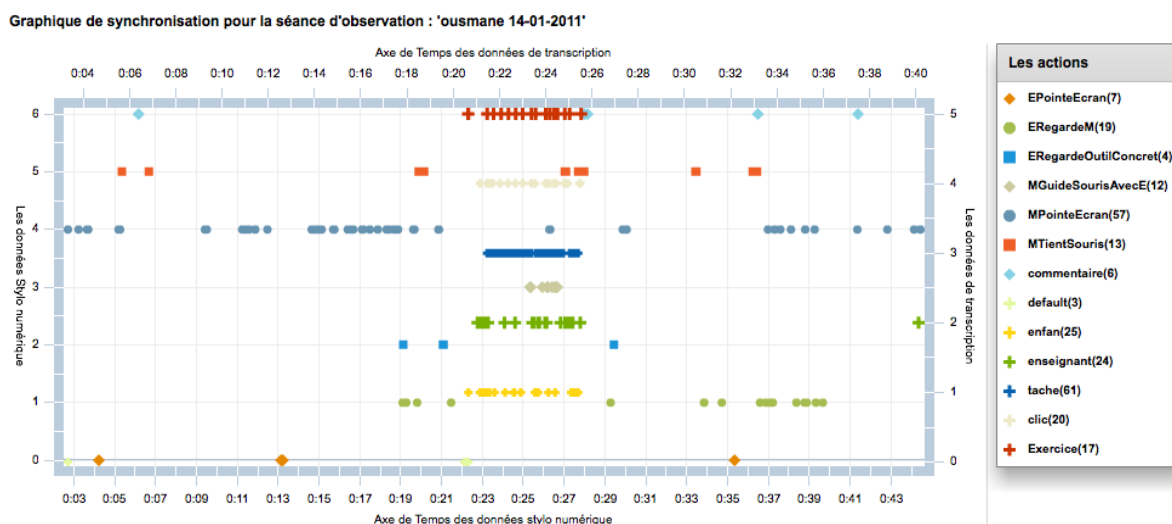


Figure 65 : Représentation des actions, notées par le stylo numérique, synchronisées avec les données de transcription.

Les actions issues des données de transcription sont en forme de croix, et les actions issues des données stylo numérique sont en forme différente. Pour la synchronisation, nous avons vérifié que les temps de début des premières notes pour chaque type de données sont presque les mêmes (quelques secondes de différence).

Nous remarquons d'après cette représentation que l'annotation de la vidéo se concentre essentiellement sur un intervalle de temps au milieu. En fait, nous avons vérifié dans la vidéo correspondante que, sur cet intervalle, l'enfant a commencé une activité différente que celle effectuée au début. C'est peut être un choix du chercheur pour noter les interactions enfant - enseignante au cours de l'usage d'une activité spécifique.

Dans tous les cas, le graphique permet de visualiser en alternance la dispersion dans le temps des actions annotées de la vidéo et les données notées avec le stylo numérique. Il est possible d'effectuer un zoom sur les deux axes de temps.

Graphique de synchronisation pour la séance d'observation : 'ousmane 14-01-2011'

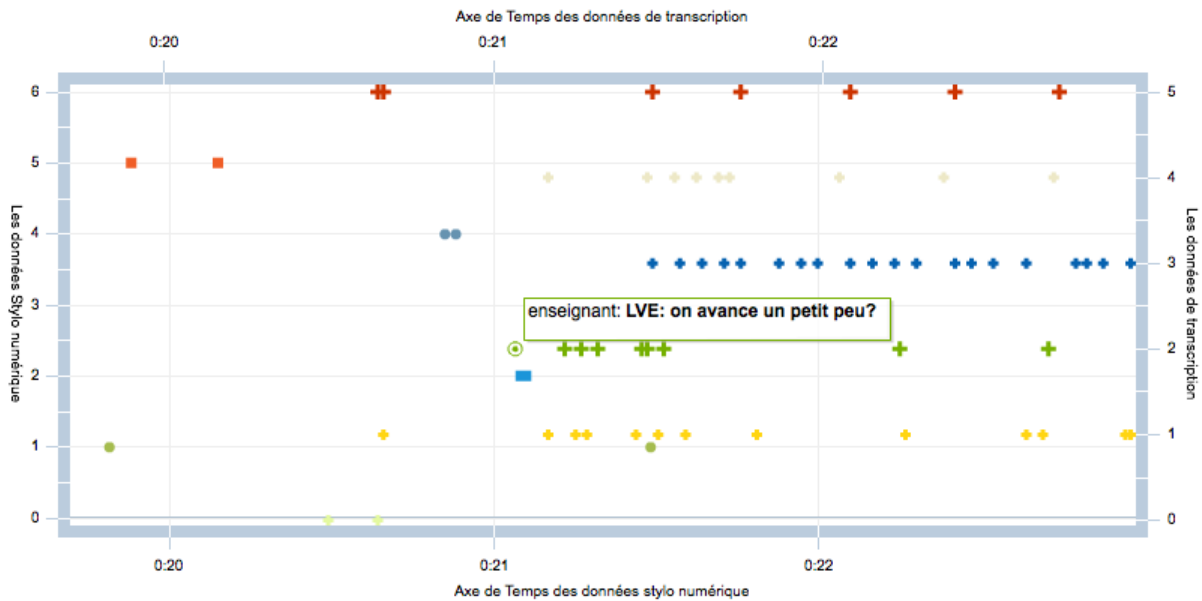


Figure 66 : Illustration d'une partie zoomée de la représentation des données synchronisées.

Par survol de la souris sur un point correspondant à une donnée de transcription, apparaît son nom ainsi que le commentaire noté par le chercheur sur le logiciel d'annotation. L'affichage de cette information peut aider le chercheur à savoir plus sur l'interaction verbale ou non verbale entre les individus par rapport à une action noté à l'aide du stylo numérique et vice versa.

4.7. Feedback de l'expérience

Etant différente des expériences précédentes, l'instrumentation et le traitement des données de cette expérience sont ainsi différents.

En effet, la phase d'instrumentation et de conception des grilles représente une étape de pré analyse pour le chercheur, du fait qu'il doit déterminer le codage de ses notes. De plus, la feuille instrumentée, étant une interface utilisateur, doit être adaptée au contexte de la prise de notes d'observation.

En utilisant la technologie disponible de Kayentis, nous avons disposé des zones de type champ de texte sur une vingtaine de lignes dans la feuille de prise de notes. Dans chaque ligne existent des zones devant chaque action. Ce choix était contraint par le fait que nous ne pouvons pas récupérer le temps de modification de l'écriture dans un champ de texte. Nous

récupérons uniquement l'instant de début et l'instant de fin d'écriture.

En conséquence, nous avons un nombre très élevé de champs instrumentés. Ce qui constitue une lourdeur pendant l'instrumentation et au moment du traitement. D'ailleurs, nous avons dû regrouper les champs en « Actions » à travers l'interface du prototype.

De plus, la tâche de regroupement des instances d'une séance d'observation à travers l'interface du prototype pourrait être évitée si l'instrumentation des grilles s'étale sur le nombre de feuilles nécessaire pour une séance d'observation. Ce qui permettra d'avoir une instance pour une séance d'observation.

Nous avons observé que le chercheur a utilisé en moyenne entre 4 et 5 feuilles par séance d'observation. Ce chiffre pourrait être fixé selon la durée et le rythme de déroulement d'une séance d'observation.

Concernant l'interface du prototype, des fonctionnalités comparables à celles produites pour les autres expériences étaient utiles dans ce contexte comme par exemple l'importation des fichiers des données, la visualisation des formulaires et des instances, et la personnalisation des libellés pour les champs représentant des actions. Cependant, nous avons ajouté la fonctionnalité de regroupement des instances en observation. C'est à dire, définir pour une séance d'observation les instances (qui représentent des feuilles de grilles) qui lui correspondent. Ce regroupement était nécessaire pour la génération des graphiques souhaités.

Pour la synchronisation des données de stylo numérique avec des données issues de la transcription des vidéos, nous avons ajouté la possibilité d'importer des fichiers autres que les données XML du stylo numérique et contenant des données temps. Pour cela, nous avons effectué des modifications sur la base des données. Cependant, une fonctionnalité de formatage des données temps importées de différentes sources est nécessaire afin d'éviter l'extension de la base de données.

Pour l'interactivité des graphiques générés, nous avons utilisé une technique permettant de zoomer et de faire défiler les axes de temps. Cela permet au chercheur de visualiser les zones représentant des données denses. De plus, nous avons utilisé les techniques de survole avec la souris sur les éléments du graphique afin d'afficher plus d'information sur les données représentées.

4.8. Conclusion

Comme l'utilisateur du stylo numérique est le chercheur lui-même, nous nous sommes concentré sur le traitement et la visualisation des noms des champs et le temps d'écriture au lieu de nous concentrer sur le comportement du transcritteur. La représentation la plus utile dans ce contexte est la dispersion des actions dans le temps. De même, l'instrumentation des grilles d'observation est d'une utilité considérable, à condition de déterminer un codage adapté au contexte de l'expérience au préalable.

Le prototype développé pour ce cas d'usage a permis de récupérer et d'analyser des données de notes d'observation en classe. Les fonctionnalités, les représentations et les interactions dans ce prototype sont développées selon l'analyse des besoins du chercheur. Toutefois, d'autres fonctionnalités restent à faire afin d'améliorer l'utilité de ce produit.

Dans le chapitre suivant, nous présentons le prototype de la plateforme SEGELL en terme d'implémentation et spécifications techniques. Nous analysons l'interface d'utilisation de ce prototype selon plusieurs facteurs.

Chapitre 5. L'implémentation et l'analyse de la plateforme SEGELL

5.1. Introduction

Dans ce chapitre nous décrivons ce que nous avons réalisé en terme de développement informatique. Dans le cadre d'une conception participative, nous avons réalisé quatre versions du prototype de la plateforme SEGELL. En effet, pour chaque version nous nous sommes focalisé essentiellement sur le besoin de l'utilisateur ainsi que sur ses habitudes de traitement des données. Au fur et à mesure, des fonctionnalités de traitement et de représentation ont été généralisées et améliorées d'une version de prototype à l'autre. Nous décrivons donc les spécifications techniques et fonctionnelles de l'outil ainsi que la synthèse de l'analyse des interactions dans l'interface utilisateur.

5.2. Spécification fonctionnelle

Nous voulons concevoir une plateforme d'aide au traitement et à la représentation des données traces d'écriture enregistrées avec un stylo numérique. La solution est destinée aux chercheurs en sciences de l'éducation désirant exploiter des données recueillies avec un stylo numérique dans le cadre d'une expérimentation. L'analyse concerne les données quantitatives et les données qualitatives notamment les données temps que contiennent ces traces.

5.2.1. *Le but*

Le but de cette plateforme est double. Il doit instrumenter la visualisation de la trace d'écriture avec un stylo numérique, mais également offrir un ensemble de traitements automatiques et permettre un ensemble de manipulations de données pour définir des indicateurs spécifiques en fonction du choix de l'utilisateur.

Cet outil doit pouvoir traiter des formulaires instrumentés avec la technologie Anoto et ne pas dépendre d'une conception spécifique. Pour cela, nous devons prévoir un paramétrage raffiné pour sélectionner les données à traiter. Ce paramétrage tiendra compte du contexte d'usage et du domaine d'expertise de l'utilisateur. Ainsi, ces données peuvent être synchronisées avec

d'autres types de média (son, vidéo, grille d'observation). Notre problématique sera de concevoir un modèle général pour gérer ces données.

5.2.2. *Le principe de la solution*

La solution consiste à importer la trace dans une base de données relationnelle et à transformer automatiquement les données temps en indicateurs. L'utilisateur peut visualiser et gérer les données importées et peut également introduire des indicateurs spécifiques à son expérience. L'interface utilisateur permet de paramétrer la représentation tabulaire et graphique des données. A terme, pour enrichir ces analyses, il sera possible de synchroniser ces données avec d'autres données.

5.2.3. *Les processus de fonctionnement*

Comme nous venons de le décrire plus haut, la conception de la plateforme SEGELL est passée par plusieurs phases de prototypage. A chaque phase, les fonctionnalités et les traitements indispensables se sont dégagés peu à peu.

A l'aide d'un programme web service, Les données sont exportées depuis la plateforme *Forms2.0* de Kayentis. Ce programme permet de spécifier le nom du formulaire, la date, la version et d'autres paramètres d'exportation. Le dossier zippé exporté contient pour chaque formulaire deux fichiers XML, un qui contient la description de la structure du formulaire et un autre qui contient les données trace d'écriture.

Le fonctionnement de l'analyse à l'aide de SEGELL commence par l'importation du fichier de la structure d'un formulaire et le fichier des instances de réponses. Un formulaire sur SEGELL représente un questionnaire, une enquête ou une grille d'observation. Le but de l'analyste est de produire des tableaux de données éditables et des graphiques interactifs.

Les principaux processus de l'application sont :

- l'importation de la trace,
- le traitement des données temps,
- la gestion des données,
- la visualisation des données et l'édition des données.

Nous distinguons dans la Figure 67 les processus automatiques et les processus nécessitant une interaction avec l'utilisateur.

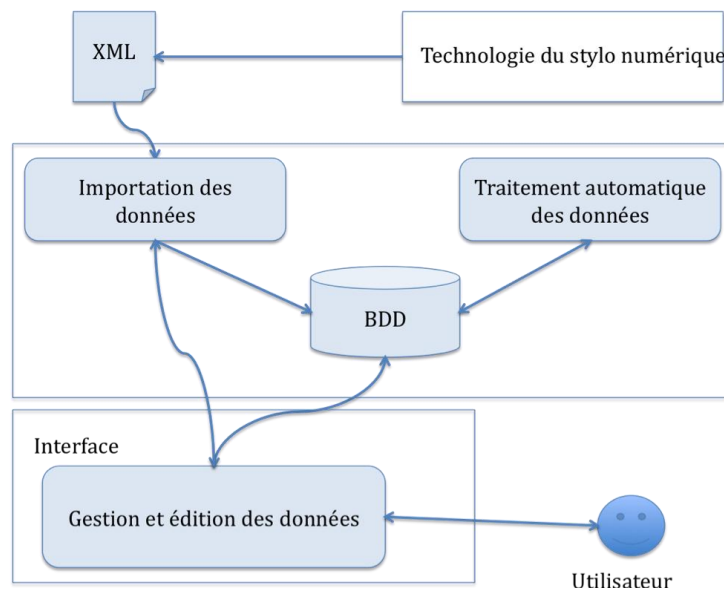


Figure 67 : le fonctionnement global de SEGELL.

Pour le choix de la technologie du langage de développement nous avons opté pour développer en php et MySQL pour la base de données. Des langages auxquels nous sommes déjà familiarisé et qui sont adaptés pour développer une plateforme en ligne, une application web.

Dans la section suivante, nous décrivons un scénario d'utilisation de la plateforme SEGELL. Les fonctionnalités de l'interface d'utilisation sont illustrées à travers cette description.

5.3. Scénario d'utilisation

Pour utiliser la plateforme SEGELL, le chercheur, muni d'un login et d'un mot de passe, doit **s'authentifier** (voir Figure 68). Les comptes utilisateurs sont pour le moment affectés par une session d'administrateur. Le login est composé des initiales des utilisateurs. Pour chaque connexion, l'utilisateur n'a accès qu'aux formulaires chargés dans sa session.

→ Fonctionnalité d'authentification.

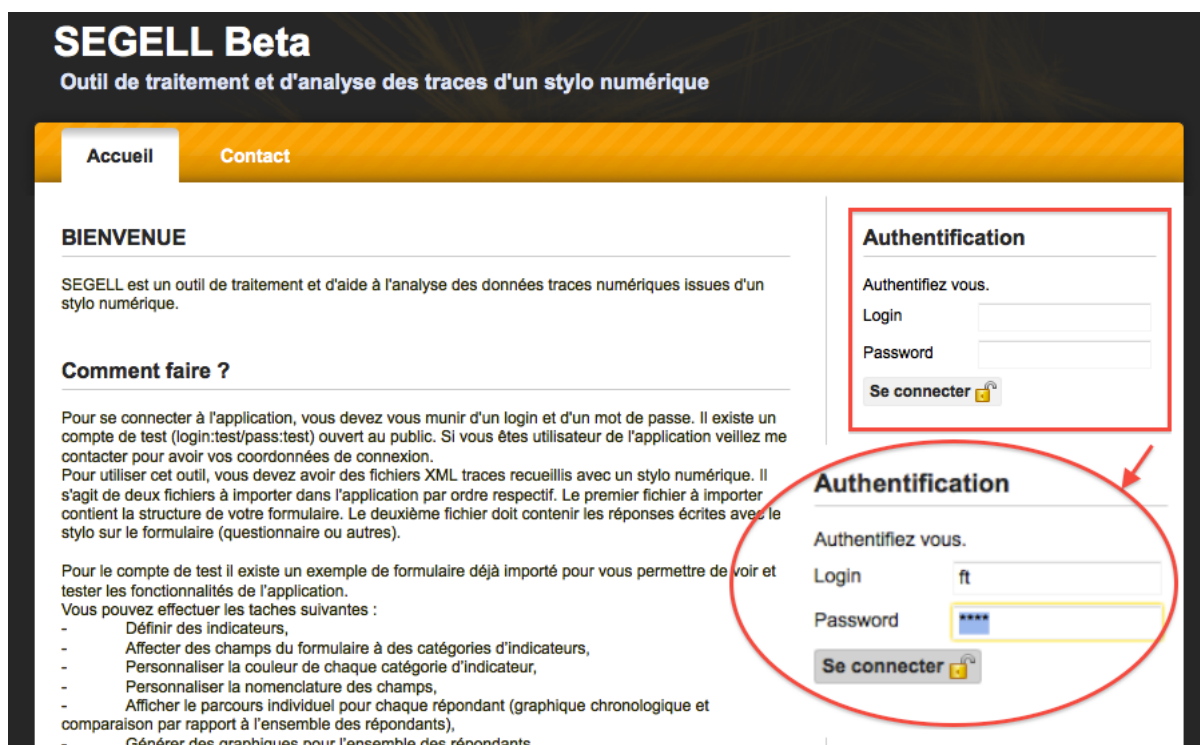


Figure 68 : Interface de Bienvenue de SEGELL et zone d'authentification.

Après l'authentification, l'interface présente deux nouveaux onglets en haut, qui sont l'onglet Formulaire et l'onglet Visualisation.

La première phase consiste à **importer les données** depuis les deux fichiers XML dans la base de données à partir de l'interface de SEGELL. L'utilisateur doit cliquer sur l'onglet **Formulaire** dans l'interface pour le faire.

Dans l'écran qui apparaît (voir Figure 69), l'utilisateur trouve par défaut l'écran de **gestion de formulaire**. Dans la **barre de navigation** qui se trouve à gauche de l'écran (voir Figure 69), le menu Formulaire est grisé. En bas de cette barre, se trouve la **liste des formulaires déjà chargés par l'utilisateur** connecté.

→ Fonctionnalité d'importation des données.

→ Fonctionnalité de gestion de formulaire.

→ Visualisation : menu de navigation contextuelle.

→ Visualisation : liste des formulaires déjà chargés par l'utilisateur.

Au milieu de l'écran, se trouve l'interface qui permet d'importer les fichiers XML. **Un ou plusieurs fichiers** peuvent être ajoutés à la **liste de chargement**. Pour chaque fichier, l'utilisateur doit spécifier **le type s'il s'agit d'un fichier de structure de formulaire ou bien d'un fichier de trace d'écriture (réponses)**. Dans le cas de données de notes d'observation, il y a un type de fichier de données de plus qui est Transcription (voir Figure 70).

→ Interaction : parcourir pour ajouter un ou plusieurs fichiers.

→ Interaction : liste déroulante statique pour choisir le type de fichier.

→ Visualisation : affichage de la liste des fichiers sélectionnés.

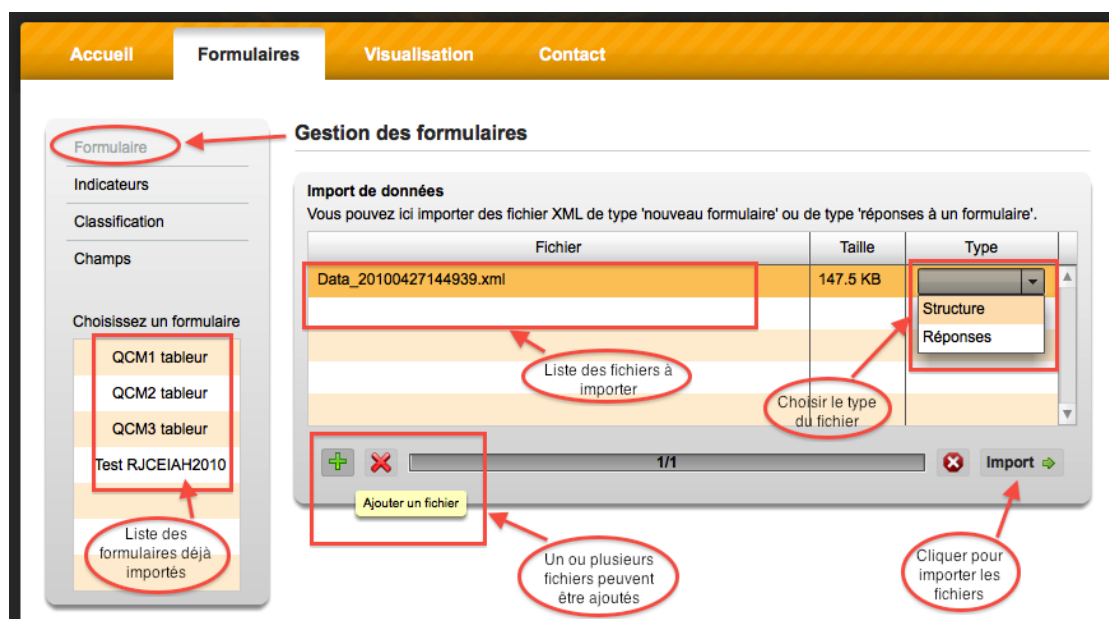


Figure 69 : L'écran d'importation des formulaires.

