

De la difficulté à établir des comparaisons non biaisées dans les enquêtes internationales

Emmanuel Dion

Dans la plupart des enquêtes internationales, l'un des intérêts majeurs de la démarche, la comparaison entre différentes cultures, se heurte à un problème méthodologique récurrent: l'impossibilité où se trouve l'analyste d'attribuer sans ambiguïté les différences constatées au fond (le sujet étudié) ou à la forme (la méthode d'investigation retenue). Au-delà de la simple question de la traduction en effet, il existe ce qu'on pourrait appeler une forme culturellement dominante du jeu des questions et des réponses, qui biaise ou teinte les données recueillies, et dont on ne peut se débarrasser tout à fait.

Prenons le cas de la différence de "style culturel de réponse"¹ entre les Anglais et les Italiens. Le cliché habituel est que les Italiens font preuve d'exagération ou d'outrance dans leurs réponses (en moyenne et quelles que soient les questions posées si l'on part du principe qu'il s'agit d'un trait culturel dominant) alors que les Anglais émettent de préférence des jugements nuancés, voire équivoques.

Dans le cas de l'enquête OCHA, cette vision stéréotypée est confirmée par l'analyse objective des données: si l'on observe les indicateurs usuels de dispersion statistique, on observera qu'à l'ensemble des 16 questions posées sous forme d'échelle d'évaluation numérique de 0 à 10, la variance moyenne est de 9,46 pour les Italiens (n=904) et de seulement 7,13 pour les Anglais (n=900; p<0,001)².

Que cela signifie-t-il? A l'évidence que les Italiens notent différemment des Anglais. Mais cela signifie-t-il que les Italiens *pensent* différemment des Anglais? Il est bien difficile de le dire. En réalité, le biais du questionnement existe quelle que soit la forme utilisée (une échelle d'évaluation, une question à choix multiple ou une question ouverte) et la distance du langage dresse inévitablement un obstacle entre la pensée du répondant et l'image que l'analyste peut s'en

¹ C'est la formulation de synthèse, simplificatrice mais pratique, que nous retiendrons dans la suite du texte.

² Les variances mentionnées ici sont les moyennes des variances mesurées sur chaque individu pris séparément. Elles n'indiquent donc pas une dispersion au sein du groupe (qu'on pourrait imaginer comme un indicateur de désaccord entre ses membres) mais la propension de chacun à exprimer des opinions plus ou moins tranchées. La variance calculée sur la totalité des individus de même nationalité est de 9,65 pour les Italiens et de 7,60 pour les Anglais. On voit qu'on retrouve entre les deux groupes une différence de variance d'environ 2; simplement, les deux variances se situent l'une et l'autre à un niveau légèrement supérieur à celui atteint pour le calcul individuel, ce qui se justifie par l'existence d'un style de réponse individuel s'ajoutant au style de réponse culturel. Si tous les individus d'un même groupe utilisaient la même fréquence de notation, toutes les variances individuelles seraient égales entre elles, et auraient la même valeur que la variance du groupe. La force relative des styles de réponse individuels et culturels (autrement dit la propension de certains individus à se singulariser au sein même de leur culture) peut être exprimée par ce rapport entre les variances moyennes des individus et la variance du groupe.

Par ailleurs, et bien que les deux approches soient défendables, nous avons préféré travailler dans toute cette recherche sur des variances estimées à base d'échantillon (formule 1) plutôt que sur des variances de population (formule 2). Les deux approches donnent de toute façon des résultats très proches.

Formule 1 : "VAR" sous Excel,
$$\frac{n\sum x^2 - (\sum x)^2}{n(n-1)}$$

Formule 2 : "VAR.P" sous Excel,
$$\frac{n\sum x^2 - (\sum x)^2}{n^2}$$

faire. Aussi bien, il n'y aura jamais de réponse définitive à la question de l'attribution des différences constatées à l'objet étudié ou au style culturel des réponses.

Tout au plus peut-on chercher à décrire précisément les principales caractéristiques de ces styles culturels de réponse de façon :

- 1 - à en abstraire la partie commune, indépendante des thèmes abordés ou des questions posées, qui pourrait alors servir de base de calcul à un redressement des résultats
- 2 - à en avertir le lecteur de l'étude pour que, les ayant alors clairement présents à l'esprit lors de l'examen des données, il se sente libre de les prendre ou non en considération.

