

Efficiencce et représentation temporelle sur papier

Préparé pour : Mme Guénand, M. Githalla et M. Gapenne
Préparé par : Émilie Coquard, Jean-François Mougnot

16 janvier 2013

Table des matières

Introduction	1
Mise en contexte.....	1
Questionnement.....	1
Hypothèses.....	2
Expérimentation	3
Dispositifs.....	3
Outils de recueil	4
Résultats	4
Interprétation.....	5
Conclusion	5
Perspective.....	5
Pièces jointes	7
Bibliographie.....	8
Ressources littéraires	8
Ressources web	8

Introduction

Mise en contexte

La data-visualization est une représentation de données sous une forme visuelle, graphique, de dessins ou de cartes. Elle doit reprendre précisément les données et les montrer de manière claire et évidente pour le lecteur. Le but étant que ces données, par leur représentation soient le plus intelligible possible. Cela peut se concrétiser par des graphiques, des camemberts, des diagrammes, des cartographies, des chronologies, des infographies ou même des créations graphiques inédites ou des photos. La présentation sous une forme illustrée rend les données plus lisibles et compréhensibles.

Au delà de l'illustration, on peut même aller jusqu'à présenter des données sous la forme d'animations, voire de vidéos. La plus célèbre d'entre elles est le Billion-Dollar-A-gram réalisée par David McCandless le maître en la matière.

La data-visualization commence toujours par une collecte de données et de chiffres bruts. C'est parfois l'opération la plus délicate car il n'est pas toujours aisé d'obtenir les données que l'on recherche. Une fois les données brutes sélectionnées, traitées et formatées dans un tableur par exemple, il appartient au journaliste, à l'infographiste ou au designer d'imaginer la meilleure façon de les présenter.

De nombreux outils permettent alors de mettre en forme ces données : logiciels (à commencer par le célèbre Illustrator d'Adobe), applications et services web, librairies javascript etc. On peut même réaliser des data-visualizations sommaires avec Excel et son outil de génération de graphiques et diagrammes.

Cette visualisation doit respecter quelques règles :

Montrer les données. Amener à réfléchir autour des idées et non pas au éléments graphiques ou de production du support. Ne pas déformer le dire des données. Présenter le plus d'informations possibles dans un espace réduit. Faire en sorte qu'elles soient cohérentes entre-elles. Encourager les lecteurs à comparer les différentes parties de ces données. Révéler plusieurs niveaux de détail, global dans un premier temps et plus fin dans un second. Servir un but clair et précis avec description, exploration, classification et décoration. Intégrer les descriptions statistiques et verbales des données.

Ces règles sont issues du livre de Edward R. Tufte, *The Visual Display of Quantitative Information*, second édition.

Questionnement

Dans cette partie, nous allons nous interroger sur les différents supports de data-visualization, bien que notre expérimentation se concentre exclusivement sur le support papier. Nous nous demandons si un dispositif interactif intègre plus le lecteur, l'immerge davantage et donc lui apporte plus d'informations en le rendant acteur qu'un dispositif où il serait juste spectateur (poster, vidéo).

Aura-t-il un ressenti plus subjectif dans un dispositif interactif ? Se sentira-t-il plus impliqué ? Aura-t-il pu prendre un point de vue personnel ? Créer sa propre lecture ?



Un aspect primordial de notre expérimentation sera de limiter le déterminisme individuel en s'assurant que tous les sujets possèdent les mêmes connaissances en matière de représentation de l'information, de design, de graphisme et d'esthétisme. Nous essayerons de limiter au maximum cela. Cette tâche est à la fois primordiale et très complexe car il est difficile de jauger et d'estimer la culture générale d'un sujet par rapport à un autre. Il s'agit là d'une variable très forte.