Après l'importation, le formulaire **s'ajoute** dans la **liste des formulaires** au nom de l'utilisateur connecté.

→ Visualisation : confirmation d'ajout du nouveau formulaire.

→ Fonctionnalité de gestion d'indicateurs.

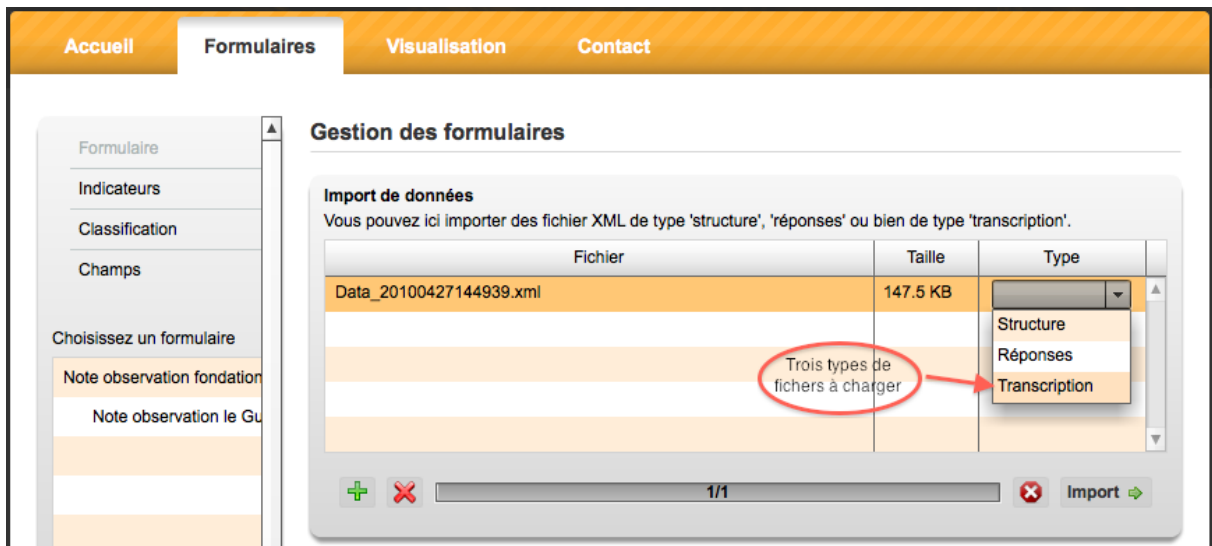


Figure 70 : L'écran d'importation de données de transcription pour notes d'observation.

Pour passer à l'écran de **gestion d'indicateurs**, l'utilisateur doit cliquer sur le menu Indicateurs à gauche de l'écran.

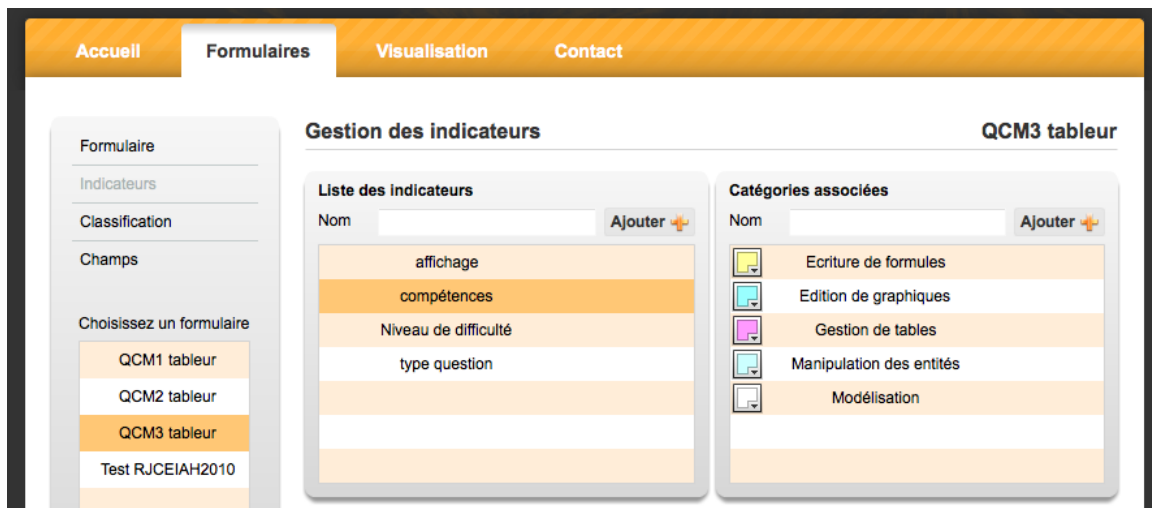


Figure 71 : Ecran de gestion d'indicateurs.

Cet écran permet à l'utilisateur **d'ajouter les indicateurs** personnalisés pour son étude (voir Figure 71). Après avoir choisi un formulaire parmi la liste disposé à gauche en bas de l'écran, une zone apparaît au milieu de l'écran. A partir de cette zone, l'utilisateur peut ajouter un nouvel indicateur. Les indicateurs correspondants au formulaire sélectionné **apparaissent dans une liste**.

→ Fonctionnalité : ajouter/ supprimer un indicateur.

→ Visualisation : liste d'indicateurs existants.

Lorsque l'utilisateur clique sur un indicateur de cette liste, une autre zone spécifique pour **ajouter les catégories** qui correspondent à cet indicateur apparaît à droite. Pour chaque catégorie ajoutée, l'utilisateur peut **choisir une couleur depuis une palette de sélecteur de couleurs**. Cette couleur va paraître dans les graphiques que l'utilisateur va générer dans l'interface de Visualisation.

→ Interaction : sélection d'indicateur.

→ Fonctionnalité : ajouter/ supprimer de catégorie.

→ Manipulation : Saisir le nom d'une catégorie.

→ Manipulation : Sélection de couleur pour chaque catégorie.

Les indicateurs que l'utilisateur ajoute, vont lui servir pour **classer les champs de réponse**. Ce qui permet par la suite d'afficher les données selon une classification et de pouvoir **analyser ou comparer différentes catégories de champs**.

→ Analyse : classer les champs de réponse pour comparer les valeurs de différentes catégories de champs.

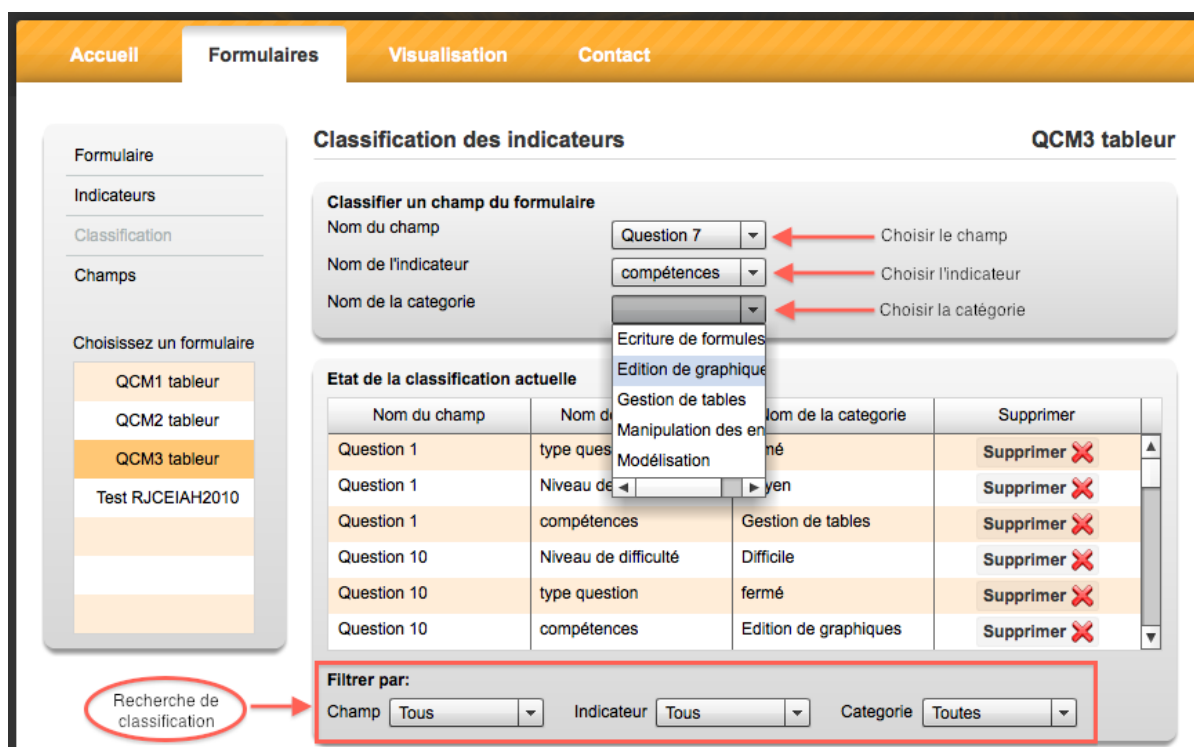


Figure 72 : Ecran des classifications des champs par indicateurs.

L'affectation des champs aux indicateurs définis se fait à partir du menu Classification à gauche de l'écran. Ce qui permet de passer à l'écran suivant de classification des champs (voir Figure 72).

→ Manipulation : Affectation des champs.

Finalement, à partir du menu Champs à gauche de l'écran, l'utilisateur peut accéder à l'écran de **gestion des champs** (voir Figure 73).

→ Fonctionnalité : gestion des champs.

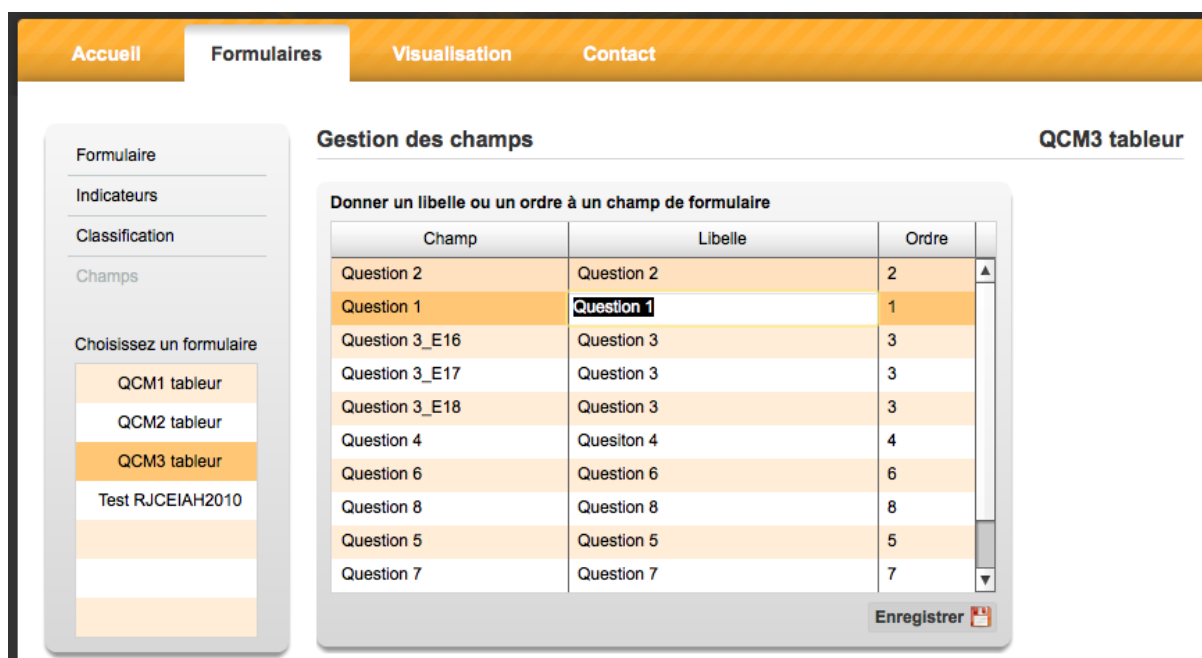


Figure 73 : Ecran de gestion des champs.

L'écran de gestion des champs permet à l'utilisateur de **définir deux critères** pour les champs qui vont permettre de mieux **organiser** l'affichage des représentations des données. En effet, les valeurs nom du champ importé depuis les fichiers XML sont ceux définis à la phase d'instrumentation. Cet écran permet de redonner **un nom de libellé** à chaque champ et de lui **affecter un ordre** selon lequel il sera affiché dans les représentations. La fonction de libellé peut servir également pour **regrouper des champs** sous un seul nom.

→ Manipulation : définir un nom de libellé à chaque champ.

→ Manipulation : affecter un ordre.

→ Analyse : regrouper les champs.

Dans le cas de traitement des données de notes d'observation, cette interface permet de classer les champs en « Action ». Ces actions, défini par le chercheur au préalable, représentent des indicateurs d'analyse (voir Figure 74).

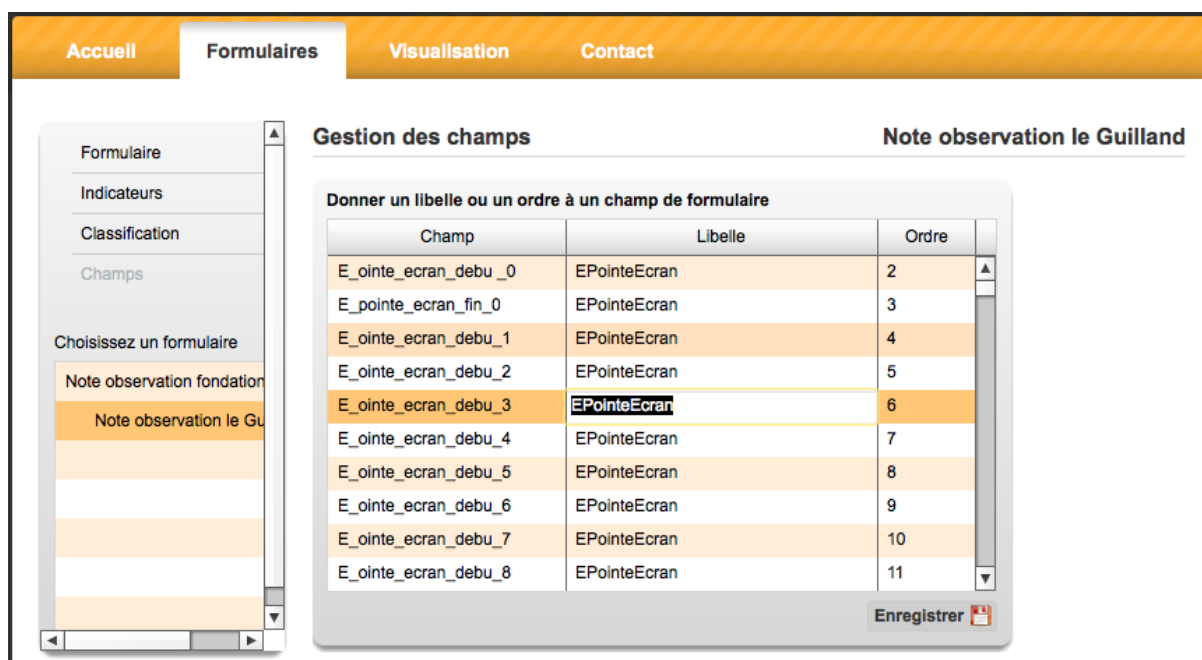


Figure 74 : Ecran de gestion des champs (cas de formulaire de notes d'observation).

A ce stade, l'utilisateur a personnalisé les données qui correspondent à la structure d'un formulaire. Il peut passer à la phase de **visualisation** des résultats des réponses au test. Pour cela, il doit cliquer sur l'onglet Visualisation dans la barre des menus en haut.

L'interface de visualisation permet de sélectionner un formulaire puis de **visualiser la liste des répondants** (instances du formulaire remplis avec le stylo numérique) qui ont participé à ce test. L'utilisateur peut visualiser les données recueillies soit par **individu** soit pour le **groupe de répondants** à ce formulaire. Pour chaque visualisation, nous avons des représentations **tabulaires** et des représentations **graphiques**.

→ Visualisation : afficher la liste des formulaires.

→ Interaction : sélection d'un formulaire.

→ Manipulation : cliquer pour générer des représentations graphiques/tabulaires.

Pour **visualiser les données qui correspondent à un individu**, l'utilisateur doit choisir un dans la **liste des répondants** au milieu de l'écran (voir Figure 75). Plus bas, deux boutons vont être affichés, un pour **visualiser une représentation tabulaire des réponses de**

l'individu et un autre pour **visualiser une représentation graphique** du déroulement chronologique du test.



Figure 75 : Ecran de visualisation des données.

- Visualisation : Afficher la liste des répondants/instances.
- Interaction : Sélection d'un individu.
- Manipulation : Cliquer pour générer des représentations graphiques/tabulaires d'individu.
- Fonctionnalité : Générer des représentations graphiques/tabulaires d'un individu.
- Manipulation : Cliquer pour supprimer un individu/instance.

Pour visualiser toutes les réponses d'un individu, l'utilisateur doit cliquer sur le bouton 'Tableau de Réponses'. Une fenêtre de type pop-up va être affichée sur l'écran principal. Cette fenêtre contient un **tableau de données** (voir Figure 76) qui contient les noms des champs dans lesquels le répondant a écrit, la réponse écrite, le temps de début d'écriture, le temps de fin d'écriture, la durée d'écriture, la durée de nom écriture, la durée totale et la note. Pour cette dernière donnée l'utilisateur doit **l'affecter par lui même** dans ce tableau puis il doit **cliquer sur le bouton 'Enregistrer notes'**.

Tableau réponses par individu							
Répondant: BADU							Formulaire: QCM1 tableur
Champ	Réponse	Début d'écriture	Fin d'écriture	Durée totale	Durée E	Durée NE	Note
Question 1	formuleA6plus1	2009-10-05 15:21:46	2009-10-05 15:21:48	00:00:02	00:00:02	00:00:00	0
Question 2	format/cellule/nbre	2009-10-05 15:22:27	2009-10-05 15:22:29	00:00:02	00:00:02	00:00:39	1
Question 3	E6x(1-F\$2)	2009-10-05 15:23:36	2009-10-05 15:24:30	00:02:01	00:00:54	00:01:07	1
Question 4	A3xB3	2009-10-05 15:25:32	2009-10-05 15:25:51	00:01:21	00:00:19	00:01:02	1
Question 5	largeurcellule	2009-10-05 15:26:26	2009-10-05 15:26:27	00:00:01	00:00:01	00:00:35	1
Question 6	SOMME(E6:E10)	2009-10-05 15:27:28	2009-10-05 15:27:52	00:01:25	00:00:24	00:01:01	1
Question 7	fraction	2009-10-05 15:29:09	2009-10-05 15:29:10	00:01:18	00:00:01	00:01:17	0
Question 8	"le montant à payer est:"8	2009-10-05 15:29:47	2009-10-05 15:33:06	00:01:47	00:01:10	00:00:37	1
Question 9	1fois	2009-10-05 15:30:36	2009-10-05 15:30:41	00:00:05	00:00:05	00:00:49	1
Question 10	SI(E6>35000,E6x(1-F\$2);E	2009-10-05 15:31:00	2009-10-05 15:31:56	00:01:15	00:00:56	00:00:19	1

Figure 76 : Fenêtre de tableau de données de réponses d'un individu.

Ce tableau est **exportable en format xls** par clic sur le bouton 'Exporter' en bas de l'écran.

→ Visualisation : Afficher un tableau des données pour un individu.

→ Manipulation : Affectation d'une note.

→ Manipulation : Cliquer pour enregistrer une note.

→ Fonctionnalité : Edition des données en format xls.

→ Manipulation : Cliquer pour exporter les données en format xls.

Pour **visualiser la représentation graphique du déroulement chronologique du test**, l'utilisateur doit **cliquer sur le bouton 'Graphique chronologique'**. Une fenêtre pop-up s'affiche sur l'écran (voir Figure 77).

→ Visualisation : Afficher la représentation graphique du déroulement chronologique du test.

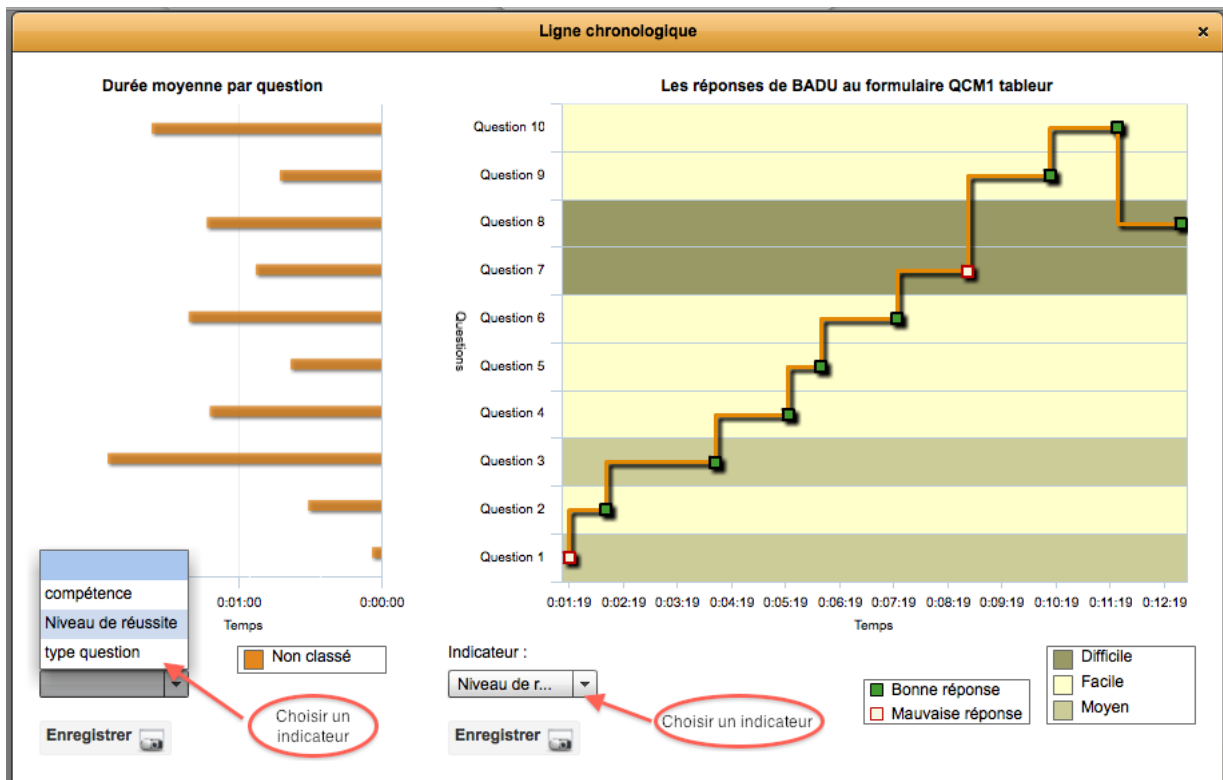


Figure 77 : Fenêtre de représentation graphique composé et paramétrable.