Pour réaliser efficacement une telle description, l'idéal serait sans doute de disposer d'un questionnaire spécifiquement étudié pour aborder un certain nombre de sujets culturellement neutres, de façon que le fond n'ait pas d'incidence sur les variations de forme (ce qui est, pour des raisons méthodologiques évidentes, exactement l'inverse de ce qu'on cherche habituellement à obtenir). Cependant, nous rencontrons là encore la difficulté majeure qu'il y a à vouloir dissocier le fond de la forme, difficulté qui se transmue en impossibilité pure et simple si nous faisons l'hypothèse que tout sujet recèle une composante, fût-elle minime, de culturalité.

En l'absence d'un véritable étalon de mesure permettant de trancher l'affaire une fois pour toutes³, nous ne pouvons donc guère faire autrement que d'accepter de nous servir tout de même des données disponibles (celles du questionnaire OCHA), en reconnaissant délibérément le biais de fond qu'elles présentent. L'analyse peut alors partir, assez classiquement, de l'observation des différences de moyennes, pour en arriver ensuite à une comparaison des variances et à la définition, au final, des styles culturels de réponse agrégés dans leurs multiples composantes. Dans un souci de simplification, nous présenterons d'abord les résultats comparés des Anglais et des Italiens, particulièrement contrastés, avant de livrer en appendice quelques résultats plus complets sur l'ensemble des groupes.

1 - Les styles de moyenne

Il existe pour commencer une différence évidente qui s'exprime par la propension, au sein de chaque culture, à utiliser les différentes notes possibles avec une fréquence variable. La figure 1, qui donne un aperçu de ces différences, mérite quelques éclaircissements car sa lecture n'est pas immédiate. Plusieurs biais (connus) s'y juxtaposent en effet:

- Un biais de polarité: les notes supérieures à la moyenne sont largement plus employées que les notes inférieures à la moyenne
- Un biais de centralité: la réponse neutre (5/10) recueille un nombre anormalement élevé de suffrages, ce qui se traduit par un pic au milieu de l'histogramme des notes
- Un biais de parité: la réponse "8/10", par exemple, est beaucoup plus utilisée que les réponses "7/10" ou "9/10"
- Un biais d'extrémisme: les deux notes les plus extrêmes, 0/10 et 10/10, sont davantage distribuées qu'elles ne le seraient sur un histogramme normal ou Gaussien

³ Une recherche bibliographique sur le sujet n'a de plus donné aucun résultat probant.

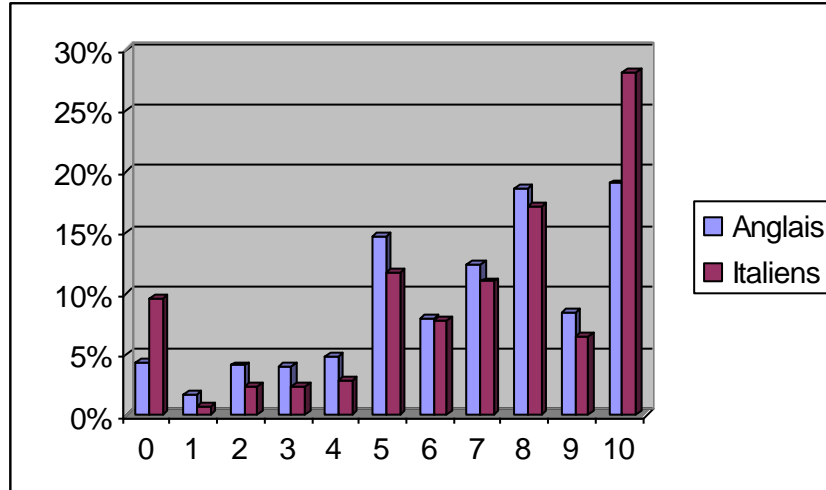


Figure 1 - Comparaison de la distribution des fréquences de notes entre Anglais et Italiens

Cependant, tous ces biais ne sont pas inquiétants. La théorie de la mesure en sciences sociales les a largement étudiés⁴ et a aidé à comprendre, d'une part comment on pouvait en partie les limiter, d'autre part pourquoi on devait en partie les accepter. Dans le cas qui nous préoccupe, ils sont d'autant moins inquiétants qu'on les retrouve à la fois chez les Anglais et chez les Italiens (bien que dans des proportions variables).

D'ailleurs, si l'on extrayait les quatre biais listés du tableau des fréquences, ce qu'on peut faire numériquement, mais aussi par un simple effort d'abstraction visuelle, on retomberait sur un histogramme quasi Gaussien.