Différents types de dispositif :

- Le poster : Une image fixe, le regard est alors attiré par des lignes et des zones colorées. Dans notre société occidentale, la lecture se fait de gauche à droite et de bas en haut (formant un "Z"). Le sujet n'a que très peu d'interactions possible avec ce dispositif, il est guidé et est totalement passif face à celui-ci.
- La timeline : La timeline représente l'évolution de données dans le temps. Le sujet à une certaine possibilité de manipulation, il peut revenir en arrière pour observer en détail les changements s'appliquant au passage du curseur. Il s'agit alors d'une activité perceptive capturée puisqu'il agit sur ce qu'il voit. Il est acteur de sa compréhension.
- La vidéo : Elle possède un sens. De part sa lecture unilatérale (pas de retour possible) mais aussi par sa forme (son montage). Ce sens est déjà créée par une tierce personne, de son point de vue, limitant les interactions possibles par le sujet avec celle-ci, allant jusqu'à guider la perception de celle-ci. Le temps "d'exposition" est alors limité, la capture des informations par le sujet est très limité puisqu'on agit à sa place.
- L'interface permettant une manipulation des données : Ce dispositif est celui qui possède la plus grande liberté d'interactions. Le sujet à un grand champ de liberté. Cependant les actions faites par le sujets sont dépendants de son expérience, ainsi que des actions qu'il aura eu l'occasion de réaliser avec un tel dispositif.

D'une manière générale, le sens des choses est dans la contrainte, une trop grande liberté peut nuire à la compréhension. Parce que nous sommes guidés par des symboles, des images, des lignes et des stéréotypes d'usages, nous comprenons. Nous avons là des dispositifs possédant différent degré de liberté, liberté modulable en plusieurs points.

Le principal support de représentation de data-visualization est une représentation figée, subjective, créé et interprétée par une personne et soumise aux utilisateurs (aujourd'hui car il tend à devenir de plus en plus interactif pour être immersif et manipulable, et alors mieux compris par le public grâce à la manière ludique d'apporter les infos). Le plus souvent elle prend la forme d'un poster ou d'une image numérique fixe (type .jpg, .png & co).

À l'heure actuelle, on subit un poster, on est simplement un spectateur de celui-ci, nuancer par la personne qui éventuellement le présente. C'est un support très limité (dans la compréhension d'un sujet, message orienté car issu d'un point de vue, mais libre dans la lecture).

Une première partie du travail a consisté à identifier les types de représentation du temps existantes. (cf. document joint «Representation_temps.pdf»). Nous avons alors identifier trois représentations principales et pertinentes pour notre expérimentation : circulaire, linéaire ou par couches ou claques.

Hypothèses

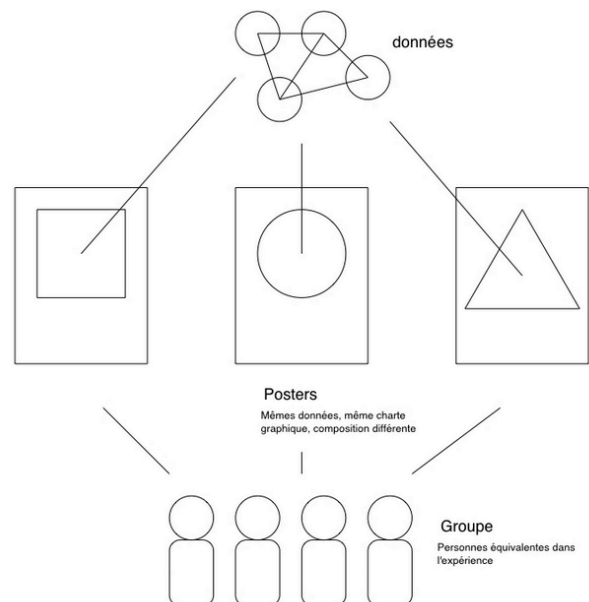
La représentation du temps en superposition permet une meilleure mémorisation des données car elle implique plus le récepteur par la manipulation des calques.

