

Les classements des universités : mise en perspective critique **(Version provisoire)**

Fouad FERHAT
CES-LAEP, université Paris 1 Panthéon - Sorbonne
Juin, 2008

Sous la direction de Monsieur le professeur :
François FACCHINI

Mots-clés : Classements des universités. Evaluation de la qualité dans l'enseignement supérieur. Evaluation de la qualité d'enseignement. Evaluation de la qualité de recherche.

Résumé :

Dans ce texte, on propose une mise en perspective critique des classements des universités. On réponds aux questions suivantes : Quels sont les différents types de classements des universités ? Comment les classements des universités sont ils créés ? Quels sont les indicateurs utilisés par les classements pour évaluer la qualité des universités ? Et quelles sont les critiques adressées aux classements ?

Plan :

Introduction

Problématique :

« Dans quelle mesure les classements des universités permettent d'évaluer la qualité dans l'enseignement supérieur ? »

1) Les différents types de classements des universités

2) La construction des classements des universités

3) Une analyse des indicateurs les plus utilisés dans les classements des universités

4) Est-ce qu'il existe une définition universelle de la qualité des universités ?

5) Les indicateurs bibliométriques et évaluation de la qualité de recherche

6) Les critiques adressées aux classements des universités

7) Les réactions des parties prenantes envers les classements des universités et le rôle des classements dans le choix des universités par les étudiants

Conclusion

Bibliographie

Introduction :

Les classements des universités sont apparus dans les années 80 dans un contexte de massification d'enseignement supérieur, des problèmes de fiscalité publique et d'augmentation des droits d'inscriptions. Les consommateurs (étudiants, parents) et les autres parties prenantes (gouvernements, employeurs) demandent plus d'informations et de transparence aux universités, ce qui a créé une demande pour les classements des universités. Un autre contexte lié à l'apparition des classements, notamment les classements mondiaux des universités est celui de l'importance accrue du marché global international d'enseignement supérieur qui est en forte expansion, le nombre des étudiants étrangers dans les pays d'OCDE a doublé entre 1980 et 1999 (1,47 millions), ces étudiants étrangers dépensent chaque année dans les pays OCDE 30 milliards de \$ qui comprennent les droits d'inscriptions et les coûts de vie universitaire. Ca devient de plus en plus intéressant pour les pays et les universités d'attirer ces étudiants étrangers. (Dill et Soo, 2004, p. 3).

Dans ce texte, on propose une mise en perspective critique des classements des universités, on essaiera de répondre aux questions suivantes : Quels sont les différents types de classements des universités ? Comment les classements des universités sont ils créés ? Quels sont les indicateurs utilisés par les classements pour évaluer la qualité des universités ? Et quelles sont les critiques adressées aux classements ?

Notre problématique est la suivante : « Dans quelle mesure les classements des universités permettent-ils d'évaluer la qualité dans l'enseignement supérieur ? »

1) Les différents types de classements des universités :

La plupart des classements des universités sont construits et publiés par les médias (par exemple : US News aux Etats-unis, Times en Grande Bretagne, Maclean's au Canada...etc.). Mais on trouve de plus en plus des classements construits et publiés par les universités (par exemple classement mondial de Shanghai, classement de Melbourne en Australie) ou par des agences gouvernementales. On trouve aussi des collaborations entre les médias qui publient les classements et les agences gouvernementales ou universités qui construisent les classements (par exemple le classement allemand CHE est une collaboration entre CHE qui construit le classement et le magazine Stern qui publie le classement)

La plupart des classements des universités sont produits par des entreprises commerciales de publication, qui ont pour objectif d'informer leurs lecteurs. Usher et Savino (2007, p. 6, p. 15) expliquent que les classements sont très coûteux à produire, ils demandent beaucoup de temps et d'efforts de collection et d'analyse de données, ce qui rend leur production non commerciale peu probable. Ils ajoutent que c'est le manque de transparence et d'informations comparatives sur la qualité des universités de la part des gouvernements a conduit les média de proposer les classements des universités aux consommateurs pour satisfaire leur demande. Surtout dans un contexte d'augmentation continue des droits d'inscriptions.

Il existe différents types de classements des universités¹ (voir **le tableau 1**). Les classements peuvent être nationaux, régionaux ou mondiaux. Les classements peuvent évaluer les universités, les départements, les disciplines, les programmes ou des activités spécifiques des universités. Bellon (2007, p. 30) estime qu'il existe actuellement environ 73 classements des universités (dont 15 mondiaux et 53 nationaux ou régionaux)

L'objectif principal des classements est d'aider les étudiants à faire des choix informés (Eccles, 2002). En effet les classements reflètent la demande des consommateurs de plus d'informations et de transparence. Puisque le choix d'une université correspond à un achat rare d'un service (enseignement supérieur) de plus en plus important dans la vie des consommateurs (étudiants). Cet achat est aussi de plus en plus coûteux. (Dill et Soo, 2004).

¹ Dans les pays anglophones les classements des universités (Rankings) sont aussi appelés des League Tables (une traduction littéraire serait les tableaux ou classements du championnat), se référant au système de classement des équipes de football dans les championnats.

Les classements des universités peuvent aussi servir comme un instrument de responsabilité au gouvernement, puisque les informations contenues dans les classements aident les étudiants à faire des choix informés, cela augmente la concurrence entre les universités pour attirer les meilleurs étudiants ce qui mène à une amélioration de la qualité des universités (Stella et Woodhouse, 2006, p.2). Van Dyke (2005, p. 104) note que même si les classements des universités à leurs débuts avaient comme objectif principal d'informer les étudiants. Aujourd'hui les objectifs des classements sont beaucoup plus larges (informations à toutes les parties prenantes, financement, transparence, responsabilité au gouvernement...etc.).

Tableau 1 : Les différents types de classements des universités avec un échantillon d'exemples :

	Classements nationaux	Classements régionaux	Classements mondiaux
Classements des universités (dans leur globalité)	<p><u>Classement des universités américaines :</u> -US News and World Report (America's Best Colleges) (première apparition 1983, annuel depuis 1987)</p> <p><u>Classement des universités britanniques :</u> -The Times Good University Guide (depuis 1992/annuel) -The Guardian University Guide (depuis 1999/annuel)</p> <p><u>Classement des universités canadiennes :</u> -Maclean's University Rankings (depuis 1991/annuel)</p> <p><u>Classement des universités australiennes :</u> -Melbourne institute (International Standing of Australian Universities) -The Good Universities guide (depuis 1992/annuel)</p> <p><u>Classement des universités allemandes :</u> -CHE ranking</p> <p><u>Classement des universités chinoises :</u> -Education 18 (Hong kong) -Guangdong Institute -Netbig -Wuhan</p> <p><u>Classement des universités espagnoles :</u> -Excelesia</p> <p><u>Classement des universités italiennes :</u> -Republica</p> <p><u>Classement des universités polonaises :</u> Perspectwy/Rzeczpospolita Uniwersytet</p>	<p><u>Classement des meilleures universités en Asie :</u> -Asia Week (arrêté en 2001, archives disponibles sur le site internet...)</p>	<p>-Classement de Shanghai (depuis 2003/annuel)</p> <p>-Classement du Times Higher Education Supplement (THES) (depuis 2004/annuel)</p>
Classements des départements/disciplines/ Programmes universitaires	<p><u>-Classement des écoles de commerce :</u> (en France, Etats-Unis, Grande Bretagne...etc.)</p> <p><u>-Classement des départements</u> de médecine, de droit...etc. dans différents pays du monde</p> <p>-CHE ranking (Allemagne) -The Times Good University Guide -The Guardian University Guide -Republica (Italie)</p>		<p>-Classement de Shanghai (par discipline)</p> <p>-Classement du Times (par discipline)</p> <p><u>Classement des écoles de commerce :</u></p> <ul style="list-style-type: none"> - Classement du Business Week - Classement de The Economist - Classement de The Financial Times - Classement de The Wall Street Journal
Classement d'activités spécifiques des universités	<p><u>L'activité de la recherche :</u> -Best American Research Universities -Center for Science and Technology Studies (en Bern, Suisse)</p> <p><u>Les universités américaines les mieux connectées à Internet et à leurs réseaux informatiques locaux :</u> -Yahoo Internet Life Magazine (Best Wired Colleges)</p> <p><u>L'intégration des étudiants issus de milieux sociaux différents :</u> -Ethnic Diversity Rankings (Journal of Black Higher Education)</p>		<p>-Le classement Webometrics (recherché et visibilité de l'université sur internet)</p> <p>-Professional Ranking of World Universities (classement français de l'école des mines de Paris, le critère c'est le nombre d'anciens étudiants dans un poste de CEO dans le top 500 entreprises mondiales)</p>

2) La construction des classements des universités :

On distingue 4 étapes de construction des classements des universités. La première étape consiste à choisir les indicateurs à utiliser dans le classement. Dans la deuxième étape les données sont collectées. Dans la troisième étape les indicateurs sont calculés et dans la quatrième étape une méthode de présentation du classement est choisie.