Cette fenêtre contient un **graphique** chronologique qui **représente le processus de résolution du test par le répondant**, à droite, et un **graphique histogramme** qui représente **la durée moyenne passée par tout le groupe de répondants sur chaque question**, à gauche. Ce dernier graphique sert à **comparer le temps passé par le répondant sur chaque question par rapport à la durée moyenne pour tout le groupe**.

→ Analyse : processus de résolution du test d'un répondant.

Le graphique à droite est de type courbe avec marques personnalisées (form = « reverseStep » en langage Flex signifie dessiner les lignes entre les points comme des segments horizontaux et verticaux). L'axe horizontal est de type temps. Les valeurs temps (LS) seront affichés dans l'axe des X. l'axe des Y comprend les Identifiants des champs des réponses, mais les noms des champs seront **affichés dans cet axe comme libellé**. Les points du graphique sont colorés **selon les notes pour chaque champ**.

→ Analyse : durée passé par le répondant sur une question.

→ Analyse : le résultat de la réponse (bonne ou bien mauvaise réponse).

→ Analyse : instant de début et instant de fin d'écriture d'une réponse.

→ Visualisation : graphique en courbe avec marques personnalisées.

→ Analyse : durée moyenne passé par tout le groupe de répondants sur chaque question.

→ Visualisation : graphique en histogramme en barre.

→ Analyse : Comparer le temps passé par le répondant sur chaque question / à la durée moyenne pour tout le groupe.

En bas, se trouvent **deux listes déroulantes** qui permettent de **choisir un indicateur** dans chaque graphique. Ces indicateurs auxquels sont affectés les champs de réponses vont **changer l'apparence des graphiques**.

→ Interaction : Listes déroulantes dynamiques d'indicateurs.

→ Manipulation : Choisir un indicateur.

Dans le graphique chronologique par exemple, le choix d'indicateur va se manifester dans le **changement de la couleur du fond du graphique devant chaque champ**. Ce qui ajoute **une nouvelle valeur affichée** dans ce graphique, en plus des valeurs de **données temps** et **résultat de réponse**. Les couleurs des barres devant chaque valeur de champ dans l'autre graphique changeront selon la catégorie à laquelle le champ appartient. Ces nouvelles valeurs qui **colorent les graphiques** permettent **d'enrichir l'analyse par croisement et comparaison des indicateurs**.

→ Visualisation : Changement de la couleur du fond de graphique devant chaque champ de réponse.

→ Analyse : Croisement et comparaisons des données.

→ Visualisation : Afficher les légendes des valeurs représentées dans les graphiques.

L'utilisateur peut **exporter ces graphiques en format jpg** en **cliquant sur le bouton 'Enregistrer'**.

→ Fonctionnalité : Edition des graphiques en format jpg.

→ Manipulation : Cliquer pour exporter les graphiques en format jpg.

Nous venons à la **visualisation de l'ensemble des données** qui correspondent à un formulaire chargé. Il existe deux types de représentations, une tabulaire et une graphique.

Pour la **représentation tabulaire**, l'utilisateur doit **cliquer sur le bouton 'Génération de tableaux'** à gauche de l'écran de visualisation pour afficher une fenêtre pop-up (voir Figure 78).

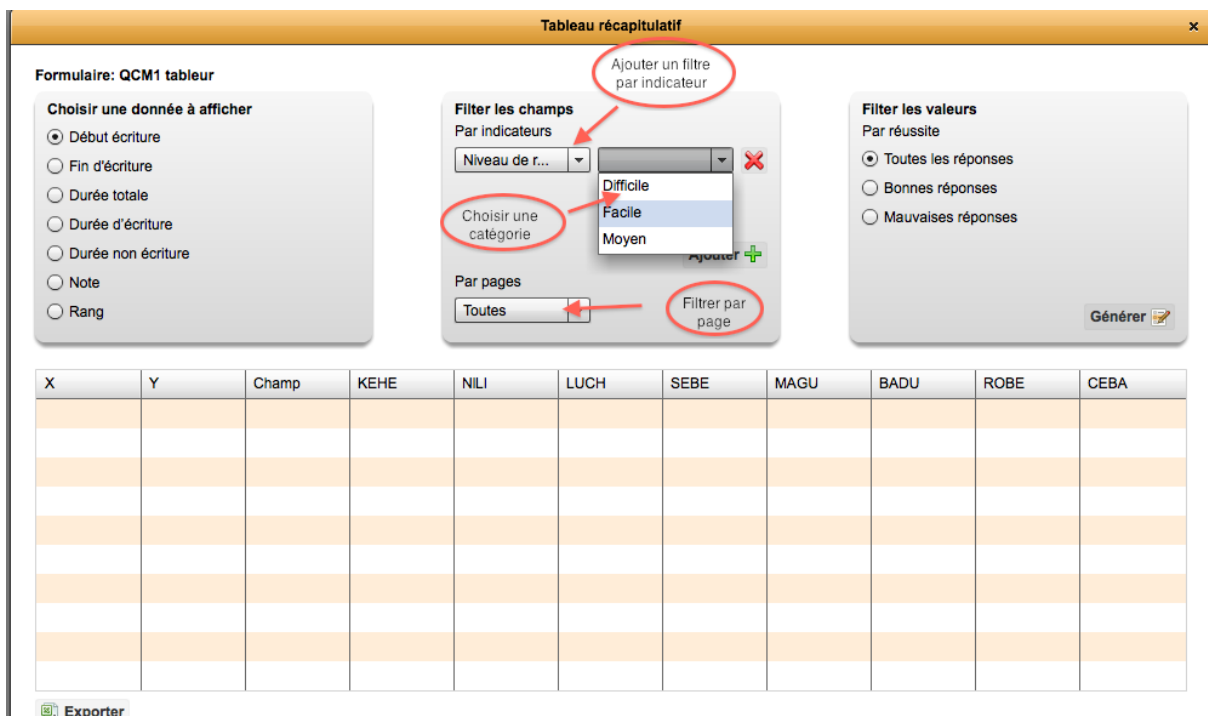


Figure 78 : Fenêtre paramétrable de génération de tableaux pour toutes les données.

Cette fenêtre **paramétrable** permet de **choisir les données** qui vont être **affichées dans le tableau** ainsi que de **choisir des filtres** pour ces données.

→ Manipulation : Paramétrage du tableau.

→ Interaction : Listes déroulantes dynamiques pour choisir les données et les filtres.

→ Manipulation : Choisir une donnée temps.

→ Manipulation : Choisir un autre indicateur.

La première zone à gauche permet donc de **choisir la donnée**, ou bien indicateur temps qui va être sélectionné pour tous les champs et tous les répondants au formulaire. La deuxième zone au milieu permet de **filtrer les valeurs champs** selon leurs appartenances à une classification d'indicateur ou leur appartenance à une page du formulaire. En fin la troisième zone à droite permet de **filtrer selon le résultat du test**, bonnes ou bien mauvaises réponses.

Une fois le paramétrage défini, l'utilisateur doit **cliquer sur le bouton 'Générer'** (voir Figure 79) pour afficher le résultat de la sélection dans le tableau en bas. Il peut ensuite **exporter ce tableau en format xls** en **cliquant sur le bouton 'Exporter'**.

→ Fonctionnalité : Génération du tableau.

→ Fonctionnalité : Edition des données en format xls.

→ Manipulation : Cliquer pour exporter le tableau en format xls.

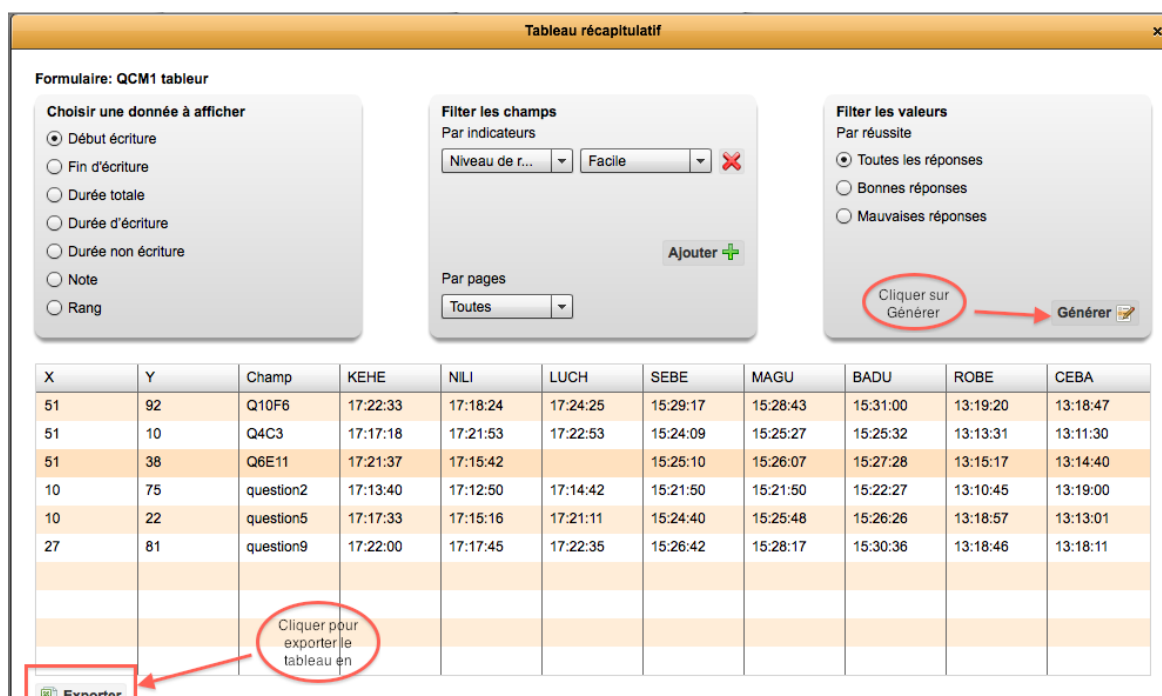


Figure 79 : Fenêtre de tableau des données contenant le résultat de la sélection.

Pour la **représentation graphique de l'ensemble des données**, l'utilisateur doit **cliquer sur le bouton 'Génération de graphiques'** à gauche de l'écran de visualisation pour **afficher** une fenêtre pop-up. Cette fenêtre contient en haut une **barre de menu horizontal** composée de quatre menus, chacun permet de visualiser un type de représentation (voir Figure 80).

→ Fonctionnalité : Génération de graphiques collectifs.

→ Interaction : barre de menu horizontal.

Le premier menu, qui apparaît par défaut (voir Figure 81), permet de visualiser la représentation de la **dispersion des données temps pour l'ensemble des répondants** au formulaire.



Figure 80 : fenêtre de représentation graphiques pour l'ensemble des données.

→ Analyse : Dispersion des données temps.

→ Visualisation : Graphique en nuage de points temps/question.

→ Visualisation : Afficher la légende.

→ Manipulation : paramétrage de la représentation.

Dans cette représentation, l'utilisateur peut choisir entre **deux types de graphiques** ; nuage de points et ligne. Il est possible de **choisir d'afficher les données par page** de formulaire et également de **choisir la donnée temps à afficher** (temps de début/fin d'écriture).

→ Interaction : liste déroulante pour choisir le type de graphique.

→ Interaction : liste déroulante dynamique pour choisir la page du formulaire.

→ Interaction : liste déroulante pour choisir la donnée temps.

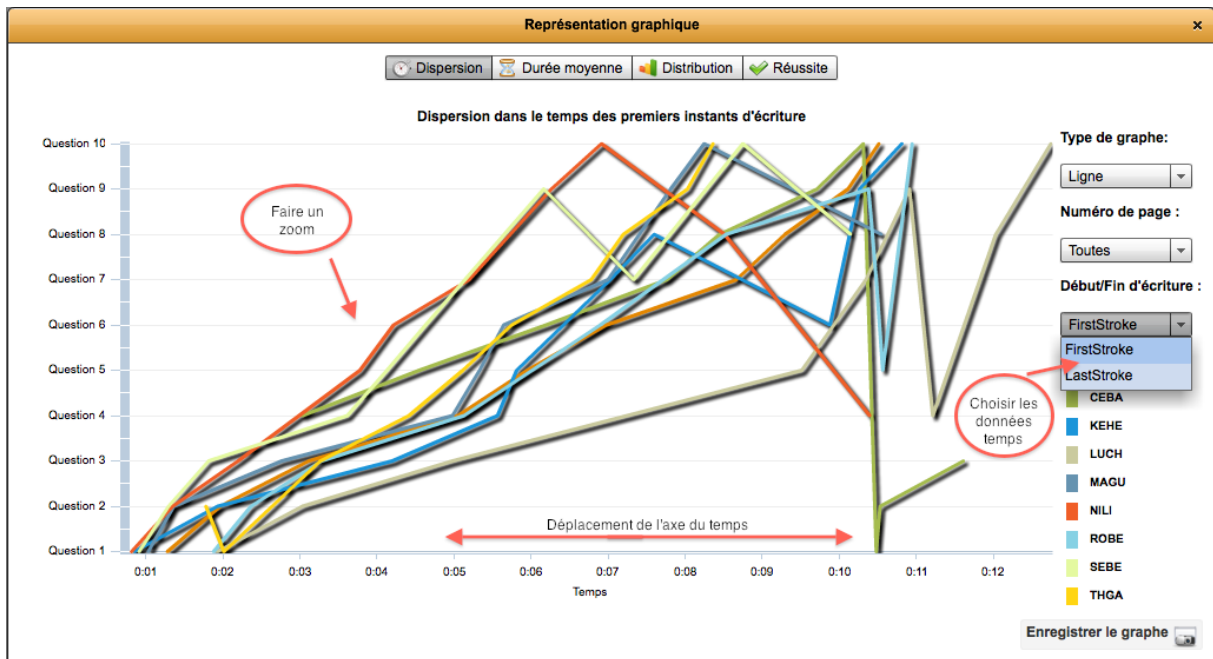


Figure 81 : Représentation en forme de ligne de la dispersion des données temps.

L'utilisateur peut effectuer un zoom sur le graphique, il peut également se déplacer sur l'axe de temps dans les deux sens (voir Figure 82).

→ Interaction : Effectuer un zoom sur le graphique.

→ Interaction : Se déplacer sur l'axe de temps.

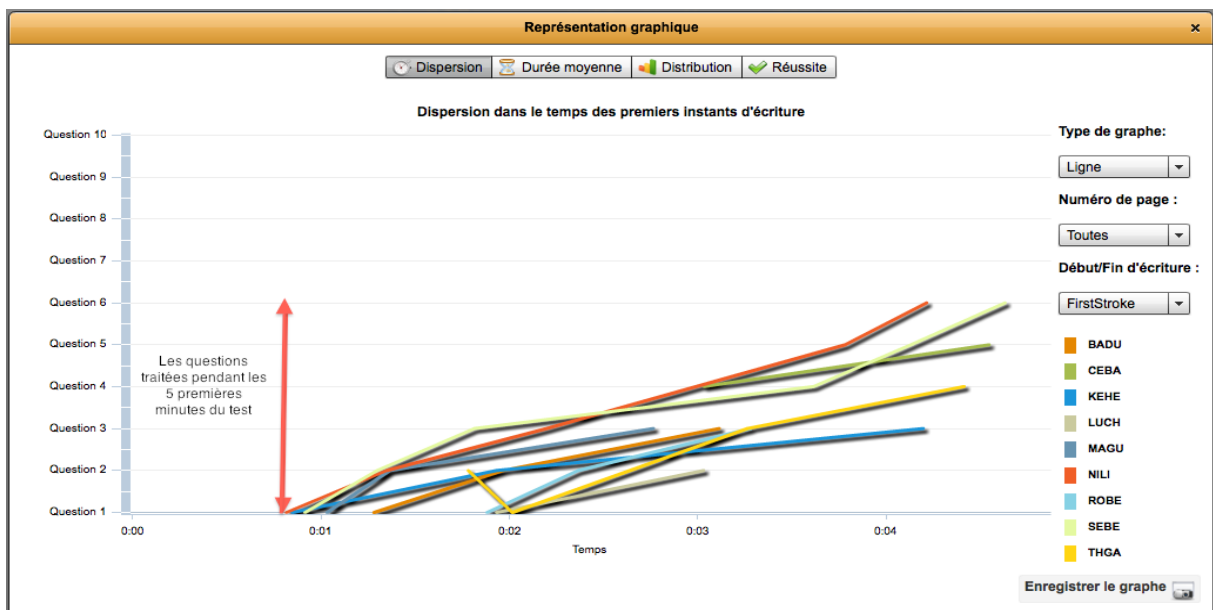


Figure 82 : Représentation de la dispersion des données temps : zoom sur les 5 premières minutes.

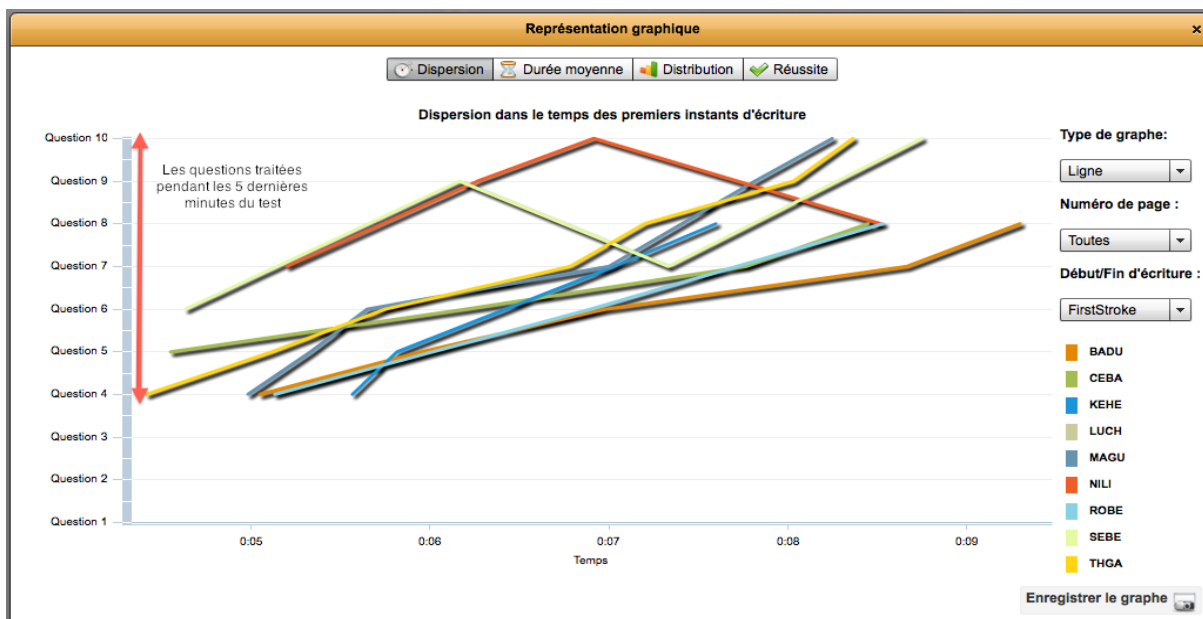


Figure 83 : Représentation de la dispersion des données temps : zoom sur les 5 dernières minutes.

Le deuxième menu permet de visualiser la représentation, sous forme d'histogramme, **des durées moyennes passées sur chaque champ par les répondants** (voir Figure 84).

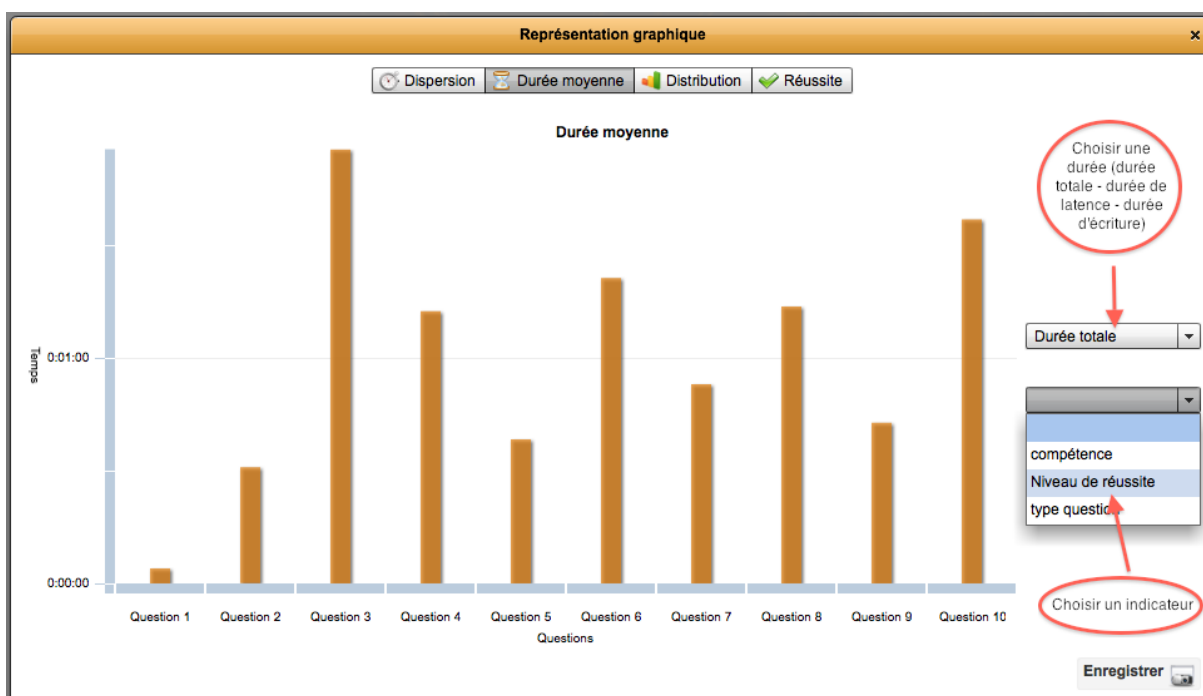


Figure 84 : représentation des durées moyennes.