La représentation linéaire du temps sera plus pertinente que la représentation circulaire, car nous avons comme usage, en occident, de représenter et de lire le temps de cette façon.

Expérimentation

Dispositifs

Notre protocole expérimentale sera composé de trois salles avec une horloge dans chaque salle. Chaque salle comprendra un type de poster sur lequel sera affiché des données ainsi qu'une forme de visualisation de celles-ci. Dans un souci de neutralité, ces données seront les mêmes sur chacun des dispositifs, leur traitement graphique aussi. Afin que l'expérimentation soit la plus juste possible, nous aurons besoin de trois sujets possédant la même expérience de la data-visualization (information, design, graphisme et sensibilité esthétique) afin de limiter les variations inter-individuelles ; nous présenterons ainsi à chacun de ces sujets un et un seul dispositif pendant 10 minutes ou moins sur demande du sujet testé. À l'issue de ce temps chaque sujet sortira de sa salle et devra répondre à un rapide questionnaire. Nous voulons ainsi jugé la pertinence des différentes représentation choisies en matière de compréhension de l'information, en s'appuyant sur les ressentis, l'interprétation et ce qu'en aura retenu chaque sujet.



L'étape suivante a été de réaliser les trois posters reprenant chacun une des représentations temporelles précédemment identifiées.





Outils de recueil

Pour recueillir les impressions et essayer de valider ou falsifier nos hypothèses, nous avons mis en place un questionnaire comprenant deux parties. La première partie de celui-ci est composée de 10 questions (cf. document joint «questionnaire_ICX03_corrige.pdf») portant sur le dispositif, son contenu et sa forme. De difficultés différentes, afin d'appréhender le degré d'exploration et d'appropriation des données offert par chaque dispositif, les questions choisies ne possèdent qu'une et unique bonne réponse. Certaines ont nécessité la mise en place de critère d'acceptation de notre part (citer 5 villes sur 7 par exemple suffira). La deuxième partie, mise en place dans l'optique du long terme, a pour but d'établir les profils «psychologique» et personnel afin de pouvoir identifier, le cas échéant, des variables pouvant influencer les réponses chez des individus présentant des similarités. Sur les dix premières questions posées, nous avons fait une moyenne pour comparer l'efficacité de chaque poster et ainsi constater si notre hypothèse de départ peut être validée. C'est à dire, si la représentation temporelle en superposition ou impliquant une action du récepteur permet une meilleure mémorisation.

Résultats

Résultats en pourcentage pour la première partie du questionnaire en fonction du type de poster (circulaire, linéaire, ou en couches) sur 15 personnes.

	groupe 1 : circulaire	groupe 2 : linéaire	groupe 3 : couches
	30	70	70
	50	60	90
	40	60	80
	40	70	70
	85	65	
		70	
Total :	49 %	65,83 %	77,5 %

Interprétation

Conclusion

D'une manière générale, il est ressorti de cette étude statistique succincte que la représentation temporelle par couches est celle qui favorise le mieux la compréhension et l'appropriation des données sur support papier (feuille, affiche ou poster). Les différents profils testés avec ce type de représentation ont répondu jusqu'à 50% de bonnes réponses supplémentaires par rapport à la représentation circulaire (77,5% contre 49%). Ces profils sont différents et variés et aucune similarité entre eux, ni point commun ne semblent avoir influencer ce pourcentage de bonnes réponses. Cette représentation, en plus de proposer une vision d'ensemble des données, semble tout à fait pertinente car elle permet une interactivité des données visuellement présentes. En effet, on peut simplement prendre la feuille faisant office de base et y ajouter un seul calque, permettant une lecture claire et précise des informations choisies. Elle permet également une forme de comparaison qu'elle seule semble offrir. En ajoutant deux calques, les données peuvent alors être comparées ou accumulées. Les utilisateurs ont le choix dans l'affichage de celles-ci, ils peuvent personnaliser celui-ci et donc se projeter et s'approprier plus facilement les informations présentées.