Les sources de données :

Les classements des universités utilisent 3 sources de données : les universités (qui fournissent leurs propres données), les enquêtes (des parties prenantes sur la qualité de l'université) et des parties tiers (gouvernement, agences de financement). (Voir le **tableau 2** pour les résultats d'un survey sur les sources de données utilisées dans un échantillon de 18 classements)

Chaque source de données a des avantages et des inconvénients. **Les données des parties tiers** (gouvernement, agences de financement) sont considérées le standard pour l'évaluation, elles sont précises, impartiales et considérées comme objectives. Mais elles ne sont pas créées dans l'objectif de servir les classements des universités, donc généralement ces données sont peu adaptées à ce genre d'exercice (comparaison des universités) de plus le choix des indicateurs qui utilisent ces données sera plus basé sur la disponibilité de données qu'a leur pertinence pour les classements. **Les données des universités** (obtenues en adressant des questionnaires aux universités) peuvent fournir des informations très précises (et individuelles) impossible d'obtenir par les enquêtes ou par les données du gouvernement et d'agences de financement. Mais le problème c'est que certaines universités peuvent refuser de répondre aux questionnaires ce qui remet en cause l'utilité de la comparaison un autre problème c'est que les données peuvent ne pas être consistants dans le temps. Ainsi que leur objectivité, puisque les universités ont intérêt de ne fournir que les données qui leur sont bénéfiques. **Les données des enquêtes** (des académiques, des étudiants et des employeurs), leur valeur est discutable, puisque les enquêtés même si on suppose qu'ils possèdent des informations sur leurs propres universités, ils n'ont pas assez d'informations sur les autres institutions dont on leur demande d'évaluer et ne peuvent donc donner des opinions objectives. (Usher et Savino, 2007, p. 8) (Je note que les données des enquêtes des étudiants ont l'avantage de permettre une évaluation qualitative de la qualité d'enseignement dans les

Tableau 3 : Le nombre total d'indicateurs et les sources de données pour un échantillon de 17 classements des universités :

Les classements des universités	Nombre d'indicateurs	Total	Sources de données		
			Enquêtes effectuées par les auteurs des classements	Universités	Des tiers (autorités publiques, base de donnée ISI et autres fournisseurs de données sur l'enseignement supérieur)
Asiaweek (Inde/Asia)	18		-	18	-
Daily Telegraph (GB)	1		-	-	1
Education18.com (Hong Kong)	9		3	2	4
Excelencia (Espagne)	71		-	-	71
Financial Times (GB)	17		-	-	17
Guangdong Institute (Chine)	17		-	3	14
Guardian University Guide (GB)	7		-	5	2
La Repubblica (Italie)	23		2	-	21
Maclean's (Canada)	24		1	18	5
Melbourne Institute (Australie)	26		3	-	23
Netbig (Chine)	18		1	7	10
Perspektywy/Rzeczpospolita (Pologne)	18		1	15	2
Shanghai Jiao Tong University (Chine)	6		-	-	6
The Times Good University Guide (GB)	9		-	-	9
Times World University Rankings (GB)	5		1	3	1
US News and World Report (US)	15		1	11	3
Wuhan (Chine)	45		2	21	22

Source : (Usher et Savino, 2007, p. 11, tableau1)

universités, mais ce type de données est surtout utilisé par les agences d'assurance qualité et très peu dans le cas des classements des universités² qui réalisent généralement des enquêtes des académiques seulement pour évaluer la réputation des universités, dans ce contexte la critique de Usher et Savino est pertinente).

Les différentes formes de présentation des classements :

On distingue deux méthodes principales de présentation des classements. La première méthode consiste à choisir des pondérations et agréger les indicateurs pour créer un score global pour chaque université et présenter un classement par ordre décroissant. Dans la deuxième méthode les indicateurs ne sont pas agrégés ni pondérés, le classement des universités se fait pour chaque indicateur individuellement. La plupart des classements des universités utilisent la première méthode, qui est d'ailleurs très critiquée à cause de choix de pondération arbitraire et le principe même que la qualité d'une université peut se résumer en un seul chiffre est remis en cause. (Voir le tableau 3, pour un échantillon des classements et la méthode de présentation utilisée)

Tableau 3 : la méthode de présentation des classements des universités :

Classement	Pondération, agrégation, score global et classement global	Sans pondération, classement pour chaque indicateur
Asia Week	X	
The center	X	
CHE/stern		X
Good Guides		X
The Guardian	X	
Maclean's	X	
Melbourne	X	
Perspektwy	X	
The Times	X	
US.News	X	

Source : (Nina Van Dyke, 2005, tableau5, p.113)

² En effet certains classements (comme celui du Times en GB) utilisent des données d'enquêtes des étudiants sur la qualité d'enseignement réalisées par des agences d'assurance qualité.

3) Une analyse des indicateurs les plus utilisés dans les classements des universités :

Le nombre d'indicateurs utilisés dans les classements varie fortement d'un classement à l'autre, mais certains types d'indicateurs sont plus utilisés que d'autres, on analysera dans cette partie les indicateurs les plus utilisés dans les classements des universités.

Les différents classements utilisent différents indicateurs, pondération des indicateurs et méthodologies et par conséquent ils produisent différents résultats. Voir le **tableau 4** pour une vue d'ensemble des indicateurs les plus utilisés dans les classements des universités.

(Van Dyke, 2005) a réalisé un survey de 10 classements des universités qui utilisent un total de 72 indicateurs répartis dans 7 larges catégories (voir le **tableau 5** pour les catégories d'indicateurs, le nombre d'indicateurs et le poids des indicateurs des différents classements surveyés). Aucun indicateur n'est utilisé par tous les classements (ce qui montre qu'il n'y a pas un consensus sur une définition globale et unique de la qualité des universités) l'indicateur le plus utilisé et le plus proche du consensus est celui de la qualité des étudiants à l'entrée à l'université (8 sur 10 classements surveyés l'utilisent)

Tous les 10 classements incluent au moins 1 seul indicateur dans la catégorie de la qualité du personnel académique.

Peu d'indicateurs de qualité d'enseignement, une explication possible c'est que évaluer l'enseignement est coûteux et consomme beaucoup de temps. Par exemple la Grande Bretagne dépense un budget considérable dans l'évaluation de la qualité d'enseignement dans les universités.

Il n'y a pas un consensus sur quels indicateurs de ressources à utiliser, certains classements incluent les bibliothèques, d'autres non. Pourquoi ces différences dans le choix des indicateurs des ressources entre les différents classements ? Ca peut refléter des différences

Tableau 4 : les indicateurs les plus utilisés dans les classements des universités répartis en input/output :

Les indicateurs input			Les indicateurs outputs			
Qualité des étudiants à leur entrée à l'université	Personnel académique	Ressources matérielles/financières	Qualité de l'enseignement	Qualité de La recherche	Les effets de l'enseignement/recherche sur l'individu/société	Réputation de l'université (Proxy de la qualité des outputs)
-les notes des étudiants à la fin de leurs études secondaires -les notes des étudiants dans les tests d'entrée à l'université -Taux de sélectivité des étudiants	-Ratio étudiants/enseignants -Taille des classes -La qualification des enseignants (leurs diplômes)	-Dépenses annuelles/étudiant -Fonds de dotation (endowments) -Dépenses sur la recherche -Equipement informatique des universités (nombre d'ordinateurs, connexions internet et dépenses) -Taille des bibliothèques (nombre de livres/revues) et dépenses	-Taux de réussite, d'échec ou d'abandon des études -Nombre de diplômes de master/doctorat attribués annuellement -Enquêtes sur les opinions des étudiants/employeurs	- <u>Les indicateurs bibliométriques :</u> -Nombre de citations -Nombre de publication - facteur d'impact des journaux	-les salaires de départ des nouveaux diplômés -le taux d'emploi/chômage des nouveaux diplômés	-Enquêtes sur les opinions des académiques (doyens, responsables de départements, présidents des universités)

Source : L'auteur

Tableau 5 : Nombre et poids des indicateurs répartis dans 8 larges catégories dans un échantillon de 10 classements nationaux des universités :

Les classements	Qualité du personnel académique*		Qualité des étudiants entrants		Qualité des programmes undergraduate		Qualité des programmes graduate		Ressources		Opinion des parties prenantes		Autres		Nombre total des indicateurs
	Nombre	Poids	Nombre	Poids	Nombre	Poids	Nombre	Poids	Nombre	Poids	Nombre	Poids	Nombre	Poids	
Asia Week	6	23%	2	14%	0	0%	0	0%	8	32%	2	20%	2	10%	20
The center	4	33%	1	11%	0	11%	1	11%	3	33%	0	0%	0	0%	9
CHE/stern	1	n/a	0	n/a	2	n/a	5	n/a	20	n/a	2	n/a	3	n/a	33
Good Guides	5	n/a	3	n/a	3	n/a	0	n/a	7	n/a	1	n/a	7	n/a	26
The Guardian	1	40%	1	10%	2	15%	0	0%	2	20%	0	0%	0	15%	6
Maclean's	4	17%	4	15%	3	10%	1	1%	7	27%	1	16%	3	15%	23
Melbourne	4	40%	1	11%	2	14%	1	16%	2	11%	2	8%	0	0%	12
Perspektwy	2	23%	0,5	2,5%	0	0%	0,5	0%	10	20%	2	50%	2	5%	17
The Times	2	37%	1	9%	3	27%	0	0%	3	27%	0	0%	0	0%	9
US.News	1	3%	3	16%	3	23%	0	0%	4	23%	1	25%	3	11%	15
Total	30		16,5		18		8,5		66		11		20		170
Moyenne	3	27%	2	11%	2	14%	1	4%	5**	24%	1	15%	2	7%	

* Pour la qualité du personnel académique, Van Dyke inclut aussi les indicateurs de la qualité de recherche.

**la moyenne du nombre de ressources a été faite sans les indicateurs du classement CHE (20 indicateurs de ressources, un nombre considéré comme outlier par Nina Van Dyke)

Source : (Nina Van Dyke, 2005, tableau 9 p.113, tableau 10, p. 114, tableau 12, p. 116)

dans les systèmes d'enseignements supérieurs nationaux, des variabilités des ressources entre les universités, des différences culturelles sur la question de ce qui affecte le plus la qualité académique ou la disponibilité ou fiabilité de données. (Van dyke, 2006, P.111)

Les indicateurs de la qualité du staff académique dans l'université :

Les indicateurs dont les classement assignent le plus de poids sont les indicateurs des ressources et de la qualité du personnel académique (qui inclut aussi la qualité de recherche) (Van Dyke, 2005, p. 114) à l'exception du classement perspektwy et du classement US news qui donnent le plus de poids aux opinions des parties prenantes.