→ Analyse : Durée moyenne passé sur chaque champ par les répondants.

→ Visualisation : graphique en histogramme.

Dans cette fenêtre, l'utilisateur peut choisir dans une **liste déroulante** la **durée à afficher** dans le graphique (durée totale, durée d'écriture, durée de nom écriture). De plus, il peut **choisir un indicateur** pour **comparer les durées moyennes par rapport à une classification d'indicateur**.

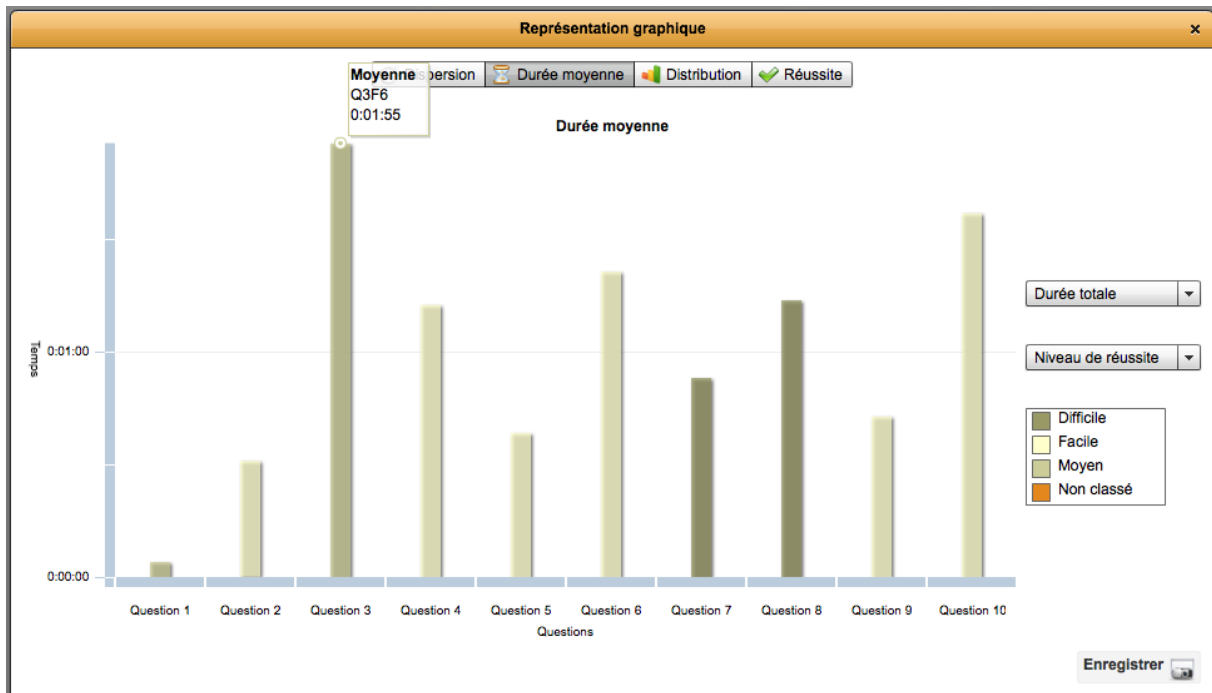


Figure 85 : Représentation des durées moyennes : résultat de choix d'indicateur.

→ Interaction : Liste déroulante statique pour choisir les données temps.

→ Interaction : Liste déroulante dynamique pour choisir les indicateurs.

→ Fonctionnalité : Génération du graphique des durées moyennes.

→ Visualisation : Changement de la couleur des barres du graphique pour chaque choix d'indicateurs.

→ Fonctionnalité : Edition du graphique en format jpg.

→ Manipulation : Cliquer pour enregistrer le graphique en format jpg.

Le troisième menu permet d'afficher **la distribution des durées totales passées sur les champs** par chaque répondant (voir Figure 86).

→ Analyse : La distribution des durées totales passées sur les champs par chaque répondant.

→ Visualisation : afficher un graphique en histogramme empilé.

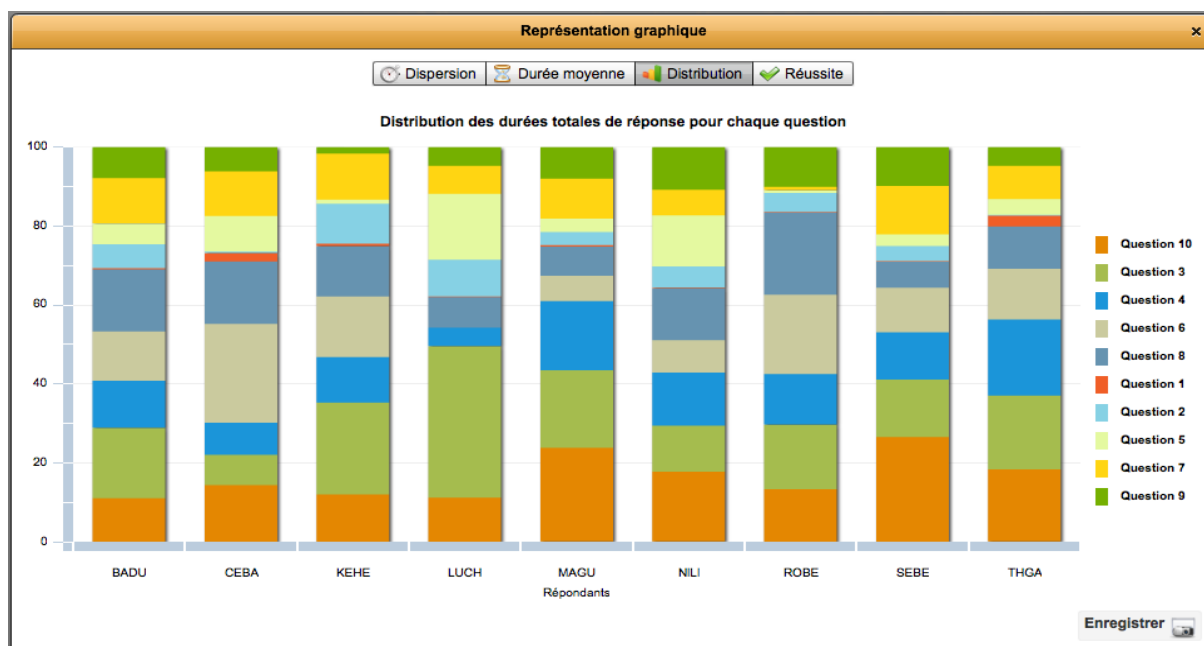


Figure 86 : Représentation de la distribution des durées de réponse au champs par chaque répondant.

Ce graphique permet **d’observer les réponses** anormalement courtes ou bien longues que se soit pour l’ensemble des répondants ou bien pour un seul répondant. Ce qui permet **d’évaluer les questions du test et les répondants.**

→ Analyse : Observer/Evaluer les durées de réponses pour les répondants.

→ Fonctionnalité : Edition du graphique en format jpg.

→ Manipulation : Cliquer pour enregistrer le graphique en format jpg.

Le dernier menu permet de **visualiser la représentation du taux de réussite au test pour chaque question** (voir Figure 87).

→ Analyse : Le taux de réussite au test pour chaque question.

→ Visualisation : Afficher un graphique en histogramme groupé.



Figure 87 : Représentation graphique du taux de réussite au test.

Pour chaque représentation, l'utilisateur peut enregistrer le graphique généré en format jpg en cliquant sur le bouton 'Enregistrer' en bas.

→ Fonctionnalité : Edition du graphique en format jpg.

→ Manipulation : Cliquer pour enregistrer le graphique en format jpg.

Dans le cas d'analyse des données de notes d'observation, l'interface de visualisation (voir Figure 88) permet de **d'effectuer des manipulations sur les instances afficher en liste**. Ce paramétrage consiste à **regrouper les instances en** « Séance d'observation » pour chaque sujet observé.

L'utilisateur **sélectionne une instance** puis **saisit le nom/date** de la séance d'observation à laquelle correspond l'instance. Puis, **clique sur le bouton « Enregistrer »**.

→ Analyse : Regrouper les instances.

→ Interaction : Sélectionner une instance.

→ Manipulation : Saisir le nom/date d'une séance.

→ Manipulation : Cliquer sur le bouton d'enregistrement.



Figure 88 : Regroupement des instances en « séance d'observation ».

Ce regroupement permet de visualiser les données qui correspondent à une séance d'observation d'un sujet dans un même graphique.

Pour **visualiser la représentation graphique de la dispersion dans le temps des actions** observées durant une séance d'observation, l'utilisateur doit **cliquer sur le bouton « Génération de graphiques »** à gauche de l'écran de visualisation pour **afficher** une fenêtre pop-up (voir Figure 89).

→ Analyse : La dispersion dans le temps des actions observées.

→ Manipulation : Cliquer sur le bouton « Génération de graphiques ».

→ Fonctionnalité : Génération du graphique de la dispersion des actions observées dans le temps.

Cette fenêtre contient deux graphiques. Le premier, à droite en séries de nuages de points, représente la dispersion des actions observées et notées au fil du temps.



Figure 89 : Représentation graphique de la dispersion et de la fréquence d’actions observées durant une séance d’observation.

Il est possible de **choisir une séance d’observation dans la liste déroulante à droite. La liste des actions** observées au cours de la séance choisit est alors **affiché en légende.**

→ Visualisation : graphique en nuage de points.

→ Manipulation : Choisir une séance d’observation.

→ Interaction : Liste déroulante dynamique de séances d’observation.

→ Visualisation : Liste des actions en légende.

A gauche il y a un **graphique en barre** qui représente **la densité/ Fréquence d’apparition d’une action durant la séance.** Ce qui permet **d’observer les actions qui se manifestent plus que des autres.**

→ Visualisation : graphique en histogramme en barre.

→ Analyse : Densité/Fréquence d’apparition d’une action.

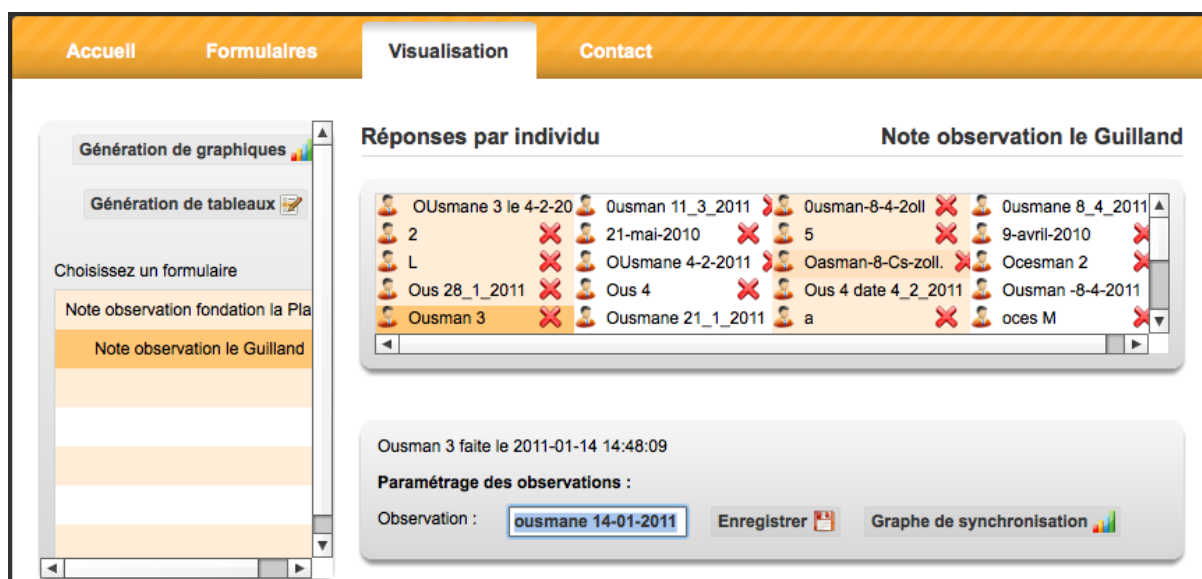


Figure 90 : des instances avec des données de transcription.

Dans le cas où l'utilisateur a importé des données de transcription qui correspondent à une séance d'observation, un **bouton de génération du graphique de synchronisation** va paraître (voir Figure 90).

→ Visualisation : Bouton contextuel de génération de graphique de synchronisation.

Pour **afficher la représentation** de la **dispersion des données d'une séance synchronisée avec des données de transcription**, l'utilisateur doit **cliquer sur le bouton** « Graphe de synchronisation » pour **afficher** une fenêtre pop-up (voir Figure 91).

→ Analyse : Dispersion des données synchronisées avec des données de transcription.

→ Visualisation : Graphique en nuage de points.

→ Manipulation : Cliquer sur le bouton « Graphe de synchronisation ».

→ Fonctionnalité : Synchronisation des données.

Ce graphique représente en nuage de points, les actions notées par le stylo numérique et celles notées depuis la transcription (en forme de croix). A droite de ce graphique se trouve **la liste de toutes les actions en légende**.

→ Visualisation : Afficher la liste des actions en légende.

Par **survol** des points qui représentent les actions, l'utilisateur peut **afficher l'acteur et le commentaire correspondant**.

→ Interaction : survoler les points.

→ Visualisation : Afficher des info-bulles contenant le nom de l'acteur et le commentaire correspondant.

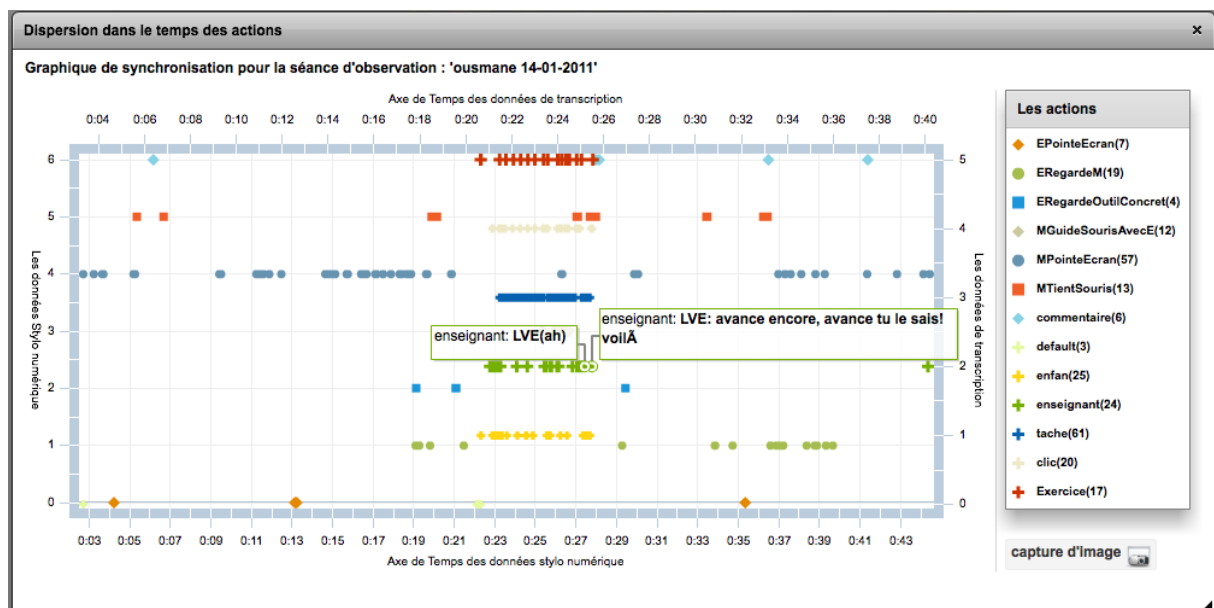


Figure 91 : Représentation de la dispersion des données stylo numérique synchronisées avec des données de transcription.

Ce graphique, formé de 4 axes dont deux temporels, permet à l'utilisateur de **se déplacer par rapport à chaque axe horizontalement**. Il est également possible **d'effectuer un zoom** sur une partie de la représentation.

→ Interaction : se déplacer sur les axes horizontalement.

→ Interaction : Effectuer un zoom sur une partie du graphique.

En se reposant sur cette description de l'interface en utilisation, nous déduisons dans la section 1.5 les caractéristiques de l'interface utilisateur. Dans la section suivante, nous présentons l'architecture de l'application et de la base des données, ainsi que les fonctionnalités de transformation d'indicateurs temps.

5.4. Modélisation de l'architecture et de la base des données

5.4.1. L'architecture de l'application

Dans le but de concevoir une application web qui permet de fluidifier les manipulations utilisateurs, nous avons utilisé de nouvelles technologies du développement web. Notre objectif est d'introduire plus d'interaction au niveau de l'interface utilisateur et d'offrir des fonctions ergonomiques notamment pour la visualisation et la représentation graphique des données.

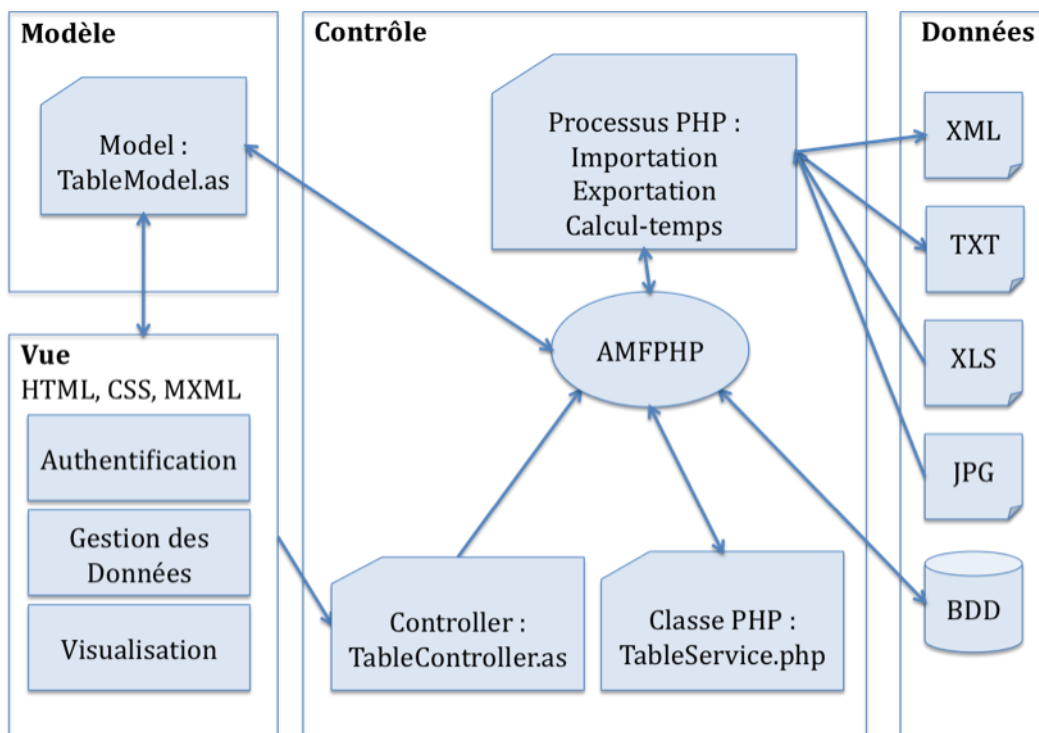


Figure 92 : L'architecture globale de l'application SEGELL.

En se basant sur une modélisation MVC (Modèle – Vue – Contrôle), nous présentons l'architecture de notre application (voir Figure 92). Commenant par la droite, la partie « Données » représente le flux d'échange des données en entrée (fichiers XML et TXT) et en sortie (fichiers XLS et JPG).

Au milieu, la partie « contrôle » représente les traitements centraux de l'application. L'importation et l'exportation des données sont des processus déclenchés par l'utilisateur. Tandis que le calcul des données temps est un processus automatique qui se fait au moment de l'importation des données. Les classes PHP contiennent les requêtes SQL nécessaires pour

la manipulation des données. Pour chaque classe il y a une correspondance « controller » en Action Script permettant la communication avec l'interface.

Dans la partie « Modèle » il y a des classes Action Script qui expriment les tables des données en objets. Des scripts déclenchés depuis la partie « Vue » permettent de manipuler les objets modèles et d'exécuter les requêtes nécessaires. Le résultat du traitement est visualisé dans des représentations.

Une description plus technique du développement du prototype est décrite en annexe B.

5.4.2. La base de données

Pour la conception de la base de données, nous nous sommes basé sur la structure du fichier XML de la trace d'écriture avec le stylo numérique. Essentiellement nous avons repris les éléments indispensables pour les traitements et les analyses auxquels s'intéressent les utilisateurs. Ces éléments de base constituent les informations nécessaires pour la manipulation d'un formulaire. En effet, un formulaire est défini par une structure, chaque exemplaire rempli avec le stylo numérique est une instance de ce formulaire, pour chaque réponse écrite dans une instance correspondent des données trace d'écriture et temps d'écriture.