La représentation circulaire à l'inverse semble être celle qui est la plus compliquée à comprendre. Avec seulement 49% de bonnes réponses sur l'ensemble des profils testés avec cette représentation, il s'agit là du score le plus faible. Sa forme graphique peu commune, ainsi que la difficulté à comparer les données semble en être la source. En plus de devoir découvrir et comprendre les données, les personnes testées sont toutes passées par une phase d'apprentissage compliquant la situation.

La représentation linéaire, plus classique et répandue en occident, ne propose pas non plus ce type d'interaction mais permet tout de même une comparaison facilitée et obtient donc un score de bonnes réponses (65,83%), s'intercalant entre la représentation circulaire et la représentation par couche. Il s'avère en réalité que nous avons démontré la pertinence de la manipulation et de l'interaction lors des phases de découverte, d'apprentissage et d'appropriation de données.

Cette première expérimentation nous a donc permis de valider nos deux hypothèses.

Perspective

Cette expérience met en avant l'interactivité comme moyen d'appréhender les données, de les manipuler et de les mémoriser plus facilement. Il serait intéressant d'appliquer la représentation par couche à un dispositif linéaire ou circulaire, afin d'en mesurer l'action effective. Il semblerait aussi pertinent d'étudier la transparence dans le dispositif à couches pour mesurer la aussi son impact. Pour se faire il suffirait de reproduire notre expérience en proposant des calques opaques. C'est pourquoi, il serait envisageable de refaire une telle expérience en confrontant trois supports de visualisation des données. La confrontation de dispositifs fixes (posters), animés (vidéos) ou interactifs permettrait de confirmer ou non notre hypothèse de départ.

Si celle-ci tend à se confirmer, nous pourrions alors dire que la manipulation des données nous permet de mieux les comprendre, les mémoriser et donc de les transformer en savoir.



Il serait tout à fait logique d'envisager de nouveaux outils prenant en compte ses résultats afin de les rendre plus pertinents. Pour être plus exhaustif, il faudrait augmenter le nombre de profils testés ainsi que les diversifier encore plus. Cela permettrait alors d'affirmer ces premières observations de façon quasi-certaine et d'identifier les facteurs variants d'un profil à un autre pouvant influencer sur la compréhension des données., en exploitant la deuxième partie de notre questionnaire. Notre travail pourrait faire office de base à la rédaction de règle permettant de maximiser l'impact de ce genre de donnée sur support papier, chose inexistante à l'heure actuelle. En effet, lors des différents tests, de nombreuses réactions nous ont été partagées. D'une manière générale, les personnes s'attendent à trouver le titre en haut, cherchent une échelle (cf. représentation linéaire) pour quantifier les données et veulent ainsi aller plus loin que l'ordonnement offert par notre poster. Celui-ci étant dépourvu d'échelle, il est difficile de connaître le nombre de passagers en se basant simplement sur la représentation visuelle des données (représentées par des bonhommes dans notre cas) puisqu'il n'y a pas d'échelle annotée. Il aurait aussi été intéressant de relever le temps d'exposition pour chaque sujet, afin de mesurer l'impact de cette variable sur notre problématique.



Pièces jointes

- Nos recherches sur les formes de représentations temporelles : «Representations_temps.pdf»
- Le poster utilisé lors de la présentation : «PosterFinal.pdf»
- Nos trois affiches : «PostersA4.pdf»
- Le questionnaire corrigé : «questionnaire_ICX03_corrige.pdf»

Bibliographie

Ressources littéraires

- The Visual Display of Quantitative Information, Edward R. Tufte
- [Sketching user experience](#), Saul Greenberg, Sheelagh Carpendale, Nicolai Marquardt, Bill Buxton
- [A comparison of two methods of scaling on form perception via a haptic interface](#), Mounia Ziat, Olivier Gapenne, John Stewart, Charles Lenay.

Ressources web

- [Smashing magazine](#)
- [datavisualization.fr](#)