L'indicateur qualité des étudiants à leur entrée à l'université :

L'indicateur qualité des étudiants à leur entrée à l'université est un des indicateurs les plus utilisés dans les classements des universités. Dill et Soo (2004, p.4) pensent que trois arguments sont à l'origine de l'utilisation des classements de cet indicateur input comme mesure de la qualité des universités : Le premier argument c'est que les meilleurs étudiants réalisent les meilleurs résultats à la fin de leurs études universitaires (le problème avec cet argument c'est que les résultats des étudiants à la fin de leurs études universitaires ont deux sources : la qualité de l'enseignement de l'université (la valeur ajoutée) et les qualités personnelles des étudiants indépendantes de la qualité d'enseignement) le deuxième argument ce sont les effets de pairs (les meilleurs étudiants enseignent d'autres étudiants, donc avoir les meilleurs étudiants améliore la qualité de l'enseignement produite par l'université). Le problème avec cet argument c'est l'importance des effets de pairs s'ils sont significants ou non. Le troisième argument c'est un argument considéré par Dill et Soo comme tautologique c'est que l'université qui réussit d'attirer les meilleurs étudiants devrait forcément être bonne. Dans cet argument la qualité de étudiants est un proxy de la qualité d'enseignement de l'université. Le problème majeur avec cet indicateur c'est qu'il ne représente pas la qualité d'enseignement offerte par l'université, mais seulement une petite partie (effets de pairs).

Les indicateurs de réputation des universités :

Cet indicateur est très utilisé dans les classements des universités (le classement mondial du Times par exemple)

Il prend généralement la forme de résultats d'enquêtes réalisés sur les académiques. En demandant leur opinion sur la question : Quelles sont selon vous les meilleures universités ? Ou Quelles sont les universités les plus renommées ?

Plusieurs classements, utilisent un indicateur qui mesure la réputation de l'université, généralement cet indicateur est obtenu en réalisant une enquête des académiques (doyens et présidents des universités) en demandant leur opinion sur la qualité ou la réputation des différentes universités.

Selon Williams et Van Dyke (2007, p. 822-823), la réputation d'une université dépend en même temps sur ses performances actuelles ainsi que sur celles du passé (parce que ça prend du temps pour que les performances deviennent bien reconnues). Ils pensent que grâce à Internet cette intervalle de temps entre les performances et la réputation est en train de diminuer. Et donc la réputation représente de plus en plus les performances actuelles des universités. En plus, plus les répondants à l'enquête sont mieux informés, plus la réputation est corrélée avec les performances actuelles des universités qu'aux performances du passé. Williams et Van Dyke ajoutent que ce qui détermine le plus la réputation d'une université c'est la contribution de ses chercheurs dans la recherche internationale, Même si d'autres facteurs jouent aussi un rôle : La qualité d'enseignement et la formation des chercheurs. Cependant, ils affirment que c'est la production totale de la recherche qui compte le plus pour la réputation internationale et non la production par chercheur³.

Pour toutes ces raisons, Williams et Van Dyke défendent la grande utilisation des indicateurs de recherche dans les classements mondiaux des universités (notamment celui de Shanghai) vu qu'au niveau internationale c'est la recherche qui procure le plus de réputation de visibilité et d'image aux universités sur la scène internationale.

Selon Williams et van Dyke (2008, p. 2-3) on peut évaluer les universités soit par des indicateurs de performances ou par des indicateurs de réputation. Selon eux, l'avantage des indicateurs de performances c'est qu'ils sont objectifs, leur désavantage c'est que le choix des indicateurs et leur poids est subjectif. L'avantage de réputation c'est que le choix des indicateurs et leur poids est implicitement laissés a un grand nombre de personnes interrogées. Donc il n'est pas arbitraire ou personnelle. Par contre le désavantage c'est la que les personnes interrogées possèdent des informations limités sur les universités qu'on leur demande à évaluer. Dans quelle mesure les résultats d'évaluation par réputation sont différents des résultats d'évaluation par performances, dans leur étude ils montrent que les résultats des deux méthodes sont assez proche (en comparant les résultats des universités

³ Ce qui peut être expliqué pourquoi les grandes universités sont généralement plus réputées par rapport aux petites universités.

australiennes dans le classement Melbourne australien qui utilise des indicateurs de réputation et dans le classement de Shanghai qui n'utilise pas des indicateurs de réputation).

Mais Selon Van Dyke (2005), si on veut mesurer les performances actuelles des universités on a pas besoin de réaliser des enquêtes d'opinions sur la réputation des universités.

4) Est-ce qu'il existe une définition universelle de la qualité des universités ?

Dans leur survey de 18 classements des universités, Usher et Savino (2007, p. 12) concluent qu'il y'a très peu de consensus sur la définition de la qualité des universités, certains classements utilisent plus d'indicateurs inputs que d'outputs et d'autres le contraire. Avec des différences dans les choix des inputs et outputs et leur pondération. Ce résultat contredit le résultat de (Dill et Soo, 2004) qui ont conclu qu'il existe une convergence de la définition de la qualité des universités mais leur échantillon était relativement petit (un survey de 5 classements seulement comparé au survey de 18 classements de Usher et Savino ce qui est beaucoup plus représentatif). En effet, selon (Usher et Savino, 2007, p. 5) chaque classement impose sa propre définition de la qualité dans l'enseignement supérieur sans qu'il mentionne les autres alternatives possibles. (Différentes pondérations et donc d'importance donnée à chaque indicateur et donc aussi des visions différentes de la qualité. Ainsi que différents types d'indicateurs et différents nombres d'indicateurs, voir **le tableau 6** pour une vue d'ensemble des différentes pratiques de pondérations utilisées dans les classements ainsi que le nombre des indicateurs utilisés)

Dans son survey de 10 classements, Van Dyke (2005, p. 115) conclut que les classements partagent quelques larges principes et approches : L'importance des programmes undergraduate (premiers cycles universitaire), la pondération des indicateurs pour créer un chiffre unique pour chaque université. L'importance d'un personnel académique de haute qualité. L'importance des ressources. Et l'importance de la qualité des étudiants. Mais les classements diffèrent dans les détails : Différents poids sont donnés à la recherche/enseignement et autres indicateurs, différents choix des ressources à mesurer et l'importance des programmes graduate (de deuxième et troisième cycles) certains classements considèrent ces programmes comme importants pour la qualité de l'université contrairement à d'autres classements.

Tableau 6 : Le poids des indicateurs Inputs/outputs dans un échantillon de classements des universités :

	Le poids des indicateurs Inputs (ressources humaines/ressources matérielles)				Le poids des indicateurs Outputs					Total (en %)
	Qualité des étudiants à leur entrée à l'université	Personnel académique	Ressources matérielles/ financières	Le poids total des inputs (en%)	Qualité de l'enseigne- ment	Qualité de recherche	Les effets de l'enseignement/ recherche sur l'individu/société	Réputation de l'université (Proxy de la qualité des outputs)	Le poids total des outputs (en %)	
Asiaweek (Inde/Asia)	25	28,3	10	63,3	0	0	16,7	20	36,7	100
Daily Telegraph (GB)	0	100	0	100	0	0	0	0	0	100
Education18.com (Hong Kong)	20	15	5	40	0	0	20	40	60	100
Excelencia (Espagne)	0	25	25	50	25	0	25	0	50	100
Financial Times (GB)	9	19	15	43	10	27	20	0	57	100
Guangdong Institute (Chine)	0	0	0	0	57,1	0	42,1	0	99,2	99,2
Guardian University Guide (GB)	28	35	10	73	10	17	0	0	27	100
La Repubblica (Italie)	10	44,4	15,6	70	10	0	20	0	30	100
Maclean's (Canada)	15	20	44	79	5	0	0	16	21	100
Melbourne Institute (Australie)	11	3,5	11	25,5	12,6	4,8	40	17,1	74,5	100
Netbig (Chine)	12	21,8	6	39,8	0	0	45,2	15	60,2	100
Newsweek (US)	10	20	10	40	0	0	60	0	60	100
Perspektywy/Rzeczpospolita (Pologne)	8	20,5	11,5	40	0	0	0	50	50	90
Shanghai Jiao Tong University (Chine)	0	0	0	0	10	0	90	0	100	100
The Times Good University Guide (GB)	3,3	53,3	6,7	63,3	3,3	3,3	30	0	36,6	99,9
Times World University Rankings (GB)	5	25	0	30	0	0	20	50	70	100
US News and World Report (US)	15	20	15	50	25	0	0	25	50	100
Wuhan (Chine)	10,6	8,5	16,6	35,7	3,4	0,6	48,6	11,7	64,3	100

Source : (Usher et Savino, 2007, p. 11, tableau2) j'ai juste réorganisé les données en inputs/outputs et calculé le poids total des inputs et le poids total des output

Une autre question intéressante à considérer serait de savoir est ce que ces différentes définitions de la qualité impliquent ou non des résultats de classements différents. Usher et Savino (2007, p.12-13) affirment que le rang d'une université dépend largement de ce que les évaluateurs choisissent d'évaluer dans leur classement. Un exemple : Queen's University (au Kingston au Canada) dans le classement canadien Maclean's est bien classée par rapport aux autres universités canadiennes, par contre dans les classements mondiaux, elle est moins bien classée, sûrement due à sa petite taille qui la place en désavantage en terme d'output total de la recherche. Usher et Savino font une remarque intéressante qu'est l'existence d'une convergence des résultats des meilleures universités (top 10) entre les classements nationaux et mondiaux, mais dès qu'on dépasse la dixième position (ou se situe au plus bas de l'échelle) on commence à apercevoir des divergences et d'inconsistance des résultats. Pourquoi ?

Selon eux il y'a déjà deux scénarios exclus : Si les résultats des classements sont inconsistants entre tous les classements, en l'occurrence, on pourrait rejeter l'idée des classements et dire qu'ils n'ont aucun sens et ne sont fait que pour des bénéfices commerciaux. Le deuxième scénario c'est de trouver que les résultats des classements sont consistants, en l'occurrence, on peut en déduire qu'il existe quelques indicateurs de qualité partagés par tous les classements et qui conduisent à des résultats similaires. Nous pourrions dans ce cas en déduire une définition universelle de la qualité des universités.