Dans ce qui suit un extrait de la structure d'un fichier XML de la trace d'écriture exporté depuis la plateforme de Kayentis est présenté.

```
<FormInstances >
  <FormInstance FormInstanceID="" PageAddress="96.529.0.1778">
    ...
    <Current>
      ...
      <Pages>
        <Page PageNumber="1" PageName="Page 1" PageAddress="96.529.0.1777">
          ...
          <Fields>
            <Field Name="num">
              <Value>01</Value>
            ...
            <TimeFirstStroke>2009-06-23 11:20:10</TimeFirstStroke>
            <TimeLastStroke>2009-06-23 11:20:11</TimeLastStroke>
          </Field>
        </Fields>
      </Page>
```

```

</Pages>
</Current>
</FormInstance>

```

La principale est donc la balise ‘*FormInstance*’ qui renseigne sur le nom du formulaire et sa version ainsi que son identifiant. Ensuite, pour chaque exemplaire du formulaire papier rempli, ou instance de formulaire, on trouve la balise ‘*Current*’. Cette balise contient des informations sur l’utilisateur qui a vérifié cet exemplaire sur la plateforme de Kayentis, le stylo avec lequel l’exemplaire a été renseigné et la date d’exportation en fichier XML à l’aide du web service.

```

<FormInstance FormInstanceID="" PageAddress="96.529.0.1766">
  <ProjectName>Eneide - Compétences tableur</ProjectName>
  <FormName>Eusprig Test</FormName>
  <FormVersion>v1</FormVersion>
  <FormID>Eneid_Tableur_Eusprig_V1</FormID>
  <Current>
    <Date>2009-06-23 09:27:53</Date>
    <UserName>Zaara Barhoumi</UserName>
    <PenID>AR5-AAG-JH3-BX</PenID>
    <Status>1</Status>
    <EventType>1</EventType>
  </Current>
</FormInstance>

```

Figure 93 : Extrait du fichier XML ; la balise 'FormInstance' et la balise 'Current'

Dans la balise ‘*Current*’ on trouve la balise ‘*Page*’ qui contient les pages de l’exemplaire. Pour chaque page, on a le nom et le nombre de pages ainsi que son identifiant (celui-ci est défini par la technologie Anoto). On a également les instants de la première et la dernière écriture dans la page.

```

<Pages>
  <Page PageNumber="1" PageName="Page 1" PageAddress="96.529.0.1765">
    <Status>3</Status>
    <EventType>1</EventType>
    <Date>2009-06-23 09:27:53</Date>
    <TimeFirstStroke>2009-06-22 19:31:38</TimeFirstStroke>
    <TimeLastStroke>2009-06-22 19:31:44</TimeLastStroke>
  </Page>
</Pages>

```

Figure 94 : Extrait du fichier XML; la balise 'Page'.

Dans une balise ‘*Page*’ il y a les balises ‘*Field*’ qui contiennent le contenu de chaque champ de saisie, là où on a écrit avec le stylo. Dans chaque balise on a le nom du champ, la valeur écrite (la balise ‘*Value*’) et les instants de la première et la dernière écriture dans le champ (les balises ‘*TimeFirstStroke*’ et ‘*TimeLastStroke*’).

```

<Field Name="FMath">
  <Value>4</Value>
  <Status>2</Status>
  <EventType>2</EventType>
  <Date>2009-06-23 09:42:25</Date>
  <TimeFirstStroke>2009-06-23 09:41:27</TimeFirstStroke>
  <TimeLastStroke>2009-06-23 09:41:31</TimeLastStroke>
</Field>

```

Figure 95 : Extrait du fichier XML : la balise 'Field' d'un champ de saisie.

Concernant les données de description de la structure d'un formulaire, nous avons utilisé un autre fichier d'export XML depuis la plateforme de Kayentis. Ce fichier contient les informations sur les zones d'écriture sensibles pour le stylo numérique.

```

<FormName>QCM3 tableur</FormName>
<IsActive>True</IsActive>
<IsMultiSessions>True</IsMultiSessions>
<Version>v1</Version>
<NumberOfPages>2</NumberOfPages>
<PublicationDate>2009-11-22 17:35:01</PublicationDate>
<Pages>
  <Page PageNumber="1" PageName="Page 1">
    <Fields>
      <Field Name="identifiant">
        <Threshold>96</Threshold>
        <PositionXmin>71,93</PositionXmin>
        <PositionXmax>87,31</PositionXmax>
        <PositionYmin>4,39</PositionYmin>
        <PositionYmax>8,38</PositionYmax>
        <Length>4</Length>
        <PossibleValues>;text:ALNUM</PossibleValues>

```

Figure 96 : Extrait d'un fichier XML de description de la structure d'un formulaire.

Ce fichier contient des informations sur le formulaire comme son nom, son identifiant, le nombre de pages contenu et la date de sa publication. Il existe également la balise 'Page' et la balise 'Field'. Cette dernière représente un élément champ de saisie dans le formulaire. Cette balise contient des informations concernant le nom du champ, sa taille et sa position dans la page ainsi que le type de données qu'il doit contenir. Ces informations nous permettent de calculer la position de chaque champ dans un formulaire et nous permettent de récupérer son type (si une zone de texte, une case à cocher ou bien un bouton radio).

Pour notre modèle de base de données (voir Figure 97), nous l'avons conçu pour supporter cette structuration des données exportées depuis la plateforme de Kayentis. Ainsi, nous avons défini une table « formulaires » de clé unique « FormID » dédié aux informations concernant le formulaire. Et nous avons défini une table « Champs » de clé auto-incrémenté qui

représente les champs d'un formulaire contenant les informations concernant un champ de saisie comme le nom du champ, sa taille et sa position dans la page ainsi que le type de données.

Les instances d'un formulaire sont représentées dans une table « Instances » avec pour clé la valeur PageAdress qui représente un identifiant unique de la première page de l'exemplaire du formulaire. Et les réponses écrites avec le stylo numérique sont représentées dans une table « Réponses ». Cette table a une double clé constituée de la clé de l'instance, à laquelle la réponse appartient, et de la clé du champ, sur lequel on a écrit cette réponse. Dans cette table, pour chaque réponse nous avons mis également les données temps de début et de fin de l'écriture ainsi que les indicateurs temps calculés automatiquement.

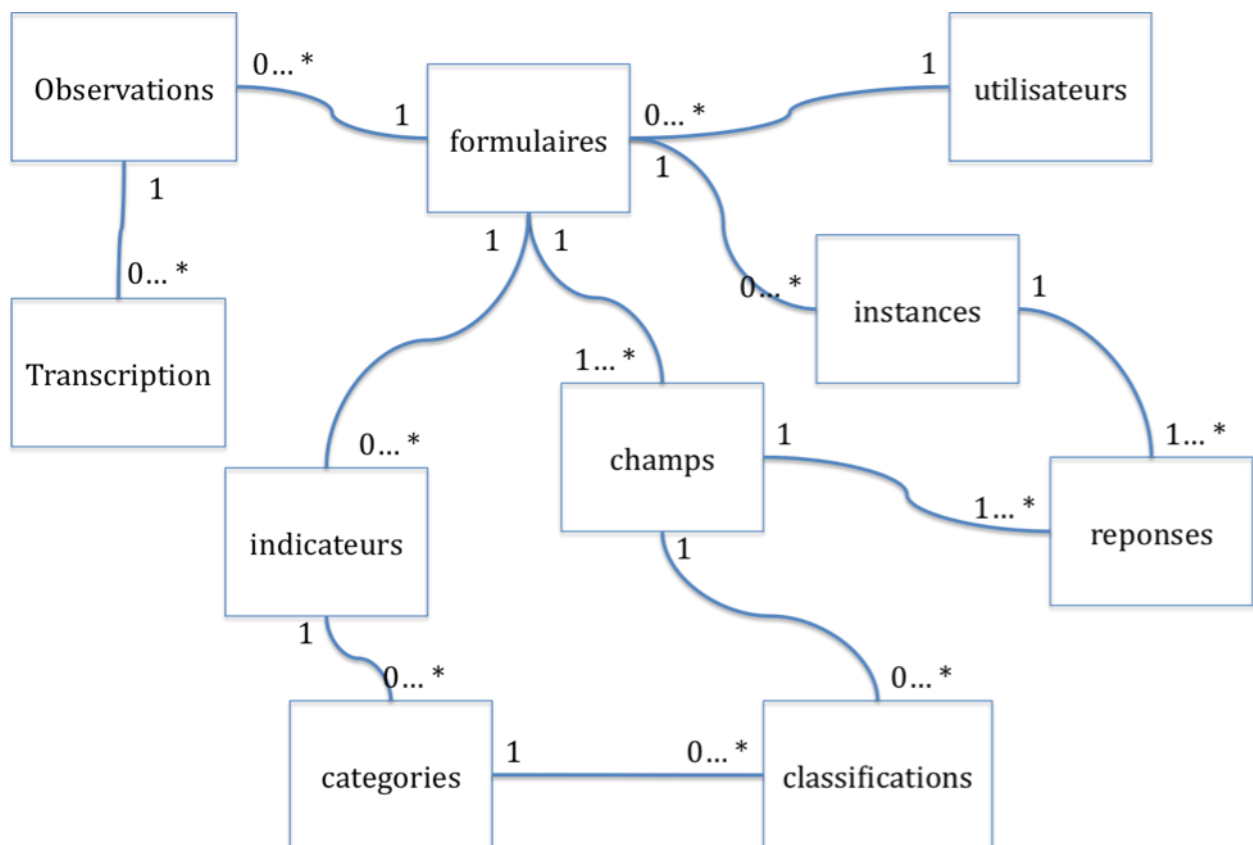


Figure 97 : Le modèle physique des données.

Pour la gestion des indicateurs personnalisés, nous avons ajouté à la base de données les tables 'indicateurs' qui correspond à un formulaire, 'catégories' qui correspond à un indicateur et 'classifications' qui est l'association entre une catégorie et un champ du formulaire.

Pour la gestion des observations, nous avons ajouté deux tables dans lesquelles nous récupérons les données importées depuis un fichier de transcription. Une observation correspond à un formulaire. Pour chaque observation, le nom de l'action et les données temps correspondantes (début, fin, durée) sont récupérés dans la table 'Transcription'.

Enfin, nous sommes conscient que cette base de données, malgré les améliorations qui lui sont apportées d'un prototype à l'autre, n'est pas encore adaptée pour gérer certaines fonctionnalités de manipulation des données comme la transformation d'une réponse en un paramètre ou un indicateur.

5.4.3. Calcul des indicateurs temps

Deux données temps brutes qui correspondent à une réponse (une écriture dans un champ) sont récupérées à partir de la trace à savoir l'instant de début et l'instant de fin de l'écriture. Ces deux instants permettent de calculer la durée d'écriture dans le champ ainsi que l'ordre de traitement.

Cependant, dans certains cas, l'écriture n'est pas ordonnée dans les champs ce qui signifie que les durées ne seront pas exactes. Nous avons discuté cette problématique dans le chapitre 2 (voir paragraphe 2.6.4 page 45). Nous avons donc suivi une stratégie permettant d'avoir les valeurs approximatives des durées d'écriture et de non écriture pour chaque champ. L'algorithme est appliqué automatiquement pour chaque donnée importée.

D'abord, des fonctions de transformation permettent d'unifier le format des données temps récupérées. Ensuite, la fonction d'ordre consiste à récupérer toutes les réponses et les instants correspondant dans un tableau. Le tri de ce tableau, selon les données instant de début d'écriture, permet d'avoir l'ordre de traitement. Puis une autre fonction permet de calculer les durées et d'affecter chaque valeur de durée (durée d'écriture ou bien de non écriture) au champ correspondant. Ceci est appliqué à chaque ensemble de réponses dans un exemplaire de formulaire.

Concernant la synchronisation, une fonction permet de déterminer un instant de début d'écriture pour un exemplaire. C'est à dire un moment qui déclenche le début du déroulement d'un test ou une observation. La valeur de cet instant, transformé en même format que les autres données temps, représente l'instant zéro dans une représentation graphique temporelle.

Ainsi, les données temps des exemplaires d'un formulaire sont synchronisées par rapport à cette valeur. De même, les données temps de transcription d'une séance d'observation sont synchronisées avec les données temps d'écriture avec le stylo numérique.

5.5. Caractéristiques de l'interface d'utilisation

Le prototype de la plateforme SEGELL présente plusieurs traitements et représentations ainsi que des tâches d'interaction.

En nous appuyant sur la description du scénario d'utilisation dans la section 5.3. nous déterminons les caractéristiques de l'interface d'utilisation par rapport à cinq facteurs qui sont : fonctionnalités, visualisation, interaction, manipulation et analyse. Ces facteurs représentent des tâches de traitement et de visualisation pour l'analyse des données. En effet, ces tâches dépendent des objectifs analytiques prédéterminés.

Dans le tableau suivant les caractéristiques du prototype sont regroupées selon les cinq facteurs.

Fonctionnalité	Visualisation	Interaction	Manipulation	Analyse
authentification	menu de navigation contextuelle	parcourir pour ajouter un ou plusieurs fichiers	Saisir le nom d'une catégorie	classer les champs de réponse pour comparer les valeurs de différentes catégories de champs
gestion de formulaire	liste des fichiers sélectionnés	liste déroulante statique pour choisir le type de fichier	Affectation des champs	processus de résolution du test d'un répondant
importation des données	confirmation d'ajout du nouveau formulaire	Sélection d'un formulaire	définir un nom de libelle à chaque champ	Observer la durée passée par le répondant sur une question
Calcul des durées d'écriture	liste des formulaires déjà chargés par l'utilisateur	sélection d'indicateur	Affecter un ordre	le résultat de la réponse (bonne ou bien mauvaise réponse)
Calcul des durées de nom écriture	liste d'indicateurs existants	Sélection de couleur pour chaque catégorie	Affecter une note	instant de début et instant de fin d'écriture d'une réponse

Calcul des durées totales	liste des répondants /instances	Sélection d'un individu	Paramétrer le tableau	durée moyenne passé par tout le groupe de répondants sur chaque question
Calcul des rangs/ordres d'écriture	tableau des données pour un individu	Listes déroulantes dynamiques d'indicateurs	paramétrer la représentation	Comparer le temps passé par le répondant sur chaque question / à la durée moyenne pour tout le groupe
gestion d'indicateurs	graphique en courbe avec marques personnalisée	Listes déroulantes dynamiques pour choisir les données et les filtres	Saisir le nom/date d'une séance	Croisement et comparaisons des données
ajouter/supprimer un indicateur / une catégorie	graphique en barre	barre de menu horizontal	Choisir un indicateur	Durée moyenne passé sur chaque champ par les répondants
gestion des champs	Changement de la couleur du fond de graphique devant chaque champ de réponse	liste déroulante pour choisir le type de graphique	Cliquer pour générer des représentations graphiques/tabulaires d'individu	La distribution des durées totales passées sur les champs par chaque répondant
Générer des représentations graphiques/tabulaires d'un individu	les légendes des valeurs représentées dans les graphiques	liste déroulante dynamique pour choisir la page du formulaire	Cliquer pour supprimer un individu/instance	Observer/Evaluer les durées de réponses pour les répondants
Edition des données en format xls	Graphique en nuage de points temps/question.	liste déroulante pour choisir la donnée temps	Cliquer pour enregistrer une note	Le taux de réussite au test pour chaque question
Edition des graphiques en format jpg	Changement de la couleur des barres du graphique pour chaque choix d'indicateurs	Effectuer un zoom sur le graphique	Cliquer pour exporter les données en format xls	Regrouper les instances
Génération du tableau collectif	graphique en histogramme empilé	Se déplacer horizontalement sur l'axe de temps	Cliquer pour exporter les graphiques en format jpg	La dispersion dans le temps des actions observées
Génération de graphiques collectifs	graphique en histogramme groupé	Sélectionner une instance	Choisir une donnée temps	Densité/Fréquence d'apparition d'une action
Synchronisation	Bouton	Liste	Cliquer sur le	Dispersion des

des données	contextuel de génération de graphique de synchronisation	déroulante dynamique de séances d'observation	bouton « Génération de graphiques »	données synchronisées avec des données de transcription
	des info-bulles contenant le nom de l'acteur et le commentaire correspondant	survoler les points	Choisir une séance d'observation	

Tableau 2 : Les caractéristiques du prototype déterminées depuis le scénario d'utilisation regroupées selon les cinq facteurs

Dans le tableau suivant, nous distinguons pour chaque facteur les tâches principales du système ainsi que les éléments graphiques de l'interface :

Facteurs	Tâches/éléments graphiques
Fonctionnalité	Importation des données Calcul des données temps Gestion d'indicateur et de catégorie Gestion et classification des champs Génération des tableaux des données (individuels ou bien collectifs) Génération de graphiques (individuels ou bien collectifs) Edition des données en format xls Edition des données en format jpg Synchronisation des données.
Visualisation	Menu de navigation Liste des fichiers (de formulaires, d'indicateurs, d'instances) Tableau de classification des champs Tableau de données pour instance Tableau de données collectif Graphique en courbe avec marques personnalisé Graphique en barre

	<p>Graphique en histogramme groupé</p> <p>Graphique en histogramme empilé</p> <p>Graphique en nuage de points</p> <p>Graphique en lignes</p> <p>Feedback de manipulation (changement de couleur du graphique et des légendes)</p> <p>Info-bulles de détails.</p>
Interaction	<p>Parcourir</p> <p>Liste déroulante statique (choisir le type de fichier, le type de graphique...)</p> <p>Liste déroulante dynamique (choisir un indicateur, une séance d'observation, une page du formulaire...)</p> <p>Sélection d'un formulaire / indicateur / instance etc.</p> <p>Zoom sur le graphique</p> <p>Déplacement horizontal sur les axes</p> <p>Survole des points dans un graphique.</p>
Manipulation	<p>Saisir un nom, date, libellé, ordre</p> <p>Affecter un champ, une note etc.</p> <p>Paramétrer un tableau / un graphique</p> <p>Choisir dans une liste déroulante (un indicateur, une donnée, une séance ...)</p> <p>Cliquer sur un bouton (pour afficher, générer, exporter ..)</p>
Analyse	<p>Classer les champs, les instances</p> <p>Comparer les champs de différentes catégories</p> <p>Le processus de résolution</p> <p>Observer les durées, la note, les instants d'écriture ...</p> <p>La durée moyenne passé par tous les répondants</p> <p>La distribution de la durée totale</p> <p>Le taux de réussite</p> <p>Dispersion dans le temps</p>

	Densité/fréquence nombre d'occurrences.
--	---

Tableau 3 : Les tâches et les éléments graphiques de l'interface.

Finalement, dans la suite, nous distinguons ce qui caractérise l'interface de l'application en terme de traitement, visualisation et interaction :

Traitement : Les fonctionnalités, coté système, et les manipulations, coté utilisateur, représentent des tâches de traitement des données :

- Traitement automatique : il s'agit des transformations de la trace effectuées au moment de l'importation des données dont le calcul des données temps et la synchronisation des données.
- Gestion : une tâche de transformation des données effectuée manuellement par l'utilisateur. Ici, il s'agit des tâches de manipulation des données : ajout d'indicateurs personnalisés, classification des champs et regroupement des données.
- Génération : un traitement semi-automatique. Il se manifeste dans l'exécution des requêtes pour obtenir un ensemble de données depuis la base des données ainsi que les calculs statistiques comme le pourcentage et la moyenne. Ces traitements analytiques préparent les données pour les représenter en tableaux ou bien en graphiques.
- Edition : il s'agit des traitements qui permettent d'exporter les données sélectionnées en différents formats.

Visualisation : la visualisation des données se manifeste dans trois types de représentation :

- La représentation en liste : pour afficher un ensemble de données et permettre la sélection d'un élément pour accéder à un détail ou bien une fonctionnalité (de traitement ou bien de représentation). Ces listes facilitent donc la navigation dans la trace.
- La représentation tabulaire : permet d'afficher les détails des informations qui correspondent à un élément, mais aussi de les modifier.

- La représentation graphique : la représentation finale des données traitées et manipulées pour réaliser l'analyse. En outre, des fonctions y sont intégrées permettant de personnaliser l'affichage et le contenu afin de raffiner l'analyse. De plus, certains graphiques représentent plusieurs valeurs pour faciliter la comparaison des données.

Pour cette dernière catégorie de représentation différentes formes de graphiques sont possibles : graphique en courbe avec marques personnalisée, graphique en nuage de points, graphique en ligne, histogrammes, histogrammes groupés, histogramme empilés, graphique en barre.

Nous retenons également les représentations sur fond personnalisé que nous avons utilisé dans le prototype 2 (voir chapitre 3 : expérience RJC) : graphique en cercle sur fond personnalisé et graphique en courbe avec marques sur fond personnalisé.

Dans le tableau suivant, nous distinguons les tâches d'analyse et les graphiques représentés correspondant.

Taches d'analyse	Graphiques
Le processus de résolution	Courbe avec marques personnalisées, courbe avec marques sur fond personnalisé
Classification	Histogramme groupé, barre, cercle sur mappe
Comparaison	Ligne, Nuage de point, Histogramme groupé, barre
Distribution	Histogramme empilé
Dispersion dans le temps	Ligne, Nuage de point
Densité/fréquence	Barre
Le taux de réussite	Histogramme groupé
La durée moyenne	Histogramme, barre

Tableau 4 : les représentations graphiques pour chaque tâche d'analyse.

Interaction : en terme d'interaction, nous distinguons les tâches suivantes :

- Navigation : pour passer d'un écran à l'autre nous avons les menus et les menus contextuels ainsi que les boutons. Le schéma de navigation sur cette plateforme n'est pas linéaire, il est conçu en étoile pour faciliter l'accès à chaque fonctionnalité depuis n'importe quel écran. En outre, d'autres moyens de

navigation sont conçus pour permettre la navigation dans les représentations des données, comme par exemple la sélection d'élément dans les listes ou bien le déplacement sur l'axe de temps d'un graphique.