Le problème c'est qu'aucun des deux scénarios n'est plausible, on remarque dans le top des classements une convergence des résultats, mais en bas de l'échelle une divergence des résultats entre les différents classements. Une explication possible de ce phénomène pourrait être le fait que les classements ne mesurent pas ce que leurs auteurs pensent qu'ils mesurent c'est-à-dire des indicateurs proxy de la qualité. Mais mesurent seulement les conséquences de certaines caractéristiques des universités. Ces caractéristiques conduisent au Top certaines universités dans tous les classements quelques soit les indicateurs choisis leur nombre ou leur poids, une sorte de matière noire exerçant une force de gravité sur tous les classements. Des candidatures plausibles pour ce rôle pourraient être : l'âge des universités (les anciennes universités sont généralement au top), la taille des universités (les grandes universités sont généralement au top) ou les dépenses par étudiant (les universités qui dépensent le plus sont généralement au top). (Usher et Savino, 2007, p. 13)

Le choix des indicateurs et poids influence fortement les résultats des classements, ce qui implique que les classements imposent leur définition de la qualité sur toutes les universités

classées. (One size fits all definition of quality). (Usher et Savino, 2007, p. 7)

Un exemple qui montre bien les différences des résultats entre les différents classements est de comparer les classements des universités françaises et américaines dans les deux classements mondiaux des universités, celui de Shanghai et celui du Times (**voir le tableau 7**)

Tableau 7 : comparaisons des résultats des classements mondiaux du Times et de Shanghai pour les universités françaises et américaines :

Universités françaises	Classement mondial de Shanghai (top 500)	Classement mondial du Times (top 200)	Universités américaines	Classement mondial de Shanghai (top 200)	Classement mondial du Times (top 500)
Université Paris VI	45e	93e	Harvard University	1e	1e
Université Paris XI	64e		Stanford University	3e	6e
Université Strasbourg I	96e		California Univ - Berkeley	4e	8e
Ecole Normale Supérieure de Paris	99e	18e	Massachusetts Institute of Technology (MIT)	5e	4e
Université Paris VII	102e-150e		California Inst Tech	6e	7e
Université Grenoble I	151e-200e		Columbia University	7e	12e
Ecole Polytechnique	201e-300e	37e	Princeton University	8e	10e
Université Lyon I	201e-300e		Univ Chicago	9e	11e
Université Méditerranée	201e-300e		Yale University	11e	4e
Université Montpellier II	201e-300e		Cornell University	12e	15e
Université Paris V	201e-300e		Univ California - San Diego	13e	44e
Université Toulouse III	201e-300e		Univ California - Los Angeles	14e	31e
Ecole des Mines Paris	301e-400e		Univ Pennsylvania	15e	26e
ESPCI Paris	301e-400e		Univ Wisconsin - Madison	16e	79e
Université Bordeaux I	301e-400e		University of Washington	17e	84e
Université Nancy I	301e-400e		Univ California - San Francisco	18e	na
Université Paris IX	301e-400e		Johns Hopkins Univ	20e	23e
Ecole Normale Supérieure de Lyon	401e-500e	72e	Michigan Ann arbor	21e	29e
Université Aix Marseille I	401e-500e		University of illinois	25e	77e
Université Bordeaux II	401e-500e		Washington university in saint louis	28e	48e

Sources : classement mondial de shanghai 2006 <http://ed.sjtu.edu.cn/ranking.htm>
classement mondial du Times 2006 <http://www.timeshighereducation.co.uk/hybrid.asp?typeCode=161>

Pour les universités françaises, le classement mondial du Times (édition 2006) place 5 universités françaises dans le top 100 (l'école normale supérieure de paris 18^{ème}, école polytechnique 37^{ème}, sciences-po 52^{ème}, école normale supérieure d Lyon 72^{ème} et pierre et marie curry 93^{ème}) et 7 universités françaises dans le top 200 (les universités du top 100 plus l'université louis pasteur de strasbourg classée 161^{ème} et l'université paris-sorbonne (paris4) classée 200^{ème}). Tandis que le classement mondial de Shanghai (édition 2006) place 4 universités françaises dans le top 100 (paris 6, 45^{ème} paris 11 64^{ème}, université Strasbourg 1 96^{ème}, école normale supérieure de paris 99^{ème}) et 6 dans le top 200 (les précédents plus l'université paris 7 classée 102^{ème} et l'université Grenoble 1 classée 151^{ème}). Et dans son total top 500 on trouve 21 universités françaises.

Dans le classement de Shanghai 51 universités américaines sont classées dans le top 100 (51% des universités du top 100 de Shanghai sont américaines) et 90 universités américaines sont classées dans le top 200 (45% des universités du top 200 de Shanghai sont américaines). Cependant, les universités américaines sont largement moins bien classées dans le classements du Times, qui place 31 universités américaines dans le top 100 (31% des universités du top 100 du Times sont américaines) et un total de 53 universités dans le top 200 (26% des universités du top 200 du Times sont américaines).

En effet comme dans la remarque de Usher et Savino (2007), les universités du top sont toujours au top dans tous les classements, tandis que pour les universités loin du top, les résultats sont très différents.

5) Les indicateurs bibliométriques et évaluation de la qualité de recherche :

Contrairement à l'évaluation de la qualité d'enseignement, l'évaluation de la qualité de recherche est loin d'être un nouveau phénomène dans le monde académique. En effet on peut même dire que dans toute l'histoire de la recherche scientifique les travaux de recherche ont toujours été évalués sous une forme ou une autre. Cependant, l'évaluation de la recherche au sens moderne du terme a commencé grâce à l'invention du processus peer-review (évaluation par les pairs).

Aujourd'hui on distingue deux méthodes d'évaluation de la qualité de recherche : une méthode qualitative classique qui est l'évaluation par les pairs (peer-review) et une méthode quantitative relativement récente et de plus en plus utilisée qui est l'évaluation par les

indicateurs bibliométriques. Quels sont les avantages et inconvénients de chacune des deux méthodes ? Est ce que ces deux méthodes sont indépendantes l'une de l'autre ? Est ce qu'elles sont substituables ou complémentaires ? C'est ce qu'on essayera de savoir dans les prochains paragraphes.

L'évaluation qualitative de la qualité de recherche ou l'évaluation par les pairs (peer-review) :

L'évaluation par les pairs (peer-review) est une évaluation qualitative par des spécialistes du champs de recherche concerné. Le peer-review peut être utilisé dans l'évaluation des chercheurs, des travaux de recherche réalisées ou des projets de recherche. Cette évaluation à 2 utilisations essentielles soit pour des décisions budgétaires ou pour des décisions de publications. Les décisions budgétaires peuvent être l'allocation des fonds aux meilleurs chercheurs et aux meilleurs projets de recherche, ou la récompense des meilleurs chercheurs et meilleurs travaux de recherche réalisés ou l'attribution des postes ou la décision des grilles de rémunération de chercheurs selon leurs performance de recherche. Pour les décisions de publications (notamment l publication dans des revues peer reviewed ne publient les articles qu'après articles, de livres)

Les critiques adressées à l'évaluation par les pairs :

On trouve dans la littérature plusieurs critiques du processus peer-review comme méthode d'évaluation de la qualité de recherche. On peut distinguer trois critiques essentielles du peer-review comme processus d'évaluation des chercheurs et des projets de recherche. *La première critique* concerne la divergence et parfois même la contradiction entre les opinions des évaluateurs. Pour l'évaluation des projets de recherche, généralement les évaluateurs sont plus d'accord sur les projets à refuser qu'aux projets à accepter. Tandis que pour les décisions de publication c'est le contraire, les évaluateurs sont plus d'accord sur l'acceptation des articles qu'a leurs rejets. Weingart répond à cette critique en disant que la divergence des opinions est nécessaire pour éviter la dominance d'une opinion sur une autre. *La deuxième critique* concerne l'inconsistance des opinions émises dans le temps, les opinions des évaluateurs sur la qualité d'un chercheur, d'un projet de recherche ou d'un travail de recherche changent au fil du temps ce qui remet en cause la crédibilité de l'évaluation initiale. *La troisième critique* porte sur le problème d'objectivité des opinions, qui peuvent être biaisées, puisque d'autres facteurs que la qualité de recherche évaluée influencent les décisions des évaluateurs, notamment le hasard et le chance, être intégré dans le bon réseau ou

appartenir à la bonne institution. (Weingart, 2005, pp. 121-124)

D'autres critiques : La subjectivité du processus, les résultats de l'évaluation dépendent sur les choix personnels d'un nombre limité d'individus. Ce qui peut résulter en conflits d'intérêts, une méconnaissance de la qualité, ou un biais contre les jeunes chercheurs peu connus les nouveaux entrants dans un champs de recherche quelconque. Il y'a un besoin d'utiliser en parallèle des indicateurs bibliométriques pour corriger les lacunes du processus peer-review. Le peer-review devrait rester l'outil principal d'évaluation de la qualité de recherche, les indicateurs bibliométriques devraient aider le processus peer-review et non le remplacer. (van Raan, 2005, pp. 135)

L'évaluation quantitative de la qualité de recherche ou l'évaluation par les indicateurs bibliométriques :

La première évaluation des institutions de recherche basée sur des indicateurs bibliométriques s'est réalisée par martin et Ivrine en 1983 en G.B. Les réactions étaient selon (weingart, 2005, pp. 117-118) prévisibles : des Critiques méthodologiques et techniques de l'évaluation mais aussi des menaces d'attaquer en justice les analystes qui ont réalisé l'évaluation ainsi qu'une crainte d'effets néfastes sur la science. En effet la communauté scientifique pensait que seulement les experts ont le droit d'évaluer les performances de la recherche en utilisant le processus classique de peer-review (l'évaluation des chercheurs, travu de recherche et projets de recherches par les pairs des chercheurs, c'est-à-dire par des spécialistes du champs de recherche concernée) selon la communauté scientifique ce système a bien fonctionné. Les décideurs public ont supporté ce rejet de la communauté scientifique de l'utililisation des indicateurs bibliométriques dans l'évaluation des chercheurs, à cause de leur désintérêt dans ce sujet à l'époque. Mais avec les problèmes financières, les autorités publiques se trouvent obligé de rationaliser les dépenses publiques (faire des choix, redistribuer efficacement les dépenses pas question de les augmenter, besoin d'indicateurs qui facilitent cette tâche)

Et aujourd'hui on a un changement (major shift) de la position de la communauté scientifique et leurs représentants (dirigeants des agences de financement de recherche, dirigeants des centres de recherche, présidents des universités...) vers une culture de responsabilité (peu ou pas de critique aujourd'hui des indicateurs bibliométriques ?)

La bibliométrie c'est la mesure de la littérature⁴ (Hood et Wilson, 2001). Elle mesure et analyse la production scientifique des chercheurs en se basant sur le nombre de publications et le nombre de citations. Les indicateurs bibliométriques sont utilisés comme une méthode d'évaluation de la qualité de recherche. On appelle des bibliométriciens, les praticiens de la bibliométrie (ceux qui collectent les données, qui construisent les indicateurs ou qui analysent les indicateurs bibliométriques). On présentera tout d'abord une introduction générale sur les indicateurs bibliométriques, pour ensuite aborder les critiques envers ces indicateurs et leur utilisation dans l'évaluation de la qualité de recherche.

Les indicateurs bibliométriques : présentation générale⁵

Les indicateurs bibliométriques reposent sur l'analyse des citations. Deux étapes pour la construction d'un indicateur bibliométrique : collection des données de citation (à partir des références bibliographiques à la fin de chaque article) et définition et calcul de la formule mathématique de l'indicateur.

Collecter et gérer des données de citations est un processus complexe très coûteux et nécessite beaucoup de temps et d'effort vu le nombre très élevé de citations. Ajouter à cela, tous les problèmes techniques et méthodologiques associés à ce processus, notamment le problème d'attribution de citations aux auteurs et aux universités. C'est pour que sont créées des bases de données fournissant ces citations et prêtes à être utilisées. (Kermarrec et al, 2007, p. 6)

On distingue des bases de données payantes et des bases de données gratuites. La base de donnée payante la plus utilisée c'est l'index **SCI (Science Citation Index)** sous sa forme papier ou **WoS : Web of Science** sous sa forme électronique sur internet. Cette base de données mondiale est construite par la société commerciale Thomson aux Etats-Unis. Dans le SCI sont indexés 8700 journaux (80% en sciences dures et 20% en sciences sociales) et 38 millions de citations mises à jour chaque semaine et qui remontent jusqu'à 1900. Mais le SCI ne couvre qu'une très faible partie de la littérature scientifique, puisqu'on estimait en 1999 qu'il existe 100 000 journaux dans le monde (dont 25 000 dans le domaine biomédical), c'est à dire que le SCI couvre moins de 10% des journaux dans le monde. Cette couverture est très inégalement distribuée entre les pays (un biais américain) et les langues (un biais d'anglais).

⁴ Voir (Hood et Wilson, 2001) pour une discussion sur les origines, définitions et différences entre la bibliométrie (mesure de la littérature), la scientométrie (mesure de la science), l'informétrie (mesure de l'information) la libramétrie (mesure des bibliothèques) et la netométrie (application des différents metrics sur internet).

⁵ Cette partie est entièrement basée sur un (document d'analyse de la commission d'évaluation de l'INRIA, 2007)

98,74% des articles indexés dans le SCI sont en anglais (seulement 0,234% en français). Le SCI indexe les lettres, les revues de livres et les introductions éditoriaux. Mais n'indexe pas les références bibliographiques aux livres ni aux conférences. Et les auto-citations (un auteur qui cite ses propres travaux) ne sont pas comptées.

Une deuxième base de donnée payante mais beaucoup moins utilisée que le SCI est celle appelée **Scopus**, lancée par Elsevier en 2004, elle contient 30 millions de citations provenant de 15 000 titres (dont 12 850 journaux, le reste sont des séries de livres, des sites internet et autres types de publications, contrairement au SCI qui n'indexe que des journaux ce qui est un avantage pour Scopus. 60% des titres sont non américains, avec plus d'articles non anglais que dans la base SCI, mais la critique essentielle pour cette base de donnée c'est qu'il manque la possibilité de faire des analyses par institution⁶.

Pour les bases de données gratuites on a **Citeseer** (une base de données spécialisée en informatique qui indexe des conférences, des journaux et des auteurs en informatique). **Google Scholar** (son but est de référencer les thèses, les livres, les articles, les rapports, les conférences...qui sont accessibles sur internet.cette base de donnée ala réputation d'une mauvaise gestion des homonymes, le même article ou auteurs d'articles peuvent être référenciés plusieurs fois sous différents noms). **Citebase** (inclut des informations sur le nombre de téléchargements des articles ainsi que le temps entre téléchargement et citation de l'article). Les problèmes communs à toutes ces bases de données gratuites c'est qu'ils sont très peu sophistiquées, les données ne sont pas bien organisée avec des problèmes d'identification des personnes, des institutions, des articles et des journaux. Ce qui rend ces bases de données très peu fiable pour une utilisation d'évaluation comparative de la qualité de recherche et des chercheurs. (Kermarrec et al, 2007, p. 6-13)

Dans toutes ces bases de données, il ne faut pas oublier que l'analyse des citations ne donne qu'une vue partielle de la production scientifique, puisque 90% des articles publiés ne sont jamais cités. Et que les citations ne prennent pas en considération d'autres aspects de la qualité de production scientifique comme les brevets par exemple, ou la qualité de transfert de connaissances entre les chercheurs et le secteur privé (Kermarrec et al, 2007, p. 6) ou entre les chercheurs et la société.

⁶ Je pense que c'est une des raisons principales de l'utilisation presque exclusive de la base de donnée SCI pour construire les indicateurs bibliométriques dans les classements des universités et autres types d'évaluation

Il existe différents indicateurs bibliométriques et chaque indicateur peut avoir plusieurs variantes, de nouveaux indicateurs continuent d'être créés par les bibliométriciens. Cependant, il y'a quelques indicateurs bibliométriques qui sont beaucoup plus utilisés que d'autres.

On distingue deux types principaux d'indicateurs bibliométriques. *Le premier type* d'indicateurs bibliométriques se sont les indicateurs quantitatifs facilement et directement obtenues à partir des bases de données de citations. Notamment, le nombre de publications et de citations par chercheur, par groupe de chercheurs ou par institution (laboratoire, département, universités) ou le nombre de publications dans les journaux à fort impact, ou indexés dans la base de donnée SCI. *Le deuxième type* des indicateurs bibliométriques se sont les indicateurs nécessitants des formules mathématiques spécifiques. Notamment, l'indicateur facteur d'impact d'un journal (JIF : Journal Impact Factor), l'indice d'immédiateté (immediacy index). L'indicateur demi vie des citations (cited half-life), le H-index...etc. (Kermarrec et al, 2007, p. 15-21)⁷

L'indicateur facteur d'impact des journaux⁸ est l'indicateur bibliométrique le plus populaire et le plus utilisé parmi les indicateurs bibliométriques qui nécessitent des formules mathématiques. Il a été introduit dans les années 60 par Eugen Garfield (fondateur de l'ISI). Cet indicateur mesure la fréquence de citation des articles pour les différents journaux (combien en moyenne les articles d'un journal sont cités). La période choisie pour calculer cette moyenne de citation est de 2 ans. La formule de calcul de cet indicateur est la suivante : (le nombre total de citations dans les deux dernières années / le nombre total d'articles publiés dans le journal dans les deux dernières année). (Kermarrec et al, 2007, p. 16)

Par exemple si on veut calculer le nombre le facteur d'impact d'un journal X en 2007, alors que ce journal a publié en 2006 300 articles qui ont été cités 6500 fois, et a publié en 2005 248 articles qui ont été cités 5300 fois. Alors le facteur d'impact de ce journal x en 2007 = n.b de citations/n.b d'articles = $(6500+5300) / (300+248) = 11800 / 548 = \mathbf{21,53}$

Le facteur d'impact du journal x en 2007 est de 21,53, c'est-à-dire qu'en moyenne chaque article publié dans le journal x dans les deux dernières années a été cité 21,53 fois.

⁷ Dans le document (INRIA, 2007, p. 15-21) différents types d'indicateurs sont présentés, j'ai retenu dans ce texte seulement les deux types que je considère essentiels.

⁸ Pour des informations plus détaillées sur l'indicateur facteur d'impact des journaux, voir (Glanzel et Moed, 2002) qui présentent une analyse détaillée de cet indicateur ainsi qu'une présentation de ses différentes variantes et substituts.

Plus le facteur d'impact d'un journal est élevé, plus on considère que ce journal est important pour la communauté scientifique puisque ces articles sont plus cités. Les chercheurs qui publient dans des journaux à fort facteur d'impact sont considérés de meilleurs chercheurs que ceux qui publient dans des journaux à faible facteur d'impact. Un article publié dans un journal à fort facteur d'impact est aussi considéré comme plus important. Cette interprétation de l'indicateur est considérée comme dangereuse et erronée pour un nombre de chercheurs bibliométriciens, c'est ce qu'on verra dans la partie des critiques des indicateurs bibliométriques.

Les raisons d'une utilisation accrue des indicateurs bibliométriques dans l'évaluation de la qualité de recherche :

Pourquoi les indicateurs bibliométriques gagnent du terrain et sont de plus en plus utilisés dans l'évaluation de la qualité de recherche ? Cela est dû à deux raisons principales : **un scepticisme accru** envers le processus d'évaluation par les pairs (la transparence et fiabilité du processus sont remises en cause) ainsi **qu'un besoin des décideurs** d'indicateurs à grande échelle pour allouer efficacement les budgets destinés à financer la recherche. Certains pays utilisent ces indicateurs bibliométriques pour évaluer les chercheurs, d'autres pays ne l'ont pas encore fait (par exemple le gouvernement fédéral américain n'utilise pas les indicateurs bibliométriques mais ses propres indicateurs couplés avec l'évaluation par les pairs pour les décisions budgétaires).