- Filtrage : les fonctionnalités de paramétrage permettent d'effectuer le filtrage des données. La sélection des paramètres depuis les listes déroulantes conduit au filtrage des données à afficher dans les tableaux ou bien les graphiques.
- Détail sur demande : cette tâche se manifeste dans la sélection des données dans les listes représentées. Par exemple, sélection du formulaire, puis sélection d'une instance. Sur les graphiques, l'affichage des info-bulles est un exemple.
- Zoom : une technique d'interaction permettant d'optimiser la visibilité et la lisibilité de la représentation graphique des données.

En somme, la méthode de réalisation, basée sur une conception participative, de cette plateforme est une méthode centrée utilisateur. Cet aspect se manifeste bien d'après l'analyse de l'interface d'utilisation de la plateforme. En effet, nous avons aperçu durant le processus de fonctionnement de l'application les tâches de l'utilisateur à savoir des tâches de manipulation, des tâches d'interaction et des tâches d'analyse.

Parallèlement, les fonctionnalités principales coté application sont la transformation des données (calcul, gestion, synchronisation), la génération des visualisations et l'édition des représentations.

5.6. Conclusion

Dans ce chapitre, nous nous sommes concentré sur la plateforme prototype SEGELL en terme d'implémentation informatique, d'une part, et en terme d'analyse de son interface à travers un scénario d'utilisation, d'autre part. Cette analyse nous a permis de définir les principales fonctionnalités et représentations illustrées dans l'application ainsi que les tâches de l'utilisateur. L'utilisation d'une future réalisation dans d'autres contextes d'usage pourrait enrichir notre définition des fonctionnalités et des tâches d'utilisateur.

Dans le chapitre suivant, nous présentons, d'une façon plus abstraite, la méthode développée depuis la phase d'instrumentation. Nous extrairons un modèle générique pour le traitement et la représentation des données recueillies.

Chapitre 6. Méthode d'analyse et modèle générique

6.1. Introduction

A partir des démarches effectuées durant les expériences précédemment décrites, nous allons abstraire une méthode générale allant de l'instrumentation à l'analyse. Notre problématique est de concevoir un modèle général pour gérer ces analyses.

6.2. La méthode d'analyse

S'agissant d'une trace numérique, en général, une approche utilisée dans des systèmes exploitants des traces numériques (Systèmes à Base de trace (Settouti et al., 2009) et Systèmes d'Analyse Automatique de l'Interaction (Dimitracopoulou, 2008)) consiste à passer par les trois fonctionnalités suivantes :

- collecte des données, qui consiste à alimenter le système avec de données issues de l'interaction d'un utilisateur avec son système informatique.
- transformation de la trace, qui consiste à produire des indicateurs appropriés et préparer les données.
- présentation des indicateurs, qui permet l'accès concret à la trace pour son exploitation.

Ainsi, nous distinguons trois processus essentiels qui constituent notre méthode d'analyse à savoir le processus de recueil des données, le processus de traitement des données et le processus de visualisation des données.

Cependant, dans le cas d'usage d'un stylo numérique comme moyen de collecte des données, le système d'analyse ne traite pas l'étape de collecte des données. Il existe une étape d'instrumentation qui consiste à définir un cadre structurel de la trace avec un codage spécifique au contexte d'usage (expérimental). Ensuite, au sein du système d'analyse, des fonctionnalités de traitement (transformations, calcul, synchronisation) et de visualisation vont être concrétisées.

6.2.1. L'instrumentation et le recueil des données

L'analyse des données issues d'un stylo numérique dépend essentiellement du contexte d'usage qui définit l'usager du stylo, les données à instrumentées et le codage nécessaire pour obtenir des données intéressantes.

6.2.1.1. Contexte d'usage

Chaque contexte d'usage est caractérisé par des objectifs expérimentaux spécifiques. Nous distinguons le cas où l'utilisateur du stylo est le chercheur lui-même, et le cas où l'utilisateur du stylo est une autre personne. Pour les deux cas, au niveau de l'instrumentation nous pouvons utiliser les mêmes outils. Il est par contre important de considérer la différence au moment du remplissage du papier. Ainsi, l'utilisateur pour le premier cas est plus « familiarisé » avec l'écriture avec cet outil. Tandis que dans le deuxième cas, il est nécessaire de prévoir un apprentissage et des explications pour l'utilisateur afin de le familiariser avec l'outil stylo numérique ou alors pour respecter un certain protocole de recueil des données.

6.2.1.2. L'instrumentation

L'instrumentation est une manipulation faisable grâce à des outils compatibles avec la technologie Anoto et consiste à délimiter des zones dans un formulaire. Ces zones de forme différentes (texte, case à cocher, bouton radio etc.) deviennent les objets sensibles au stylo numérique. Le paramétrage de chaque zone (nomenclature, type, format, et reconnaissance de la donnée capturée) ainsi que les données temps et l'écriture du stylo constituent la trace structurée.

Après une première conception d'un formulaire, il est nécessaire de le tester pour valider, d'une part, son contenu par rapport aux objectifs analytiques, et d'autre part, le bon fonctionnement de la capture des données et de la reconnaissance.

6.2.1.3. Les données recueillies

La trace du stylo numérique est une trace riche avec les différentes informations qu'elle récupère (données temps, position, vitesse, reconnaissance de l'écriture, reconnaissance de signe...). Toutefois, le contenu de la trace dépend du système d'instrumentation utilisé et du système de récupération de la trace. Dans notre cas, même si les outils de Kayentis sont adaptés pour gérer des formulaires de type QCM, nous récupérons une trace plus compacte

par rapport à celle d'origine.

La granularité de la trace récupérée représente un objet, défini en phase d'instrumentation, qui contient des données temps (instant de début et instant de fin d'écriture) et d'autres données. Après le retraitement de la trace, s'ajoutent de nouvelles données temps (durée et ordre) ainsi que des données personnalisées.

En outre, dans certains cas, la trace du stylo numérique n'est pas suffisamment complète pour des analyses expérimentales. Pour cela, il nous semble indispensable d'enrichir ces données avec d'autres données issues de sources de recueil différentes. Une phase de synchronisation est donc nécessaire.

6.2.2. *Le système de traitement et de visualisation*

Le but final de la conception de tel système est la visualisation informatique de données numériques pour atteindre des objectifs d'analyse. Ainsi, cette problématique concerne différents domaines dont l'interface utilisateur, les techniques de représentations des données (algorithmes et formes), les méthodes statistiques et les techniques d'analyse de données. Un système de visualisation est un système qui intègre des techniques de visualisation en appliquant des transformations analytiques (Qin, Zhou, & Pei, 2003).

Aujourd'hui, il y a plusieurs outils de visualisation pour différents intérêts. L'usage des taxonomies comme moyen d'assistance, tant pour l'utilisateur que pour le développeur, pour sélectionner l'outil approprié pour atteindre des objectifs d'analyse est devenu une nécessité. Ces taxonomies ont permis de classer les techniques, les modèles et les systèmes de visualisation selon des critères à savoir le type des données, les tâches de l'utilisateur et la représentation visuelle.

Dans la littérature, nous nous sommes intéressée aux travaux qui concernent les données temporelles d'une part et les tâches de l'utilisateur d'autre part, ainsi qu'aux modèles de référence des systèmes de visualisation d'une façon générale.

Concernant le type des données, (Shneiderman, 1996) distingue, à travers sa taxonomie dite (Type by Task taxonomy TTT), sept types de données qui sont 1D, 2D, 3D,

Multidimensionnelles, Réseaux, Arbre et Temporelles. Pour ce dernier type de données, souvent considéré comme un paramètre parmi d'autres dans une représentation visuelle, il est nécessaire pourtant d'accorder une considération au facteur temps (Muller & Schumann, 2003). Une donnée temporelle est définie dans (Daassi, 2003) comme étant une association de valeurs structurelles (sa dimension structurelle) et de valeurs temporelles (sa dimension temporelle).

(Muller & Schumann, 2003) proposent une classification centrée sur les données temporelles. Ils ont identifié plusieurs aspects du temps qui sont ; intervalle (présente une durée) vs instant (un point sans durée), linéaire (événements du passé vers le futur) vs cyclique (événements périodiques), ordinal (qualitative) vs continu (quantitative), ordonné (événements successives) vs ramifié (événements alternés). Ensuite, (Aigner et al., 2008) proposent une classification de visualisation basé sur trois facteurs : le temps (primitive temporelle et structure), la donnée (cadre de référence, nombres des variables et niveau d'abstraction) et la représentation (dynamique, statique et la dimensionnalité).

Un modèle de référence d'état de données pour décrire le processus de visualisation est proposé par (Chi, 2000). Ce model décrit les techniques de visualisation avec un intérêt aux données et leurs transformations. Le processus de visualisation est divisé en quatre étapes : valeur, abstraction analytique, abstraction de visualisation et la vue. Trois types de transformation de données : transformation des données, transformation de visualisation et la transformation de la vue.

Une extension de ce modèle pour tenir compte d'une dimension temporelle a été proposée par (Daassi, 2003). Le problème avec ce modèle c'est qu'il permet de décrire le processus d'une technique de visualisation, alors qu'un système de visualisation peut contenir une ou plusieurs techniques. De plus, les tâches de l'utilisateur sont peu exprimables à partir de ce modèle.

Le modèle de référence proposé dans (Card, Mackinlay, & Shneiderman, 1999) quant à lui présente presque les mêmes étapes du processus de visualisation en tenant compte des tâches d'interaction de l'utilisateur. A partir des collections de données structurées ou non structurées, des transformations sont nécessaires à faire pour avoir une visualisation de ces données. Ces transformations comprennent le filtrage, le calcul et la normalisation des données afin d'obtenir des données en structure unifié. Ensuite, des correspondances visuelles

transforment les données en structures visuelles adéquates. A partir de ces structures, un ensemble de vues, permettant la navigation dans les données, est généré.

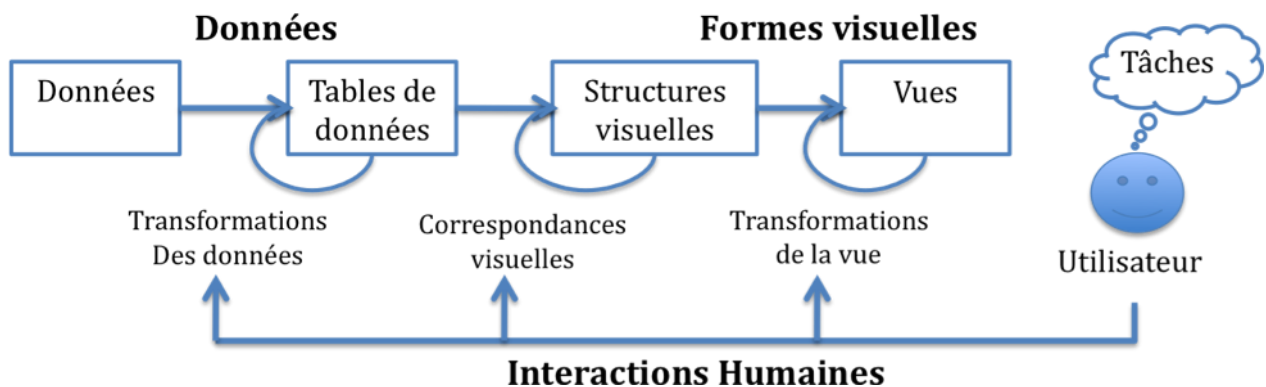


Figure 98 : Modèle de référence pour la visualisation des données présenté par (Card et al., 1999).

L'utilisateur peut donc interagir avec le processus de transformation en différentes étapes. Il peut ajuster les représentations, changer la structure visuelle ou bien effectuer des transformations aux données.

Plus récemment, Colin Ware (Ware, 2004) a schématisé les étapes de la visualisation d'une façon plus abstraite. Quatre phases sont combinées dans un cycle itératif de feedback. Les quatre étapes consistent en :

- Le recueil et la sauvegarde des données.
- Le pré-traitement conçu pour transformer les données en quelque chose compréhensible.
- L'outil visuel et les algorithmes graphiques permettant de produire des images sur l'écran.
- Le système sensoriel et cognitif humain.

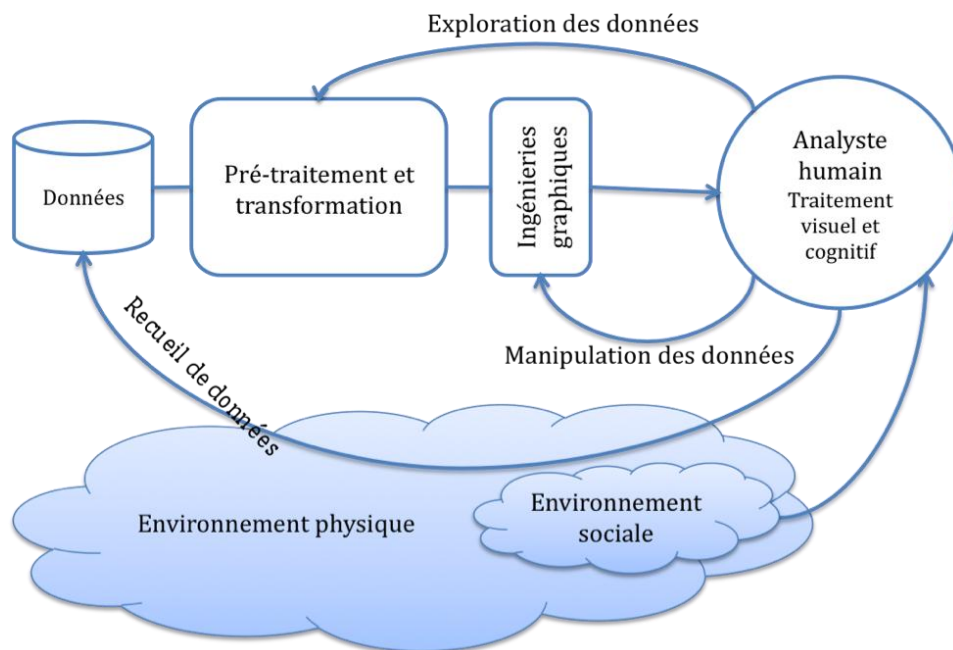


Figure 99 : schéma des étapes de visualisation proposé dans (Ware, 2004).

L'intérêt de ce modèle est qu'il tient compte des facteurs extérieurs au système de visualisation. De plus, il schématise le rôle de l'utilisateur (analyste, scientifique et chercheur d'information) dans chaque phase physique, notamment pour le recueil des données, ou bien informatisé. Sans oublier les ingénieries (transformation ou bien algorithme graphique) réalisées par le système en arrière fond. Les tâches d'interaction avec le système sont des tâches d'exploration des données et des tâches de manipulation des données.

Par ailleurs, en terme de tâche de l'utilisateur, chaque étape à son propre ensemble de tâches. Dans sa proposition, (Daassi, 2003) décrit des techniques d'interaction avec des représentations de données temporelles en se reposant sur des tâches d'analyse de données qui sont ; tâche de navigation, tâche d'observation, tâche de comparaison et tâche de manipulation. En outre, (Shneiderman, 1996) a décrit dans sa taxonomie, d'une façon plus abstraite, sept tâches que l'utilisateur peut effectuer dans une interface de visualisation. Il a également recommandé une stratégie d'interaction sous forme de Mantra de visualisation qui ordonne les tâches de la façon suivante :

- Vue d'ensemble en premier : Les utilisateurs ont besoin d'avoir un aperçu de toute la collection de données.
- Zoom : Les utilisateurs effectuent le zoom pour deux raisons, soit pour se

concentrer sur un sous-ensemble des données ou bien pour voir plus de détail graphique.

- Filtre : En plus de la réduction de l'ensemble de données graphique réalisé par le zoom, les utilisateurs ont souvent besoin de réduire la taille de l'ensemble des données en éliminant les éléments basés sur leurs attributs.
- Détails sur demande : Tout en explorant l'ensemble de données, les utilisateurs veulent voir les détails d'un élément de données particulier. Ceci est habituellement accompli en cliquant sur un ou plusieurs éléments. Les détails sont donc affichés dans une fenêtre pop-up.
- Lier à : Si les utilisateurs découvrent un élément d'intérêt, ils peuvent avoir besoin de connaître les autres éléments ayant des attributs similaires. Le fait de cliquer sur un élément soulignerait dans la représentation les éléments qui lui sont liées.
- Historique : Les utilisateurs ont besoin d'annuler et rejouer des actions. Il est donc important de conserver un historique des actions de l'utilisateur et lui permettre de les manipuler.
- Extraction : Les résultats d'exploitation d'une visualisation doivent être extraites à partir de l'interface. Les utilisateurs peuvent extraire des sous-ensembles de données pour une analyse plus approfondie.

Dans notre étude, nous avons introduit la plupart de ces tâches (navigation, filtrage, zoom, détails sur demande et extraction). La stratégie de la Mantra peut également être vérifiée. Cependant, ces tâches qui correspondent à l'outil informatique, ne tiennent pas compte des tâches de l'utilisateur qu'il pourrait effectuer en dehors du système de visualisation, comme le recueil et le codage des données.

Nous avons identifié d'autres types de tâches comme les tâches d'analyse qui consistent par exemple à regrouper et classer les données, ainsi que les tâches de manipulation qui permettent de transformer les données et/ou les graphiques.

6.3. Notre modèle

En se basant sur l'analyse de l'interface d'utilisation du prototype de la plateforme SEGELL, et en s'inspirant des travaux que nous avons cité précédemment, nous proposons notre modèle de système de traitement et de visualisation. Dans ce modèle nous mettons en avant essentiellement les processus de visualisation par rapport aux tâches de l'utilisateur.

Outre la dimension cognitive, qui se manifeste à travers des tâches de l'utilisateur, nous schématisons le processus de la visualisation par trois étapes qui sont :

- Le recueil des données
- Le traitement des données
- La visualisation des données

Nous considérons le rôle de l'utilisateur dans chaque étape du processus. Allant de la phase de recueil des données, puis le traitement et enfin la visualisation, l'utilisateur intervient à travers un ensemble de tâches. Nous distinguons trois types de tâches ; les tâches d'analyse, les tâches de manipulation et les tâches d'interaction (voir Figure 100).

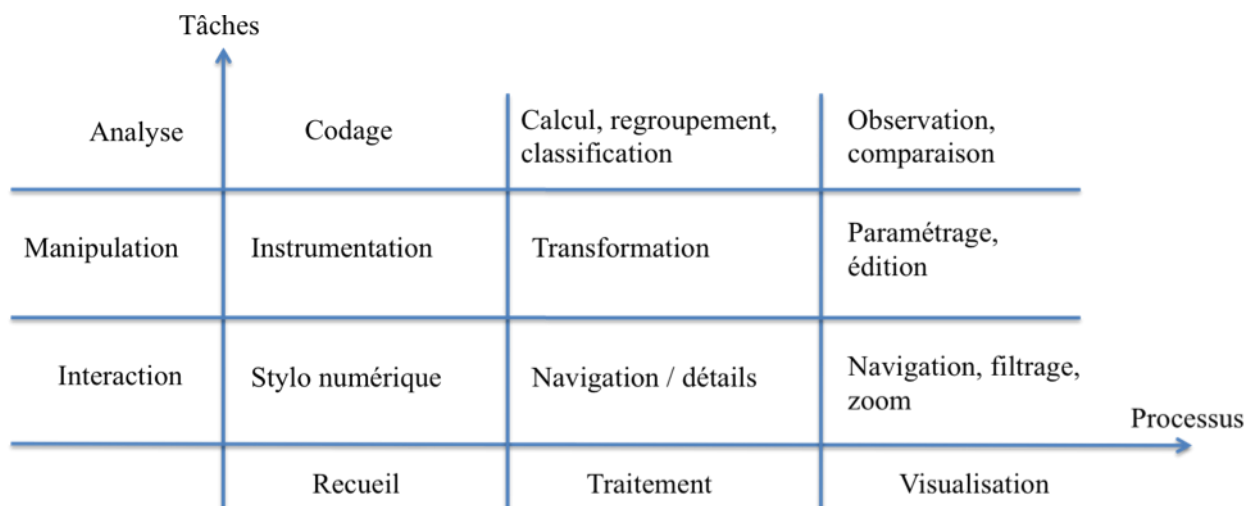


Figure 100 : Représentation de notre modèle en fonction du processus de l'analyse d'une part, et en fonction des tâches de l'utilisateur d'autre part.

Nous définissons les tâches d'interaction par l'ensemble des tâches qui concernent l'usage et la manipulation de l'interface. Les tâches de manipulation quant à elles représentent les tâches

qui permettent de manipuler les données à travers l'interface. Et les tâches d'analyse se sont les tâches issues d'un besoin analytique ou bien d'une réflexion cognitive de l'utilisateur.

Pour la phase de traitement, Il est important en premier lieu d'identifier l'aspect de la donnée structurelle et l'aspect de la donnée temporelle pour pouvoir préciser une stratégie de traitement. Pour notre cas, un élément est composé de plusieurs données structurelles (comme le nom, la position, la page, le type et la réponse) et de deux données temps (sous forme d'instant de début et instant de fin). A partir de ces données brutes, il est possible d'appliquer des transformations semi-automatiques, effectuées par l'utilisateur, ou bien des transformations automatiques, effectuées par le système. Les transformations semi-automatiques concernent notamment les données structurelles et sont appliquées à travers un ensemble de tâches. Alors que les transformations automatiques concernent notamment les données temporelles et sont appliquées au moment de l'importation des données.