Pour les pays qui utilisent les indicateurs bibliométriques dans l'évaluation de la qualité de recherche, on distingue deux types d'utilisation : Certains pays construisent eux-mêmes leurs propres institutions d'évaluation pour collecter et analyser les données sur les performances de la recherche (exemple en France l'observatoire des sciences et des techniques). D'autres pays utilisent des évaluateurs indépendants non gouvernementaux. (Weingart, 2005, p. 119)

Selon Weingart les 2 groupes dépendent d'un seul fournisseur de données bibliométriques internationales : le ISI (Institute for Scientific Information) qui est une entreprise commerciale privée, son nouveau propriétaire Thomson company of Toronto⁹, commercialise, fait de la

⁹ L'Institute for Scientific Information (ISI) a été fondé en 1960 par Eugene Garfield et a été acquis par Thomson Scientific and Health Care en 1992. L'institut offre une base de données bibliométrique de citations des articles publiés dans des revues. Ainsi que plusieurs index (Science Citation Index SCI, Social Sciences Citation

publicité et promeut fortement l'utilisation des indicateurs ISI pour tous ceux qui sont prêt à payer les prix (qui augmentent dans le temps à cause d'une demande croissante). (Weingart, 2005, pp.119-121) distingue *trois effets négatifs* de cette commercialisation agressive des données bibliométriques ISI. *Le premier effet* c'est que les chercheurs qui corrigent les données seront forcés de se retirer du marché, à cause de l'image donnée par ISI que ses données sont correctes et ne nécessitent aucun nettoyage supplémentaire coûteux par les utilisateurs. *Le deuxième effet* de cette commercialisation agressive des données ISI c'est le risque d'une utilisation irresponsable des données par des personnes non qualifiées ne connaissant pas les méthodologies utilisées pour l'obtention de ces indicateurs et qui décident exclusivement sur la base de ces indicateurs l'affectation du budget de recherche et la rémunération des différents chercheurs. *Le troisième effet* c'est que c'est très dangereux qu'une entreprise privée commerciale possède le monopole virtuel (actuellement pas de concurrents) de production de données influençant les décisions d'affectation budgétaire des universités et chercheurs dans le monde. C'est dangereux puisque quand la pression pour commercialiser des données prématurées augmente et la qualité des données diminue les décisions budgétaires basées sur ces indicateurs seront faussées.

L'utilisation accrue des indicateurs bibliométriques comme réponse à une méfiance et scepticisme envers l'évaluation par les pairs est souvent due à 2 fausses hypothèses. *La première hypothèse* c'est que les indicateurs bibliométriques étant quantitatives sont nécessairement plus objectifs que l'évaluation par les pairs qui est un processus d'évaluation qualitatif. La deuxième hypothèse c'est que les indicateurs bibliométriques sont indépendants du processus d'évaluation par les pairs. (Weingart, 2005, p. 122)

Les deux hypothèses sont erronées puisque les indicateurs bibliométriques ne sont qu'une simple représentation quantitative à grande échelle des résultats déjà établis des processus d'évaluation par les pairs et donc ils ne sont pas indépendants, de ce fait les indicateurs bibliométriques ne peuvent être plus objectif que l'évaluation par les pairs. Weingart distingue 2 avantages réels des indicateurs bibliométriques par rapport à l'évaluation par les pairs. *Le premier avantage* concerne la non réactivité des indicateurs bibliométriques dans le sens qu'étant basés sur de larges événements d'évaluation par les pairs, ces large événements

SSCI, Arts and Humanities Citation Index AHCI) et publie annuellement des (journal citation reports) qui contiennent l'indicateur facteur d'impact des revues. L'ISI publie aussi la liste des chercheurs les plus cités. 14,000 revues sont indexés dans la base de donnée ISI
(source : http://en.wikipedia.org/wiki/Institute_for_Scientific_Information)

n'étaient pas créés dans le but d'être comptés pour évaluer les chercheurs et la recherche. *Le deuxième avantage* c'est que ces indicateurs bibliométriques étant basés sur de larges événements (grandes quantités de décisions des évaluateurs) par rapport au processus d'évaluation par les pairs limité et réalisé dans des contextes spécifiques, cela évite les biais dus à un manque d'information de la part des évaluateurs. En plus étant donné un nombre large suffisant d'événements d'évaluation par les pairs, différents motifs pour citer un article neutralisent d'autres motifs ce qui diminue les biais de l'évaluation par les pairs. Mais malgré ces avantages des indicateurs bibliométriques il existe un certain nombre d'inconvénients (des problèmes méthodologiques et techniques).

Les critiques adressées envers l'évaluation bibliométrique de la qualité de recherche :

Les problèmes techniques et méthodologiques des indicateurs bibliométriques :

Les indicateurs bibliométriques peuvent avoir des problèmes techniques ainsi que des problèmes méthodologiques. *Les problèmes techniques* peuvent être des erreurs dans le traitement des données, ou des problèmes statistiques causés par exemple par une insuffisance de nombre de citations disponibles à l'analyse. *Les problèmes méthodologiques* peuvent être des problèmes dans la définition des champs de recherche, notamment interdisciplinaires ce qui peut mener à exclure certaines publications dans le calcul du nombre de citations. Un autre problème est dû à la nature sélective des indicateurs bibliométriques, ce qui mène à des biais de langue (l'anglais), de pays (anglophones), ou de disciplines (sciences durs). Le nombre de citations comme un indicateur de performance de recherche n'a pas la même importance entre les différentes disciplines, dans certains disciplines on cite plus que d'autres disciplines (par exemple les chercheurs en biomédecine citent 6 fois plus que les chercheurs en mathématiques). Enfin, de petites différences des résultats sont insignifiantes et ne représentent pas des différences d'effort, de productivité, de performance ou de qualité individuelle ou institutionnelle. Les décideurs publics ne devraient pas se fier aux classements basés sur de petites différences entre les concurrents. (Weingart, 2005, pp. 121-124)

Weingart donne trois conseils pour contrer et limiter ces problèmes méthodologiques et techniques. *Le premier* c'est qu'il faut appliquer les indicateurs bibliométriques avec un niveau élevé d'agrégation de données. *Le deuxième* c'est qu'il faut construire prudemment ces indicateurs. La liste des publications devrait être vérifiée par les universités et il faudrait construire différents types d'indicateurs (normaux, et ceux qui tiennent compte de la taille de

l'université vu que des travaux empiriques ont montré que les résultats ne sont pas les mêmes si on contrôle ou non pour la taille). Et *le troisième* c'est qu'il faut prudemment interpréter les résultats de ces indicateurs en gardant toujours à l'esprit les problèmes techniques et méthodologiques de ces indicateurs. (Weingart, 2005, p. 124 et p. 128)

(Van Raan, 2005) insiste sur *Les problèmes techniques* qui les considère comme étant les problèmes les plus importants des indicateurs bibliométriques., puisque selon lui l'analyse des citations est basée entièrement sur l'appariement entre les publications qui citent et les publications citées, commettre des erreurs dans cette phase d'appariement cause une perte de citations qui peut affecter considérablement les résultats d'évaluation de la qualité de recherche des universités. *Le premier problème* concerne donc le non appariement de citations (7% selon Moed 2002 cité par Van Raan, 2005), ce qui cause une perte du nombre de citations pour certains articles. Ce non appariement concerne généralement les articles écrits par plusieurs auteurs, ou des erreurs dans les noms d'auteurs (surtout les noms non anglophones) ou des erreurs dans les références bibliographiques (le numéro de la revue, les numéros de pages...etc). Ce non appariement est inégalement distribué, dans certains cas il peut aller jusqu'à 30%. *Le deuxième problème* technique est le problème d'attribution des publications aux bonnes institutions, ce problème est causé par l'existence de différents noms des universités (par exemple Leiden University a 5 variantes de nom si les publications de ces 5 variantes ne sont pas regroupés dans le classement de Shanguai comme une seule université de Leiden, cette université pourrait perdre une part considérable de nombre de citations et donc de qualité de recherche évaluée). Parfois l'article publié ne mentionne pas l'université (par exemple en France le CNRS est mentionné au lieu du nom de l'université où la recherche a été réalisé), ainsi que différents départements universitaires, mais aussi des problèmes d'adresses des universités. Connaître l'adresse de l'université ne suffit pas toujours à bien les identifier, puisque parfois plusieurs universités se situent dans la même ville ou dans la banlieue.

Les problèmes méthodologiques des indicateurs bibliométriques basés sur une analyse des citations sont selon Van Raan : Un biais contre les sciences humaines et sociales et en faveur des sciences dures et naturelles. Un biais linguistique en faveur de l'anglais. Un biais des données de citation en faveur des Etats-Unis (la plupart des revues indexés dans la base de données ISI sont des revues américaines. L'indicateur impact des revues pourrait être dominés par les surveys qui ne sont pas des travaux de recherche innovants. Il faut faire la

différence entre les types des travaux de recherche publiés dans les revues (article de survey de littérature, articles innovants, lettres ou notes) ils devraient pas avoir la même valeur dans l'évaluation de la qualité de recherche, publier un article innovant devrait avoir plus de valeur que de publier un survey de littérature. (par exemple le classement de Shanghai ne tient pas en compte ces différences entre les types d'articles).

L'utilisation de l'index de citations dans les classements des universités (notamment le classement de Shanghai) pose un problème majeur : l'index ne reflète pas la qualité de recherche actuelle des universités ni celle des jeunes chercheurs. Puisque cet index inclue des citations qui peuvent remonter à des dizaines d'années (old boy effect,) par exemple le classement de Shanghai donne un poids de 60% pour les indicateurs basés sur une analyse de citations, (Van Raan, 2005, p. 138) estime que cela introduit une incertitude dans les positions du classement de 5 à 10 positions pour le classement des universités européennes et une incertitude de 25 à 50 positions pour le classement mondial.

Des problèmes spécifiques de l'indicateur facteur d'impact des journaux :

Selon (Weingart, 2005, pp. 126-128) parmi tous les indicateurs bibliométriques utilisés pour évaluer la qualité de recherche, l'indicateur facteur d'impact des revues, pose des problèmes majeurs. L'évaluation des chercheurs par rapport au nombre de publications dans des revues qui ont le plus d'impact (les revues les plus citées) selon l'indicateur facteur d'impact des revues est dangereuse pour la qualité de la recherche, à cause de plusieurs problèmes : L'indicateur est peu fiable dans la détermination de la qualité de recherche à cause de la distribution inégale des citations d'articles dans n'importe quelle revue, puisque dans la même revue certains articles seront très cités, d'autres articles sont très peu cités. Un autre problème c'est que généralement la revue n'aide pas l'article mais le contraire, un article peu important ne sera pas très cité même s'il est publié dans une revue à fort impact. Aussi cet indicateur est problématique à cause de différentes pratiques de citation selon les différents champs de recherche, donc il ne peut représenter la qualité de recherche dans les champs de recherche qui utilisent peu les citations (notamment les mathématiques). Et finalement on trouve des articles très cités publiés dans des revues peu renommées et le contraire.