En se basant sur les questions posées par (MacEachren, 1995) sur les données temporelles, nous proposons un ensemble d'indicateurs temps calculés automatiquement depuis la trace brute. Nous correspondons chacune de ces questions, tel qu'elles sont présentées dans (Muller & Schumann, 2003), à un indicateur temps :

Questions	Indicateurs temps
Un élément de données existe-t-il à un moment donné ?	Instant de début et instant de fin de l'écriture
Quand est ce qu'un élément de données existe ? Y a t-il un comportement cyclique ?	Durée de latence
Quelle est la durée passée du début à la fin d'un élément de données ?	Durée d'écriture
Comment un élément de données change-t-il ou bien quelle est la différence entre deux éléments de données au fil du temps ?	Nombre de modification (si la trace contient les moments de modifications).

Dans quel ordre les éléments de données apparaissent-ils ?	Ordre de traitement
Combien de fois un élément de données se produit-il ?	Nombre d'occurrence d'une action
Est-ce que les éléments de données peuvent exister ensemble ?	Synchronisation

Tableau 5 : les indicateurs temps.

En plus de cet ensemble d'indicateurs calculés par le système, l'utilisateur peut définir un codage des données lui permettant l'automatisation d'autres indicateurs structurels. Par ailleurs, les manipulations semi-automatiques sont réalisables à travers l'interface de la visualisation permettant d'introduire d'autres indicateurs et classifications des données.

Au niveau de l'interface de la visualisation, nous distinguons trois types de représentations qui sont la représentation en liste, la représentation tabulaire et la représentation graphique.

La représentation en liste permet d'avoir un aperçu général de la trace. Le contenu textuel facilite la navigation hiérarchique. Le fait de cliquer sur un élément d'un ensemble de données permet d'afficher un sous-ensemble correspondant. Les listes déroulantes permettent de filtrer également les données.

La représentation tabulaire permet d'afficher plus de détails sur un élément. Il est possible de paramétrer ou bien de modifier le contenu d'un tableau. De plus, la fonction d'exportation permet d'extraire un ensemble structuré de données.

La représentation graphique permet de concrétiser les objectifs analytiques comme la comparaison. A ce niveau, plusieurs tâches d'interaction sont possibles comme le zoom, le filtrage, le détail et l'extraction des données.

Quant à la forme de ces représentations graphiques, nous avons utilisé dans notre étude des techniques conventionnelles basées sur des diagrammes puisque les données traitées, même si dépendantes du temps, sont essentiellement des données quantitatives.

6.4. Conclusion

L'analyse de l'interface d'utilisation du prototype de la plateforme SEGELL nous a permis d'abstraire un modèle générique d'un système de traitement et de visualisation des données. A travers ce modèle, nous avons illustré trois types de tâches de l'utilisateur durant le processus d'analyse à savoir les tâches d'analyse, les tâches de manipulation et les tâches d'interaction. Ainsi, l'élaboration de ce modèle montre l'importance de la considération du rôle de l'utilisateur, comme étant expert de la donnée à traiter, dans chaque étape du processus.

Quoique issu d'une méthode de prototypage exploratoire et d'une conception participative, ce modèle devrait être validé sur d'autres applications. Ce qui ouvre des perspectives pour d'une part effectuer des réalisations informatiques, et d'autre part améliorer son aspect conceptuel.

Chapitre 7. Conclusion et perspectives

7.1. Résumé des résultats

Nous nous sommes intéressées dans ces travaux de recherche à l'usage de l'outil stylo numérique pour la recherche en sciences de l'éducation. Nos contributions sont, essentiellement, la proposition d'un modèle de système de traitement et de visualisation des données recueillies avec un stylo numérique ainsi que le développement d'un prototype de système de traitement et de visualisation.

7.1.1. Contribution théorique

En optant pour une méthode inductive, le modèle proposé est issu de l'analyse de prototypes conçus pour traiter les données recueillies dans un contexte expérimental. Notre modèle supporte à la fois les trois phases d'un processus d'analyse (recueil, traitement et visualisation) et les différentes tâches que l'utilisateur pourrait avoir besoin de faire.

Dans le cas d'usage du stylo numérique comme moyen de recueil de données, la phase de collecte des données est considérable du fait que l'instrumentation construit pour un pré-traitement des données. Cette technologie permet à l'utilisateur de coder, selon des objectifs, la trace qui va être collectée. Les prototypes, réalisés en participation avec les utilisateurs, représentent l'interface facilitant à l'utilisateur la manipulation et la visualisation de la trace.

L'analyse des expériences et de l'interface du prototype réalisé a fait émerger essentiellement des tâches de l'utilisateur durant chaque phase du processus d'analyse. Nous avons classé ces tâches en trois catégories ; les tâches d'interaction, les tâches de manipulation et les tâches d'analyse. Ainsi, l'utilisateur peut intervenir dans chaque phase du processus à travers différents types des tâches.

7.1.2. Outil développé

Nous avons produit, en collaboration avec l'utilisateur, pour chaque cas d'usage, des solutions pour réaliser les objectifs de l'expérience. Ainsi, pour chaque expérience, nous avons produit un prototype qui permet de traiter et visualiser les données selon le contexte et l'objectif de l'expérience.

Toutefois, ces prototypes sont globalement complémentaires et comportent des fonctionnalités de base qui sont nécessaires au traitement de la trace du stylo numérique. En effet, le premier prototype implique le processus d'importation de la trace dans une base de données relationnelle et l'édition de quelques graphiques statiques.

Le deuxième prototype implique le processus de calcul des indicateurs temps et la génération de tableaux de données en format xls. Pour le troisième prototype, nous avons introduit plus d'interaction dans l'interface et nous avons intégré la génération des graphiques interactifs.

Finalement, pour le dernier prototype, nous nous sommes concentrée sur l'amélioration de l'interactivité des graphiques et la synchronisation des données traitées avec d'autres types de données temps également importées dans le système.

7.2. Réalisations futures et perspectives

Nos perspectives à court terme visent au premier lieu l'amélioration des fonctionnalités, des représentations et des interactions dans l'interface de la plateforme SEGELL. Puis, une évaluation de l'outil est également nécessaire.

En terme de réalisation informatique, nous avons identifié de nombreuses améliorations sur la plateforme SEGELL. D'abord, dans le but de pouvoir traiter sur la plateforme des traces issues d'autres sources, nous voulons convertir les données temps qui vont être importées en un formatage unique. Ce formatage définit un objet appartenant à une structure prédéfini (formulaire sur papier, formulaire sur le web) avec l'ensemble des données qui lui correspondes dont des données temps (instants de début et de fin, durées).

Nous voulons améliorer les fonctionnalités analytiques comme le filtrage des instances selon un paramètre défini. Ce paramètre pourrait être une valeur d'un champ, une réponse à une question ou bien un paramètre personnalisé.

En terme de visualisation, nous voulons intégrer d'autres techniques de visualisation pour représenter les données en fonctions du temps. De plus, l'espace de représentation dans la plateforme devrait être plus souple pour contenir et gérer des données volumineuses. Pour cela, il sera intéressant d'intégrer des composants graphiques auto-adaptés en fonction du type

et de la taille de la trace.

En terme d'interaction, les améliorations à faire consistent à rendre les graphiques plus interactifs pour pouvoir manipuler l'affichage des séries des données sur les graphiques. Nous voulons également ajouter la possibilité d'exporter un graphique en tableau de données (en format xls par exemple).

Une future réalisation de cette plateforme, qui n'est qu'un prototype pour l'instant, devrait être testé par plusieurs utilisateurs dans différents contextes d'usage. Nous voulons utiliser la plateforme SEGELL pour traiter les données d'autres expériences dans le but de (1) évaluer l'outil et (2) valider le modèle proposé.

Une autre perspective consiste à étudier le choix, notamment par le système, d'une représentation graphique selon le contexte et les objectifs analytiques de l'utilisateur.

Enfin, notre modèle, centré tâches de l'utilisateur, complète les modèles qui existent dans la littérature. Toutefois, l'application de son concept sur un outil permettra de le valider et de faire émerger d'autres dimensions, notamment cognitives, que nous devons tenir compte dans la conception de tel système.

Références bibliographiques

Aigner, W., Miksch, S., Muller, W., Schumann, H., & Tominski, C. (2008). Visual Methods for Analyzing Time-Oriented Data. *IEEE Transactions on Visualization and Computer Graphics*, 14(1), 47 - 60. doi:10.1109/TVCG.2007.70415.

Barhoumi, Z. (2010). L'apport du traitement des données-temps issues d'un stylo numérique pour l'analyse de tests d'évaluation. In *RJC EIAH2010*. Présenté à Rencontres des jeunes chercheurs en EIAH 2010. Consulté à l'adresse <http://hal.archives-ouvertes.fr/hal-00515384>.

Barhoumi, Z., & Tort, F. (2011). Processing and Analyzing Assessment Test Logs provided by Digital Pen and Paper. Présenté à 4th Annual Conference on e-Learning Excellence in the Middle East. Consulté à l'adresse <http://hal.archives-ouvertes.fr/hal-00622813>.

Bassili, J. N., & Fletcher, J. F. (1991). Response-Time Measurement in Survey Research a Method for Cati and a New Look at Nonattitudes. *Public Opinion Quarterly*, 55(3), 331- 346. doi:10.1086/269265.

Bassili, J. N., & Scott, B. S. (1996). Response Latency as a Signal to Question Problems in Survey Research. *Public Opinion Quarterly*, 60(3), 390- 399. doi:10.1086/297760.

Belgnaoui, Y. (2008). Le stylo numérique veille à la fabrication. *MESURES*, *MESURES 801*, 48.

Bertrand, R., Blais, J.-G., & Raïche, G. (2004). *Modèles de mesure: l'apport de la théorie des réponses aux items*. PUQ.

Brotherton, J. A., & Abowd, G. D. (2004). Lessons learned from eClass: Assessing automated capture and access in the classroom. *ACM Trans. Comput.-Hum. Interact.*, 11(2), 121-155. doi:10.1145/1005361.1005362.

Callegaro, M., Yang, Y., Bhola, D. S., & Dillman, D. A. (2004). Response latency as an indicator of optimizing: A study comparing job applicants' and job incumbents' response time on a web survey. In *Recent developments and applications in social research methodology. Proceedings of the RC 33 Sixth International Conference on Social Science Methodology*. Amsterdam. Consulté à cette [adresse](#).

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann.

Chi, E. H. (2000). A Taxonomy of Visualization Techniques Using the Data State Reference Model. In *Proceedings of the IEEE Symposium on Information Visualization 2000* (p. 69–). Washington, DC, USA: IEEE Computer Society. Consulté à cette [adresse](#).

Daassi, C. (2003, juillet 16). *Techniques d'interaction avec un espace de données temporelles*. Université Joseph-Fourier - Grenoble I. Consulté à cette [adresse](#).

Draisma, S., & Dijkstra, W. (2004). Response Latency and (Para)Linguistic Expressions as Indicators of Response Error. In S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, Elizabethrtin, Jeanrtin, & E. Singer (Éd.), *Methods for Testing and Evaluating Survey Questionnaires* (p. 131–147). John Wiley & Sons, Inc. Consulté à cette [adresse](#).

Etche Ogeli, R. (2010). Utilisation des TIC pour les enfants autistes dans un cadre éducatif : Mise en place des indicateurs et des observations suivies instrumentées avec un stylo numérique. In *RJC EIAH2010*. Présenté à Rencontres des jeunes chercheurs en EIAH 2010, Lyon.

Foulin, J.-N. (1995). Pauses et débits : les indicateurs temporels de la production écrite. *L'année psychologique*, 95(3), 483- 504. doi:10.3406/psy.1995.28844.

Guimbretière, F. (2003). Paper augmented digital documents. In *Proceedings of the 16th annual ACM symposium on User interface software and technology* (p. 51–60). New York, NY, USA: ACM. doi:10.1145/964696.964702.

Kobayashi, M., & Koike, H. (1998). EnhancedDesk: integrating paper documents and digital documents. In *Computer Human Interaction, 1998. Proceedings. 3rd Asia Pacific* (p. 57 - 62). Présenté à Computer Human Interaction, 1998. Proceedings. 3rd Asia Pacific. doi:10.1109/APCHI.1998.704149.

Lee, W., De Silva, R., Peterson, E. J., Calfee, R. C., & Stahovich, T. F. (2008). Newton's Pen: A pen-based tutoring system for statics. *Comput. Graph.*, 32(5), 511–524. doi:10.1016/j.cag.2008.05.009.

Liao, C., Guimbretière, F., & Hinckley, K. (2005). PapierCraft: a command system for interactive paper. In *Proceedings of the 18th annual ACM symposium on User interface software and technology* (p. 241–244). New York, NY, USA: ACM. doi:10.1145/1095034.1095074.

MacEachren, A. M. (1995). *How Maps Work: Representation, Visualization, and Design*. New York: Guilford Press. Consulté à cette [adresse](#).

Maldonado, H., Lee, B., & Klemmer, S. (2006). Technology for design education: a case study. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (p. 1067–1072). New York, NY, USA: ACM. doi:10.1145/1125451.1125654.

Miura, M., Kunifuji, S., & Sakamoto, Y. (2007). Practical Environment for Realizing Augmented Classroom with Wireless Digital Pens. In B. Apolloni, R. J. Howlett, & L. Jain (éd.), *Knowledge-Based Intelligent Information and Engineering Systems* (p. 777- 785). Springer Berlin Heidelberg. Consulté à l'adresse http://link.springer.com/chapter/10.1007/978-3-540-74829-8_95.

Miura, M., Kunifuji, S., Shizuki, B., & Tanaka, J. (2005). Augmented Classroom: A Paper-Centric Approach for Collaborative Learning System. In H. Murakami, H. Nakashima, H. Tokuda, & M. Yasumura (éd.), *Ubiquitous Computing Systems* (p. 104- 116). Springer Berlin Heidelberg. Consulté à l'adresse http://link.springer.com/chapter/10.1007/11526858_9.

Muller, W., & Schumann, H. (2003). Visualization methods for time-dependent data - an overview. In *Simulation Conference, 2003. Proceedings of the 2003 Winter* (Vol. 1, p. 737 - 745 Vol.1). doi:10.1109/WSC.2003.1261490.

Mulligan, K., Grant, J. T., Mockabee, S. T., & Monson, J. Q. (2003). Response Latency Methodology for Survey Research: Measurement and Modeling Strategies. *Political Analysis*, 11(3), 289- 301. doi:10.1093/pan/mpg004.

OECD. (2010). *PISA computer-based assessment of student skills in science*. OECD Publishing.

Olive, T., Alves, R., & Castro, S. L. (2008). La production de textes: quels processus sont activés lors des pauses de production et des phases d'écriture? Consulté à l'adresse <http://repositorio-aberto.up.pt/handle/10216/6935>.

Oviatt, S., Arthur, A., & Cohen, J. (2006). Quiet interfaces that help students think. In *Proceedings of the 19th annual ACM symposium on User interface software and technology* (p. 191–200). New York, NY, USA: ACM. doi:10.1145/1166253.1166284.

Pietrzak, T., Malacria, S., & Lecolinet, É. (2010). What do U-Note? An Augmented Note Taking System for the Classroom. Présenté à Papercomp Workshop. Consulté à l'adresse <http://hal.inria.fr/hal-00671476>.

Qin, C., Zhou, C., & Pei, T. (2003). Taxonomy of visualization techniques and systems - concerns between users and developers are different. Présenté à Asia GIS Conference 2003.

- Science & Vie. (2005). Le cahier numérique. *Science & Vie*, septembre 2005(1056), 138.
- Settouti, L. S., Prié, Y., Champin, P.-A., Marty, J.-C., & Mille, A. (2009). A Trace-Based Systems Framework : Models, Languages and Semantics. Consulté à l'adresse <http://hal.inria.fr/inria-00363260>.
- Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In *In IEEE Symposium on Visual Languages* (p. 336–343).
- Signer, B. (2005). *Fundamental Concepts for Interactive Paper and Cross-Media Information Spaces*. L'institut Federal Suisse de Technologie. Consulté à l'adresse http://www.academia.edu/175442/Fundamental_Concepts_for_Interactive_Paper_and_Cross-Media_Information_Spaces.
- Sodoké, K., Raïche, G., & Nkambou, R. (2007). *La plateforme d'évaluation adaptative des apprentissages: PersonFit*. Consulté à l'adresse <http://hal.archives-ouvertes.fr/hal-00161592/en/>.
- Sugihara, T., Miura, M., & Kunifuji, S. (2010). Practicing on Stage: Increasing Transparency and Interaction in Class Activity with Digital Pen system. In R. Setchi, I. Jordanov, R. J. Howlett, & L. C. Jain (éd.), *Knowledge-Based and Intelligent Information and Engineering Systems* (p. 457- 464). Springer Berlin Heidelberg. Consulté à l'adresse http://link.springer.com/chapter/10.1007/978-3-642-15384-6_49.
- Tabard, A., Mackay, W. E., & Eastmond, E. (2008). From individual to collaborative: the evolution of prism, a hybrid laboratory notebook., 569- 578.
- Tort, F., Blondel, F.-M., & Bruillard, E. (2009). From error detection to behaviour observation: first results from screen capture analysis. *arXiv:0908.1580*. Consulté à l'adresse <http://arxiv.org/abs/0908.1580>.
- Ware, C. (2004). *Information Visualization: Perception for Design*. Morgan Kaufmann.
- Weibel, N., Fouse, A., Hutchins, E., & Hollan, J. D. (2011). Supporting an integrated paper-digital workflow for observational research (p. 257). ACM Press. doi:10.1145/1943403.1943443.
- Wellner, P. D. (1993). *Interacting with paper on the DigitalDesk*. University of Cambridge.
- Yeh, R., Liao, C., Klemmer, S., Guimbretière, F., Lee, B., Kakaradov, B., ... Paepcke, A. (2006). ButterflyNet: a mobile capture and access system for field biology research. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (p. 571–580). New York, NY, USA: ACM. doi:10.1145/1124772.1124859.

Annexe A

1.1. Analyse vidéo d'un test court sur papier

Zaara Barhoumi – 3 décembre 2008

1.1.1. Déroulement du test

Le questionnaire comporte 2 parties. Une partie posant des questions descriptives sur l'étudiant (genre, parcours) et ses usages de l'ordinateur. Une partie d'évaluation de compétences tableur (illustré sur le logiciel Excel de Microsoft).

La deuxième partie comporte 5 exercices dont 3 à choix multiple, 2 à choix unique et 1 question ouverte, demandant l'écriture d'une formule. Le temps prévu pour le test est de 15 minutes.

Le test est passé individuellement par 3 étudiants de l'université de Gafsa (Tunisie). 2 étudiants sont inscrits en 3^e année de « Math info », le troisième est inscrit en 3^e année de « gestion d'entreprise ».

L'étudiant est installé sur une table. La caméra dans son dos, en plongée sur la table. Sur le plan de capture, il n'y a que la feuille et la main de l'étudiant.

Un court entretien est réalisé à la fin de chaque test, pour questionner l'étudiant sur le test qu'il vient de passer. Nous avons noté les principales discussions.

1.1.2. Note d'observation

Etudiant 1

Observation

Il commence par lire au début de la première page, puis il répond aux questions au fur et à mesure, dans l'ordre du questionnaire. Il a une difficulté avec l'exercice 3. Il le relit plusieurs fois. Puis il passe à l'exercice 4 puis l'exercice 5. Vers la fin il revient voir l'exercice 3.

Entretien

Quelle est la question la plus difficile ? La question 3, je l'ai passé pour faire la 4ème.

Quelle est ta stratégie pour faire un test ? Je lis chaque question, si c'est faisable je la fais, sinon je passe à la suivante. Mais généralement, dans les examens, je dois lire tout et commencer par la plus facile, comme ils le disent les profs.

Par quel type de question tu commences si tu as plus ou moins du temps ? Les plus faciles.

Le test est-il difficile pour toi ? Oui, un peu.

Etudiant 2

Observation

Il commence par lire la totalité du test. Il prend son temps pour lire attentivement. Il revient ensuite au début du test et pour chaque question, la relie et y répond.

Entretien

Quelle est la question la plus difficile ? L'exercice 5 est le plus difficile

Quelle est la question la plus facile ? Peut être la première mais tout est difficile je trouve.

Quelle est ta stratégie pour faire un test ? Je lis tout l'examen juste pour voir de quoi il s'agit. Après je commence à faire un par un.

Par quel type de question tu commence si tu a plus ou moins du temps ? Par le plus facile

Le test est-il difficile pour toi ? Il n'est pas difficile coté informatique, mais coté forme, je n'ai pas compris.

Etudiant 3

Observation

Il commence par lire la première partie et y répond. Ensuite, il passe à la deuxième partie, lit attentivement chaque question, y répond sinon il passe à la suivante. A la fin, il revient sur les questions passées.

Entretien

Le test est-il difficile pour toi ? Oui c'est difficile

De quel coté logique ou connaissance du logiciel ? Pas dans la logique mais au niveau de la connaissance du logiciel.

Quelle est la question la plus difficile ? L'exercice 3 est le plus difficile parce que je ne sais pas faire une formule de calcul.

Quelle est la question la plus facile ? L'exercice 1 est le plus facile, même si je n'ai pas bien répondu mais quand même le raisonnement est correct.

Quelle est ta stratégie pour faire un test ? Lire chaque question ; si je sens une difficulté je passe. Et je reviens après faire les plus difficiles.

Par quel type de question tu commence si tu a plus ou moins du temps ? Je fais les plus faciles, et normale si je sors sans finir une question difficile. Sans passer du temps à essayer, je sors. Des fois je fais ça, et des fois la réponse d'une question se trouve dans la question suivante. Ça arrive.