Il y'a aussi un biais de la langue (van raan, 2005, p.139) en effet l'index des citations contient des publications non anglophone, or vu que ces publications sont très peu cités, leur impact serait très faible, ce qui est en défaveur des institutions et chercheurs qui publient des publications dans d'autres langues que l'anglais. Il existe une preuve empirique qui montre

que l'inclusion de ces publications non anglophones diminue le facteur d'impact de 25 % (Van Leeuwen et al, 2001 cité par Van Raan, 2005 p. 139) en d'autres termes enlever les publications allemandes du département médecine d'une université allemande augmente son impact de 25%. (Van Raan considère qu'enlever les publications non anglophones dans calcul du facteur d'impact est équitable dans un contexte de comparaison internationale pour ne pas défavoriser ceux qui publient dans d'autres langues que l'anglais). Mais malgré tous ces problèmes l'indicateur impact de revues reste l'indicateur bibliométrique le plus utilisé dans l'évaluation de la qualité de recherche.

Les effets négatifs potentiels de l'évaluation bibliométrique sur la qualité de recherche :

Quels sont les effets de l'évaluation de la recherche par les indicateurs bibliométriques sur la qualité de recherche ? (Weingart, 2005, pp. 124-126) dit qu'on ne sait pas encore, cela dépend de leur utilisation. Cependant selon Weingart, on possède quelques informations qui peuvent nous donner des indices sur les effets potentiels de l'évaluation de la recherche par les indicateurs bibliométriques. Ce qu'on sait jusqu'ici c'est que les mesures bibliométriques sont des mesures réactives dans le sens que les chercheurs réagissent et changent leur comportement après l'utilisation de l'évaluation bibliométrique. Ce **changement de comportement est intentionnel** notamment quand ces indicateurs bibliométriques sont utilisés pour affecter des fonds de recherche, en l'occurrence, l'objectif de l'agence de financement est d'augmenter la concurrence entre les chercheurs pour améliorer la qualité de recherche. Autre chose qu'on sait c'est que chaque indicateur bibliométrique a **un objectif unidimensionnel**, par exemple l'objectif des indicateurs qui mesurent les recherches réalisées avec des fonds extérieurs est d'augmenter la concurrence entre les chercheurs pour déposer leurs candidatures et obtenir ces fonds. Un autre exemple est l'indicateur du nombre des doctorants par professeur (ou institution) qui est utilisé comme un proxy de la qualité de recherche du professeur ou de l'institution concernée, mais a aussi comme objectif d'augmenter le nombre de doctorants à l'université. Un autre indicateur est celui du nombre d'articles publiés qui a pour objectif d'augmenter la quantité de la recherche.

Cependant, les chercheurs peuvent manipuler et certains manipulent déjà les effets intentionnels de pilotage et d'orientation de l'évaluation par les indicateurs bibliométriques, ce qui cause des effets non intentionnels sont : les chercheurs font des projets de recherches conservateurs et non risqués au lieu de faire des projets de recherche risqués et innovateurs. Les chercheurs peuvent travailler sur la forme plus que sur le fond pour ainsi pouvoir publier

dans les revues les plus renommées. Les chercheurs augmentent le nombre de publications en divisant leurs publications. Et les chercheurs peuvent diminuer leurs standards pour leurs étudiants doctorants juste pour augmenter le nombre de leurs doctorants.

Il existe des effets non prévus ou non intentionnels sur la qualité de recherche des chercheurs ainsi que sur les politiques éditoriales des revues. Pour *les effets sur la qualité de recherche*, on note que l'objectif devient de publier dans les revues qui ont le plus d'impact selon l'indicateur impact des revues au lieu de se concentrer sur la qualité de recherche à publier. Un autre effet c'est l'augmentation du nombre de survey de littérature au détriment des articles de recherche innovants, puisque les surveys de littérature sont généralement plus cités. L'indicateur impact des revues devient une fin en soi. Pour *les effets sur les politiques éditoriales* des revues (les publications deviennent une affaire commerciale plus que professionnelle et scientifique). On commence à remarquer une manipulation du processus de communication (par exemple en médecine la revue shock demande à ceux qui veulent publier de citer la revue dans leurs articles ainsi que d'envoyer des copies de la revue à leurs collègues après la publication ! Même si Eugene Garfield (Garfield, 1997 cité par Weingart) défend cette pratique en soutenant l'idée que ça aide à contrer l'effet de Mathew, selon Weingart c'est un argument difficilement convaincant. Il peut y avoir d'autres effets non prévu sur le comportement des individus et des organisations. Un exemple est que évaluer par cet indicateur peut mener les chercheurs à bâcler leur travail de recherche pour publier le plus rapidement possible et obtenir des fonds, en plus c'est une pression de publier sans tenir compte de la qualité ça peut même mener à la fraude.

Il existe aussi plusieurs effets négatifs potentiels d'une utilisation de l'évaluation bibliométrique dans les décisions budgétaires d'allocation de fonds de recherche, ces effets sont : un biais de la quantité contre la qualité. Des performances qui visent le court terme au détriment du long terme, une utilisation de théories et d'approches de courants principaux conventionnels, non risqués mais surtout non innovants. La publication dans des revues indexées dans la base de donnée ISI devient un objectif en soi pour augmenter le nombre de publications et obtenir des fonds de recherche.

La communauté scientifique a connu la concurrence avant la bibliométrie, mais ils n'ont jamais connu un eaudience extérieure ni des commentateurs employants le langage des évènements sportifs quand ils parlent de l recherche scientifique et des chercheurs ! (par

exemple ISI highly cited scientists in science and watch, january-february headline reads :
“astrophycisist andrew fabian on rocketing to prominence” (Weingart, 2005, pp. 128-129)

Avec tous ces problèmes et critiques des indicateurs bibliométriques et de leur utilisation dans l'évaluation de la qualité de recherche, est ce qu'il faut les utiliser ou non dans l'évaluation de la qualité de recherche ? Et si oui comment les évaluateurs de la recherche devraient les utiliser ? Répondre à cette question nous mène à la sous partie suivante : la complémentarité entre l'évaluation bibliométrique et l'évaluation par les pairs.

La complémentarité entre les deux méthodes d'évaluation :

L'évaluation bibliométrique peut contrôler et renforcer le processus d'évaluation par les pairs (limitation des biais associés à l'évaluation par les pairs, minimisation des effets réseaux, et lutter contre l'exclusion des projets de recherché innovants qui s'inscrivent pas les domaines de recherche à la mode). Etant donné que les indicateurs bibliométriques sont basés sur une grande quantité de données, ils révèlent des comportements et schémas macro. Ils donnent la grande image et ils peuvent informer (notamment sur les connexions entre les différents champs de recherche dont on pensait qu'ils étaient non liés). Les indicateurs bibliométriques ne devraient pas remplacer le processus d'évaluation par les pairs, les deux méthodes d'évaluation devraient fonctionner ensemble, chacune complémente l'autre, il est dangereux d'utiliser les indicateurs bibliométriques indépendamment de l'évaluation par les pairs, ces indicateurs devraient être interprété par des experts ou au moins utiliser l'aide des experts dans l'interprétation des résultats des indicateurs bibliométriques.

Pourquoi il faut l'aide des experts (chercheurs bibliométriciens) dans l'interprétation des indicateurs bibliométriques ? Ils savent comment les chiffres bibliométriques sont créés et surtout ils savent ce que ces chiffres représentent et ce que ces chiffres ne représentent pas ainsi que les limites techniques et méthodologiques de ces indicateurs

Pour Weingart utiliser l'évaluation bibliométrique pour affecter des fonds sans l'aide des experts est : Trompeur, n'a pas de sens, irresponsable et non éthique puisque cela mène à de fausses décisions budgétaires dus à l'incompétence des décideurs publics et des médias. Aussi cela mène à des objectifs de court terme au détriment du long terme avec des effet néfastes profonds sur la production de connaissances sur le long terme. Par exemple en Grande Bretagne des règles de transfert des footballeurs sont appliqués aux chercheurs Selon

Weingart, le problème n'est pas dans les indicateurs bibliométriques mais dans les personnes non qualifiées qui utilisent mal ces indicateurs. (Les agences veulent rapidement des indicateurs tous prêts sans vouloir dépenser assez d'argent) (Weingart, 2005, 124-126)

6) Les critiques adressées aux classements des universités :

Les critiques adressées aux classements des universités sont très nombreuses et on trouve à peu près les mêmes critiques dans la plupart des articles traitants des classements. On a déjà abordé les critiques spécifiques pour plusieurs indicateurs, mais dans cette partie on abordera les critiques générales adressées aux classements des universités. On peut résumer ces critiques dans les points suivants :

Le choix des indicateurs et des pondérations est arbitraire et subjectif. C'est la critique la plus citée dans la littérature. D'autres critiques liées sont les suivantes : Par ce choix les auteurs du classement imposent leur propre définition de la qualité sur toutes les universités sans tenir compte ni des spécificités organisationnelles des universités ni des spécificités institutionnelles des différents pays. Par ce choix les auteurs des classements déterminent en grande partie quelles seront les universités les bien classées et les moins bien classées. Dans ce sens les résultats des classements des universités ne sont jamais objectifs.

Une des solutions au problème des pondérations est de ne pas utiliser des pondérations en laissant ce choix à l'utilisateur du classement comme le fait le classement allemand CHE. Mais cela ne règle pas le problème du choix arbitraire et subjectif des indicateurs. Une des solutions possibles pourrait être la solution adoptée dans le classement australien de Melbourne, à savoir, enquêter la communauté académique en leur demandant quels sont les indicateurs qui considèrent les plus importants pour la qualité des universités ainsi que leur pondération. Cependant, même si cette méthode règle le problème de choix arbitraire, le choix reste toujours subjectif mais sûrement beaucoup moins subjectif.