1.1.3. Déroulement chronologique (vidéo)

Etudiant 1

Heure	Questions	Gestes	Temps	Stratégies
0:00:00	Partie 1	lecture	0:00:19	Lire et remplir dans la première partie
0:00:19		écriture	0:00:02	
0:00:21		lecture	0:00:16	
0:00:37		écriture	0:00:01	
0:00:38		lecture	0:00:03	
0:00:41		écriture	0:00:01	
0:00:42		lecture	0:00:05	
0:00:47		écriture	0:00:01	
0:00:48		lecture	0:00:13	
0:01:01		écriture	0:00:01	
0:01:02		lecture	0:00:13	
0:01:15		écriture	0:00:01	
0:01:16		lecture	0:00:05	
0:01:21		écriture	0:00:01	
0:01:22		lecture	0:00:10	
0:01:32		écriture	0:00:01	
0:01:33		lecture	0:00:04	
0:01:37		écriture	0:00:01	
0:01:38		lecture	0:00:05	
0:01:43		lecture	0:00:13	
0:01:56	Partie 2 : Q1	lecture	0:00:35	Passer à la deuxième partie, lire et répondre à la question 1
0:02:31		hésitation	0:00:04	
0:02:35		écriture	0:00:01	
0:02:36		lecture	0:00:10	

0:02:46	Partie 1	lecture	0:00:04	Retour sur la première partie	
0:02:50		écriture	0:00:05		
0:02:55		recul	0:00:10		
0:03:05		passer à la deuxième feuille	0:00:05		
0:03:10	Partie 2 : Q2	lecture	0:00:45	Lire et répondre	
0:03:55		écriture	0:00:01		
0:03:56		recul	0:00:03		
0:03:59	Q3	lecture	0:00:56	Difficulté rencontré	
0:04:55	Q4	lecture	0:02:01	Lire et répondre	
0:06:56		écriture	0:00:01		
0:06:57	Q5	lecture	0:02:15		
0:09:12		écriture	0:00:01		
0:09:13		lecture	0:00:03		
0:09:16		écriture	0:00:01		
0:09:17		lecture	0:00:03		
0:09:20		écriture	0:00:01		
0:09:21		recul	0:00:03		
0:09:24		écriture	0:00:01		
0:09:25		lecture	0:00:09		
0:09:34	Q3	lecture	0:01:18		Retour sur la difficulté
0:10:52	fin				

Etudiant 2

Heure	Questions	Gestes	Temps	Stratégies
0:00:00		lecture	0:01:47	Lire toute la feuille 1
0:01:47		passer à la deuxième feuille	0:00:04	

0:01:51		lecture	0:01:15	Lire toute la feuille 2
0:03:06		Revenir sur la première feuille	0:00:08	
0:03:14	Partie 1	écriture	0:00:06	Lire et répondre à la première partie
0:03:20		lecture	0:00:09	
0:03:29		hésitation	0:00:02	
0:03:31		écriture	0:00:07	
0:03:38		recul	0:00:04	
0:03:42		hésitation	0:00:03	
0:03:45		écriture	0:00:09	
0:03:54		lecture	0:00:19	
0:04:13		écriture	0:00:01	
0:04:14		lecture	0:00:13	
0:04:27		écriture	0:00:01	
0:04:28		lecture	0:00:05	
0:04:33		écriture	0:00:01	
0:04:34		lecture	0:00:02	
0:04:36		écriture	0:00:01	
0:04:37		lecture	0:00:05	
0:04:42	Partie 2 : Q1	lecture	0:00:25	Lire et répondre
0:05:07		hésitation	0:00:13	
0:05:20		hésitation	0:00:03	
0:05:23		écriture	0:00:01	
0:05:24		passer à la deuxième feuille	0:00:06	
0:05:30	Q2	lecture	0:00:28	
0:05:58		écriture	0:00:02	
0:06:00		recul	0:00:03	
0:06:03	Q3	lecture	0:00:46	Difficulté rencontrée
0:06:49	Q4	lecture	0:01:26	

0:08:15		écriture	0:00:01	Lire et répondre
0:08:16		recul	0:00:03	
0:08:19	Q5	lecture	0:01:17	Difficulté rencontrée
0:09:36	Q3	lecture	0:01:24	Retour sur la première difficulté
0:11:00		hésitation	0:00:09	Lire et répondre après des hésitations
0:11:09		hésitation	0:00:06	
0:11:15		lecture	0:00:14	
0:11:29		hésitation	0:00:11	
0:11:40		lecture	0:00:26	
0:12:06		écriture	0:00:15	
0:12:21		recul	0:00:19	
0:12:40	Q5	lecture	0:01:11	Retour sur la deuxième difficulté
0:13:51	fin			

Etudiant 3

Heure	Questions	Gestes	Temps	Stratégies
0:00:00		commencer	0:00:20	Lire et répondre
0:00:20	Partie 1	écriture	0:00:05	
0:00:25		lecture	0:00:08	
0:00:33		écriture	0:00:02	
0:00:35		lecture	0:00:05	
0:00:40		hésitation	0:00:03	
0:00:43		lecture	0:00:44	
0:01:27	Partie 2 : Q 1	lecture	0:00:19	
0:01:46		écriture	0:00:03	

0:01:49		passer à la deuxième feuille	0:00:06	Lire et répondre
0:01:55	Q2	lecture	0:00:50	
0:02:45		hésitation	0:00:03	
0:02:48		écriture	0:00:01	
0:02:49	Q3	lecture	0:00:36	Deux difficultés rencontrées
0:03:25	Q4	lecture	0:01:20	
0:04:45	Q5	lecture	0:01:33	Lire et répondre
0:06:18		écriture	0:00:04	
0:06:22	Q4	lecture	0:00:26	Retour sur la 1ere difficulté Lire et répondre après des hésitations
0:06:48		hésitation	0:00:03	
0:06:51		hésitation	0:00:26	
0:07:17		écriture	0:00:03	
0:07:20	Q3	lecture	0:01:07	Retour sur la 2ème difficulté
0:08:27	Partie1	Revenir sur la première feuille	0:00:06	
0:08:33		lecture	0:01:05	Retour sur la première partie
0:09:38		écriture	0:00:01	
0:09:39		lecture	0:00:22	Lire et répondre
0:10:01		écriture	0:00:01	
0:10:02		lecture	0:00:12	
0:10:14		écriture	0:00:01	
0:10:15		lecture	0:00:18	
0:10:33		écriture	0:00:01	
0:10:34		lecture	0:00:11	
0:10:45		écriture	0:00:02	
0:10:47		lecture	0:00:14	
0:11:01	fin			

Annexe B

Dans cette annexe nous présentons l'aspect technique du prototype SegellBeta. L'architecture de notre application se base sur le modèle MVC (modèle, vue, contrôleur) (voir Figure 101).

Habituellement, un système de visualisation est développé sous forme de logiciel qui nécessite d'être installé sur un poste de bureau. Nous avons choisi de développer le prototype de notre système sous forme d'application web RIA (Rich Internet Applications ou Applications Internet Riches) disponible sur Internet et permettant d'enrichir l'interaction avec l'utilisateur.

Les applications web riches permettent de fluidifier les manipulations utilisateurs courantes (onglet, menu, tables de données arbre et saisie). Dans une RIA, il n'est plus nécessaire de recharger la page à chaque action utilisateur. Le modèle de navigation de page en page n'existe donc plus, et nous parlerons plutôt d'écran.

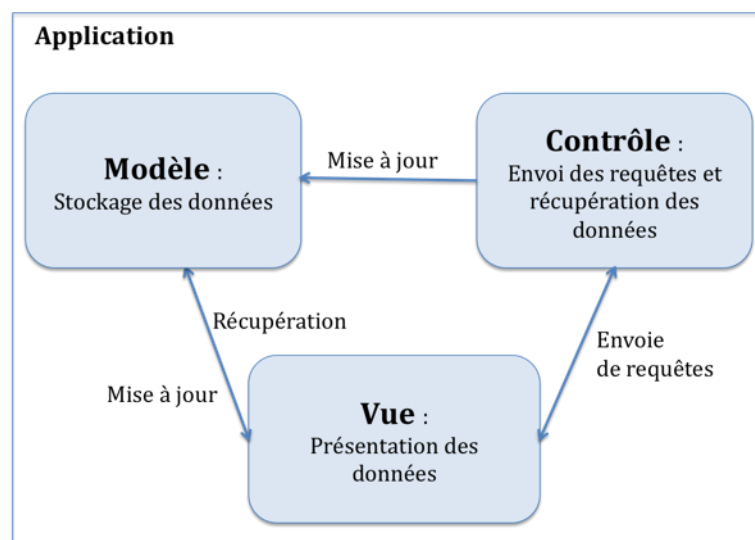


Figure 101 : la modélisation de l'application en MVC.

En terme de langage, nous avons développé cette application avec les technologies récentes du développement web. Nous avons utilisé PHP pour le traitement des indicateurs temps et

pour la communication avec la base des données à travers des requêtes. Et nous avons utilisé le Framework libre Flex¹⁸ pour l'application RIA. Nous avons donc installé un plug-in de ce Framework sur Eclipse.

Dans le tableau suivant, nous présentons trois données métriques permettant de quantifier le code du prototype SegellBeta : le nombre de packages, le nombre de classes et le nombre de lignes de code. Nous distinguons les lignes de code développées en PHP et celles développées en ActionScript et en MXML à travers le Framework Flex.

	Nombre de packages	Nombre de classes	Lignes de code MXML	Lignes de code AS	Lignes de code PHP
Contrôle PHP	1	8	-	-	1731
Contrôle AS	1	9	-	930	-
Modèle	1	9	-	469	-
Vue	5	11	4694	973	-
Rendu	1	1	454	78	-
Utilitaires AS	3	15	-	2080	-
Utilitaires PHP	-	-	-	-	435
Total	12	53	5148	4530	2166

Tableau 6 : Données métriques du prototype SegellBeta.

De plus, nous avons opté pour l'usage d'une micro-architecture interne sous Flex nommé Reflex¹⁹, une bibliothèque contenant les principales classes qui orchestrent la communication entre les modèles, les vues et les contrôles dans une application Flex en MVC.

¹⁸ http://fr.wikipedia.org/wiki/Adobe_Flex

¹⁹ www.lafabrick.com/labz/reflex/tuto/ReflexTuto1.pdf

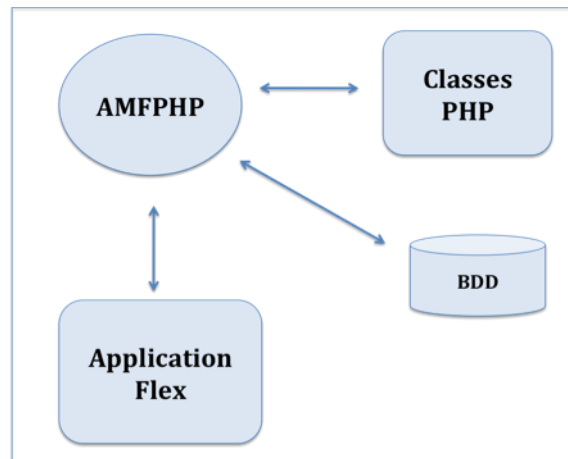


Figure 102 : le fonctionnement de la passerelle AMFPHP.

Pour les échanges serveurs, nous avons utilisé une technologie intermédiaire, AMFPHP, qui sert de passerelle entre l'application Flex et les données à travers des traitements en PHP. Tout objet envoyé depuis l'application RIA vers le service PHP sera sérialisé en type qui correspond à PHP. Il s'agit simplement d'une « passerelle » facilitant les échanges serveurs et client Flash/Flex et PHP.

Table des illustrations

Les figures

Figure 1 : Processus d'utilisation des interfaces et technologies à interaction stylo.....	8
Figure 2 : Autres exemples de dispositif de capture de l'écriture sur papier.....	9
Figure 3 : La technologie Anoto du papier et stylo numériques.....	10
Figure 4 : Un extrait de données encre numérique représenté en InkML.....	12
Figure 5 : Représentation d'une pointe d'un stylo dans la matrice de la trame d'Anoto.....	14
Figure 6 : Le stylo numérique Logitech Io2.....	16
Figure 7 : Le fonctionnement de la technologie Anoto avec les services de Kayentis.....	16
Figure 8 : L'infrastructure des PADD (Guimbretière, 2003).....	20
Figure 9 : Extrait des commandes traitées par PaperCraft pour l'annotation de document (Liao et al., 2005).....	21
Figure 10 : L'interface du navigateur de l'outil ButterflyNet.....	25
Figure 11 : L'interface de ChronoViz : l'annotation numérique s'affiche sur la frise temporelle en bas à gauche, et la représentation numérique du papier à droite (Weibel et al., 2011).....	26
Figure 12 : Méthode de travail.....	30
Figure 13 : Illustration de la zone « identifiant » depuis le premier test destiné aux étudiants de L3 Economie et Gestion de l'ENS Cachan.....	38
Figure 14 : un exemple de question ouverte depuis le premier test destiné aux étudiants de L3 Economie et Gestion de l'ENS Cachan.....	38
Figure 15 : exemple de question fermé à choix unique depuis le premier test destiné aux étudiants de L3 Economie et Gestion de l'ENS Cachan.....	39
Figure 16 : exemple de question fermé à choix multiple depuis le deuxième test destiné aux étudiants de L3 Economie et Gestion de l'ENS Cachan.....	39
Figure 17 : illustration depuis le premier test de C2i des cases à cochés pour marquer la sureté à propos des réponses aux questions.....	40
Figure 18 : La succession des FS et des LS dans le temps.....	44
Figure 19 : Ecran de gestion des indicateurs illustrant l'indicateur « compétence » et les catégories y associées pour le premier test tableur – L3 économie et gestion de l'ENS Cachan.....	48
Figure 20 : Ecran de gestion des indicateurs illustrant l'indicateur « difficulté » et les catégories y associées pour le premier test tableur – L3 économie et gestion de l'ENS Cachan.....	48
Figure 21 : Ecran de gestion des indicateurs illustrant l'indicateur « type de question » et les catégories y associées pour le premier test tableur – L3 économie et gestion de l'ENS Cachan.....	49
Figure 22 : Ecran de classification des champs de réponses par indicateurs pour le premier test tableur – L3 économie et gestion de l'ENS Cachan.....	49
Figure 23 : Ecran de personnalisation de la valeur « libellé » de champ de réponse et la valeur « ordre » pour le premier test tableur – L3 économie et gestion de l'ENS Cachan.....	50
Figure 24 : Un extrait d'un fichier xls contenant les données du stylo numérique de deux répondants au premier test tableur – L3 économie et gestion de l'ENS Cachan.....	51
Figure 25 : Tableau de réponse pour chaque individu extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.....	52
Figure 26 : Conception du graphique chronologique du parcours de réponse d'un répondant. Les numéros des questions qui sont en ordonnée permettent de distinguer l'ordre de traitement du test par le répondant.....	53
Figure 27 : Représentation multiple illustrant : (1) la progression chronologique d'un répondant, (2) le résultat de sa production, (3) le niveau de difficulté des questions, (4) la classification des questions selon les compétences, (5) le temps moyen passé sur chaque question.....	54
Figure 28 : Interface interactive qui illustre la représentation personnalisée du graphique multiple.....	55
Figure 29 : un extrait de fichier xls contenant les instants de fin d'écriture correspondant à tous les questions pour chaque répondant au deuxième test de C2i.....	56
Figure 30 : interface permettant d'afficher les données temps et les filtrer selon un indicateur spécifique	

extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.....	57
Figure 31 : représentation en nuage de points de la dispersion des données temps FS pour chaque question par tous les répondants extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.....	58
Figure 32 : représentation en lignes de la dispersion des données temps FS pour chaque question par tous les répondants extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.....	59
Figure 33 : représentation en histogramme de la durée moyenne d'écriture dans chaque question par rapport à l'indicateur « compétence » extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.....	59
Figure 34 : représentation en histogramme de la durée moyenne de non écriture dans chaque question par rapport à l'indicateur « niveau de réussite » extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.....	60
Figure 35 : représentation en histogramme de la durée moyenne totale (écriture + non écriture) dans chaque question par rapport à l'indicateur « type de question » extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.....	60
Figure 36 : représentation en pourcentage empilé de la durée totale passée dans chaque question par répondant, extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.....	61
Figure 37 : représentation en histogramme du taux de réussite pour chaque question, extrait du premier test tableur – L3 économie et gestion de l'ENS Cachan.....	61
Figure 38 : Exemple de zone d'écriture qui n'est pas validé par le système de reconnaissance.....	66
Figure 39 : Le contenu de la première page qui contient l'énoncé et la position du problème.....	70
Figure 40 : une zone pour notifier le niveau de familiarité avec la classification Rajalingham.....	70
Figure 41 : L'arbre de classification selon Rajalingham et al (2000) révisé par Purser (2006) http://www.eusprig.org/2006/purser-awareness-types-spreadsheet-errors.pdf	71
Figure 42 : La nouvelle disposition de l'arbre de classification en vertical.....	72
Figure 43 : Les zones de saisie sur chaque cercle au-dessous de la classe qui lui correspond.....	72
Figure 44 : Le nombre de personnes répartie selon leur familiarité avec la classification Rajalingham.....	75
Figure 45 : les durées globales passés dans chaque question par tout les participants.....	75
Figure 46 : les parcours chronologiques de réponses aux questions de 1 à 5 dans l'axe des ordonnées, en fonction du temps dans l'axe des abscisses, de tous les répondants.....	77
Figure 47 : les durées moyennes passées par chaque catégorie de répondants.....	78
Figure 48 : des parcours variés qui n'ont pas respecté un ordre habituel.....	80
Figure 49 : Instrumentation de la première page du test RJC.....	83
Figure 50 : Instrumentation de la deuxième page du test RJC ; les cases à cocher alignées sur la carte.....	84
Figure 51 : répartition des répondants par taux de réussite.....	87
Figure 52 : répartition du taux de réussite par pays.....	87
Figure 53 : Pour chaque image, le nombre de répondants l'ayant identifiée en premier.....	88
Figure 54 : Pour chaque image, le nombre de répondants l'ayant identifiée en dernier.....	88
Figure 55 : Parcours du répondant RKI04.....	89
Figure 56 : Parcours du répondant RTS05.....	90
Figure 57 : (a) Répartition des réponses des répondants sur la carte, (b) les bonnes réponses.....	90
Figure 58 : un model instrumenté de conception de la grille d'observation.....	97
Figure 59 : Le model finale de conception de la grille d'observation.....	98
Figure 60 : Une grille d'observation remplie.....	99
Figure 61 : illustration de la dispersion des actions observées dans le temps à droite, et à gauche le nombre d'occurrence pour chaque action.....	101
Figure 62 : Représentation de la séance d'observation Ousmane du 14-01-2011.....	102
Figure 63 : Représentation de la séance d'observation Ousmane du 21-01-2011.....	102
Figure 64 : Représentation de la séance d'observation Ousmane du 04-02-2011.....	103
Figure 65 : Représentation des actions, notées par le stylo numérique, synchronisées avec les données de transcription.....	104
Figure 66 : Illustration d'une partie zoomée de la représentation des données synchronisées.....	105
Figure 67 : le fonctionnement global de SEGELL.....	110
Figure 68 : Interface de Bienvenue de SEGELL et zone d'authentification.....	111
Figure 69 : L'écran d'importation des formulaires.....	112
Figure 70 : L'écran d'importation de données de transcription pour notes d'observation.....	113
Figure 71 : Ecran de gestion d'indicateurs.....	113

Figure 72 : Ecran des classifications des champs par indicateurs.	115
Figure 73 : Ecran de gestion des champs.....	116
Figure 74 : Ecran de gestion des champs (cas de formulaire de notes d'observation).....	117
Figure 75 : Ecran de visualisation des données.....	118
Figure 76 : Fenêtre de tableau de données de réponses d'un individu.	119
Figure 77 : Fenêtre de représentation graphique composé et paramétrable.	120
Figure 78 : Fenêtre paramétrable de génération de tableaux pour toutes les données.	122
Figure 79 : Fenêtre de tableau des données contenant le résultat de la sélection.	124
Figure 80 : fenêtre de représentation graphiques pour l'ensemble des données.	125
Figure 81 : Représentation en forme de ligne de la dispersion des données temps.....	126
Figure 82 : Représentation de la dispersion des données temps : zoom sur les 5 premières minutes.....	126
Figure 83 : Représentation de la dispersion des données temps : zoom sur les 5 dernières minutes.	127
Figure 84 : représentation des durées moyennes.	127
Figure 85 : Représentation des durées moyennes : résultat de choix d'indicateur.	128
Figure 86 : Représentation de la distribution des durées de réponse au champs par chaque répondant.	129
Figure 87 : Représentation graphique du taux de réussite au test.	130
Figure 88 : Regroupement des instances en « séance d'observation ».....	131
Figure 89 : Représentation graphique de la dispersion et de la fréquence d'actions observées durant une séance d'observation.	132
Figure 90 : des instances avec des données de transcription.....	133
Figure 91 : Représentation de la dispersion des données stylo numérique synchronisées avec des données de transcription.....	134
Figure 92 : L'architecture globale de l'application SEGELL.	135
Figure 93 : Extrait du fichier XML ; la balise 'FormInstance' et la balise 'Current'.....	137
Figure 94 : Extrait du fichier XML; la balise 'Page'.	137
Figure 95 : Extrait du fichier XML : la balise 'Field' d'un champ de saisie.....	138
Figure 96 : Extrait d'un fichier XML de description de la structure d'un formulaire.....	138
Figure 97 : Le modèle physique des données.....	139
Figure 98 : Modèle de référence pour la visualisation des données présenté par (Card et al., 1999).....	153
Figure 99 : schéma des étapes de visualisation proposé dans (Ware, 2004).....	154
Figure 100 : Représentation de notre modèle en fonction du processus de l'analyse d'une part, et en fonction des tâches de l'utilisateur d'autre part.....	156
Figure 101 : la modélisation de l'application en MVC.	175
Figure 102 : le fonctionnement de la passerelle AMFPHP.....	177

Les tableaux

Tableau 1 : un extrait des données traitées et exportées en format xls pour le chercheur.....	100
Tableau 2 : Les caractéristiques du prototype déterminées depuis le scénario d'utilisation regroupées selon les cinq facteurs	143
Tableau 3 : Les tâches et les éléments graphiques de l'interface.	145
Tableau 4 : les représentations graphiques pour chaque tâche d'analyse.....	146
Tableau 5 : les indicateurs temps.....	158
Tableau 6 : Données métriques du prototype SegellBeta.	176