Je pense que la meilleure solution serait de conduire des recherches ayant pour but de découvrir les inputs et les outputs qui influencent le plus la qualité d'enseignement et de recherche des universités et se baser sur les résultats de ces études dans le choix des indicateurs de classements.

Le regroupement de tous les indicateurs pour créer **un score unique total** qui représente la

qualité globale de l'université est très critiquable.

Les classements qui évaluent globalement les universités **ne tiennent pas compte des différences de qualité** qui puissent exister entre les différents départements à l'intérieur des universités. Pour faire des choix informés les étudiants ont besoin d'informations non seulement de la qualité et réputation globale de l'université mais aussi de la qualité et réputation des différents départements, disciplines et programmes. Mais selon (Williams et Van Dyke, 2007, p. 822) la variation de qualité à l'intérieur de la même université est en train de diminuer, à cause des recherches et collaborations interdisciplinaires et grâce à l'utilisation croissante de procédures de contrôle de plus en plus strictes et performants s'appliquant à l'intérieur de toute l'université.

Les différences minimales des scores, ne peuvent représenter des différences réelles dans la qualité des universités. Par exemple, Clarke (2002) a réalisé une étude statistique sur les résultats de deux classements : le classement US news des universités américaines et le classement US news des écoles de commerce américaines, pour les deux types de classement elle a trouvé les différences des scores attribués aux différentes universités sont insignifiants statistiquement.

Les classements **confondent entre les notions de réputation et de qualité**. Les classements mesurent en réalité beaucoup plus la réputation que les vraies performances des universités.

L'utilisation des enquêtes d'opinion des académiques sur la qualité des différentes universités est une pratique très critiquable. Les académiques (généralement les doyens et présidents des universités) qui répondent aux enquêtes, même si on suppose qu'ils ont assez d'informations sur leur propre université, il en est pas de même quand il s'agit d'évaluer d'autres universités.

L'évaluation de **la qualité d'enseignement est négligée** dans la plupart des classements des universités

Les classements **ignorent la mission des universités de servir la société** et ne tiennent pas compte des différents besoins des différents pays en terme d'enseignement et de recherche

A cause des critiques, les médias ont été forcés dans leurs classements des universités de faire de plus en plus d'efforts dans la spécification des méthodologies utilisées et d'exercer plus de transparence. Ils améliorent et révisent régulièrement leurs méthodologies et utilisent de plus en plus l'aide des experts académiques dans l'évaluation en créant des conseils consultatifs spécialement conçus pour les classements des universités. (Van Dyke, 2005, p. 116)

6) Les réactions des parties prenantes envers les classements des universités et le rôle des classements dans le choix des universités par les étudiants :

6-1) les réactions des parties prenantes envers les classements des universités :

Pour les étudiants et leurs familles, les classements des universités sont populaires et ont été bien accueillis la raison de cet enthousiasme c'est ces classements donnent des informations comparatives sur les universités auparavant inaccessibles au grand public, ce qui aide les étudiants à faire des choix informés.

Les académiques ont critiqué les classements, en montrant leur inquiétude sur les dangers potentiels liés à des utilisations abusées de ces classements. Ainsi que la possibilité que les universités manipulent les données.

Les universités ont des opinions mitigés, certaines sont favorables aux classements, d'autres s'opposent voir même boycottent les classements en refusant de participer, certaines universités ont fait un de leurs objectifs d'améliorer leur rang dans les classements des universités. En 1999 35 universités asiatiques ont refusé de participer dans le classement Asia Week, dont l'université de Tokyo pourtant numéro 1 de ce classement en 1997 et 1998, le président de l'université de Tokyo justifiant ce refus par la difficulté de quantifier et comparer la qualité à cause des spécificités de chaque université. (Stella et Woodhouse, 2006, p. 5)

6-2) le rôle des classements dans le choix des universités par les étudiants :

Est-ce que les classements des universités jouent un rôle déterminant dans le choix des universités par les étudiants ?

Selon des enquêtes en Inde et dans d'autres pays. Ce qui influence le plus le choix de l'université c'est la proximité de la maison, réputation de l'université, la disponibilité des programmes désirés, les frais d'inscriptions, les équipements sportifs, l'accompagnement des amis). Une enquête réalisée aux Etats-Unis en 1995 sur 221 897 étudiants de première année universitaire dans 432 universités de 4 ans. 60% des étudiants enquêtés trouvent les classements non importants. 30 % trouvent les classements un peu importants et 11% trouvent les classements très importants. (McDonough, 1998) cité par (Stella et Woodhouse, 2006, p. 18) Une autre enquête en 1997 par Higher Education Research Institute, parmi les 251 232 étudiants de première année universitaire surveyés, seulement 8,6% ont considéré les classements comme très importants dans le choix d'une université. Les sources de données les plus fréquemment consultés par ces étudiants sont par ordre : les parents, les amis et camarades de classes, les étudiants inscrits ou qui vont être inscrits dans des domaines du même intérêt, les brochures des universités, les matériels et guides proposés dans les lycées, en effet seulement (Stella et Woodhouse, 2006, p. 18)

Dans une enquête britannique sur 8000 étudiants dans 12 universités, a montré que les facteurs du choix d'université les plus importants pour les étudiants sont de bons programmes qui satisfont leurs préférences, une réputation dans la qualité d'enseignement, et la qualité des équipements offerts (bibliothèque, ordinateurs) qualité des bâtiments, des classes et amphithéâtres. Ces facteurs sont difficilement représentables par des chiffres numériques pourtant ils sont importants dans le choix d'université, notamment dans les universités qui n'ont pas une longue réputation de recherche. (Utley, 2002, cité par Eccles, 2002, p.428)

Si les classements des universités jouent un rôle important dans le choix des étudiants, cela impliquerait qu'un changement de position du classement d'une université devrait affecter le nombre de candidatures qu'elle reçoit l'année suivante la publication du classement, est-ce le cas ?

Eccles a construit des graphiques qui montrent la relation entre la part des candidatures que reçoivent les universités et les résultats du classement Times. Il a étudié 3 types d'universités : les universités post-1992, les universités post-1960 et les anciennes universités d'avant 1960.

Il trouve qu'il n'y a pas de relation claire entre les classements et les candidatures des étudiants. Parfois la position du classement d'une université se détériore mais le nombre de candidatures augmente.

Il conclut qu'au moins dans le court terme les classements n'ont pas d'effets sur els choix des étudiants. Les classements reflètent les choix des étudiants plutôt que d'influencer ce choix, en effet selon Eccles les classements sont destinés aux meilleurs étudiants qui ont la possibilité de choisir, ces étudiants connaissent déjà quelles sont les universités les mieux réputées et se font concurrence féroce entre eux pour être acceptés dans ces universités. Ces universités ayant le meilleurs étudiants cela renforce leur position dans les classements.

(Eccles, 2002, pp.428-431)

Conclusion :

Notre problématique de départ était de savoir dans quelle mesure les classements des universités permettent d'évaluer la qualité dans l'enseignement supérieur, on peut conclure que les classements des universités sont des évaluateurs très imparfaits de la qualité dans l'enseignement supérieur, plusieurs améliorations méthodologiques sont nécessaires surtout si les classements veulent jouer un rôle important pour l'amélioration de la qualité et dans les décisions publiques d'allocation des ressources. Cependant, il ne faut pas oublier que les classements des universités sont une approche d'évaluation parmi pleines de d'autres approches, notamment l'approche de l'assurance qualité mais on pense surtout à l'approche d'efficacité économique. On a vu que un des défauts principaux des classements des universités c'est qu'ils ne mesurent pas la valeur ajoutée de l'université. Avec un choix arbitraire des indicateurs qui confond les inputs, les processus et les outputs. On pense que l'utilisation d'une approche d'efficacité économique pourrait aider à résoudre ces problèmes et rendre l'évaluation de la qualité des universités moins confuse et plus claire.

Bibliographie :

Antony Stella et David Woodhouse, 2006, «Ranking of Higher Education Institutions », Australian University Quality Agency, AUQA Occasional Publications Number 6. Accès le 13 juin 2008
http://www.auqa.edu.au/files/publications/ranking_of_higher_education_institutions_final.pdf

A-M. Kermarrec et al, 2007, « Que mesurent les indicateurs bibliométriques? », document d'analyse de la commission d'évaluation de l'INRIA (Institut National de Recherche en Informatique et en Automatique).

Glanzel et Moed, 2002, "Journal impact measures in bibliometric research", *scientometrics*, vol. 53, No. 2, pp. 171-193

Eccles, Charles. "The use of University Rankings in the United Kingdom." *Higher Education in Europe* 27, no. 4 (12, 2002): 424.

Federkeil, Gero. "Some Aspects of Ranking Methodology--the CHE-Ranking of German Universities." *Higher Education in Europe* 27, no. 4 (12, 2002): 389.

Marguerite, Clarke. "Some Guidelines for academic Quality Rankings" *Higher Education in Europe* 27, no. 4 (12, 2002): 389.

Usher, Alex and Massimo Savino. "A Global Survey of University Ranking and League Tables." *Higher Education in Europe* 32, no. 1 (2007): 5-15.

Van Dyke, Nina. "Twenty Years of University Report Cards." *Higher Education in Europe* 30, no. 2 (2005): 103-125.

van Raan, Anthony F. J. "Fatal Attraction: Conceptual and Methodological Problems in the Ranking of Universities by Bibliometric Methods." *Scientometrics* 62, no. 1 (01/01, 2005): 133-143.

Weingart, Peter. "Impact of Bibliometrics upon the Science System: Inadvertent Consequences?" *Scientometrics* 62, no. 1 (01/01, 2005): 117-131.

Williams, Ross and Nina Van Dyke. "Reputation and Reality: Ranking Major Disciplines in Australian Universities." *Higher Education* 56, no. 1 (07/23, 2008): 1-28.

———. "Measuring the International Standing of Universities with an Application to Australian Universities." *Higher Education* 53, no. 6 (06/03, 2007): 81