Au-delà de cette première analyse, nous pouvons également observer dans les données présentées l'effet qui est au cœur de notre problématique: les Italiens donnent bien plus fréquemment des notes extrêmes (0/10 et 10/10) que les Anglais. Pour toutes les notes intermédiaires en revanche, les Anglais les devancent. Cette différence de comportement de notation a un effet sur la moyenne de chaque individu⁵. Les Italiens seront en effet sensiblement plus nombreux que les Anglais à aboutir à une moyenne élevée. Le graphique suivant (Figure 2) montre en effet que si l'histogramme des moyennes anglaises comme celui des moyennes Italiennes est très régulier, d'allure Gaussienne, toutes les catégories de moyennes inférieures à 7 sont sur-représentées pour les Anglais, alors que presque toutes les catégories de moyennes supérieures à 7 sont au contraire sur-représentées pour les Italiens.

⁴ Voir par exemple le classique: W. Torgerson, *Theory and Methods of Scaling*; Wiley, 1958.

⁵ Attention: comme pour les variances mentionnées plus haut, il s'agit bien ici d'un calcul fait sur les moyennes par individu. Un tel calcul peut sembler contre-intuitif en ce qu'il opère en sens inverse de l'analyse habituelle, les moyennes étant en général calculées sur les réponses de tous les individus à une même question. C'est précisément pour dégager les phénomènes de biais formels des mesures de fond qu'il convient ici d'analyser au contraire les moyennes que donne chaque individu pris séparément à l'ensemble des questions.

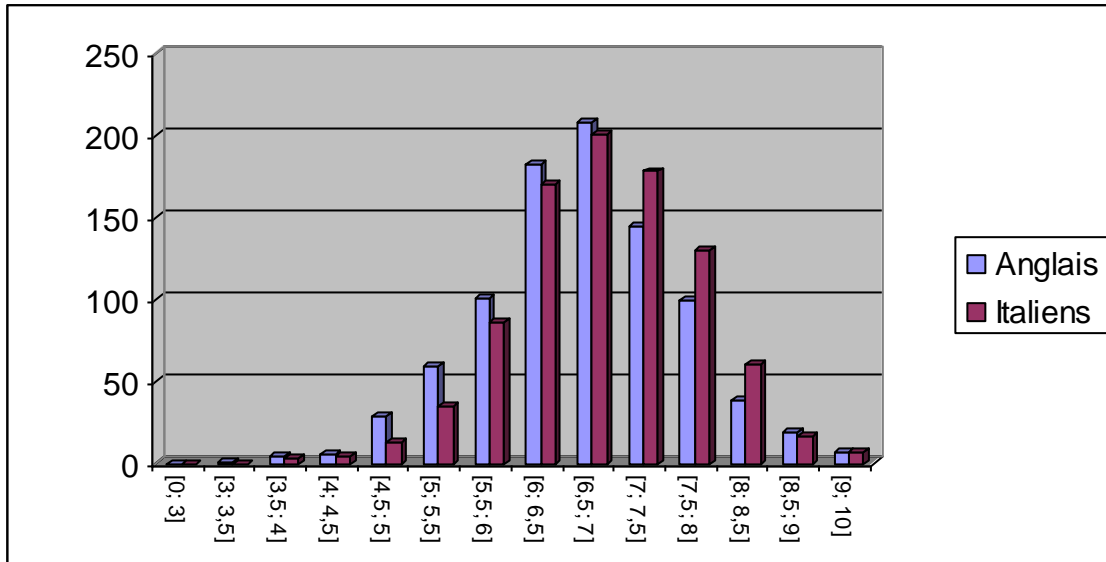


Figure 2 - Comparaison de la distribution des moyennes des individus

En conséquence, la moyenne générale des notes données à toutes les questions par tous les Anglais est de 6,64 contre 6,82 pour les Italiens. Cette différence de 0,18, modeste mais tout de même largement significative au seuil de 0,001 compte tenu du nombre important de données recueillies, doit bien être prise en compte dans l'interprétation des résultats. Toute analyse ponctuelle qui montrerait, pour une question donnée, une différence Anglais/Italiens de l'ordre de 0,2, et qui donc passerait les tests statistiques d'usage en termes de niveau de signification, serait pourtant bien plus vraisemblablement justifiée par les styles culturels de réponse que par une différence de fond.

2 - Les styles de variance

L'analyse d'autres différences de comportement en matière de variance fait également apparaître des phénomènes dont la régularité atteste la validité. Pour les classes de variances inférieures à 10, les Anglais sont en effet systématiquement plus représentés que les Italiens, situation qui s'inverse au-delà de 10. L'effet est remarquablement fort pour les variances comprises entre 2 et 6, pour lesquelles les Anglais sont près de deux fois plus nombreux que les Italiens, et plus encore pour les variances supérieures à 14, pour lesquelles les Italiens sont plus de trois fois plus nombreux (Figure 3).

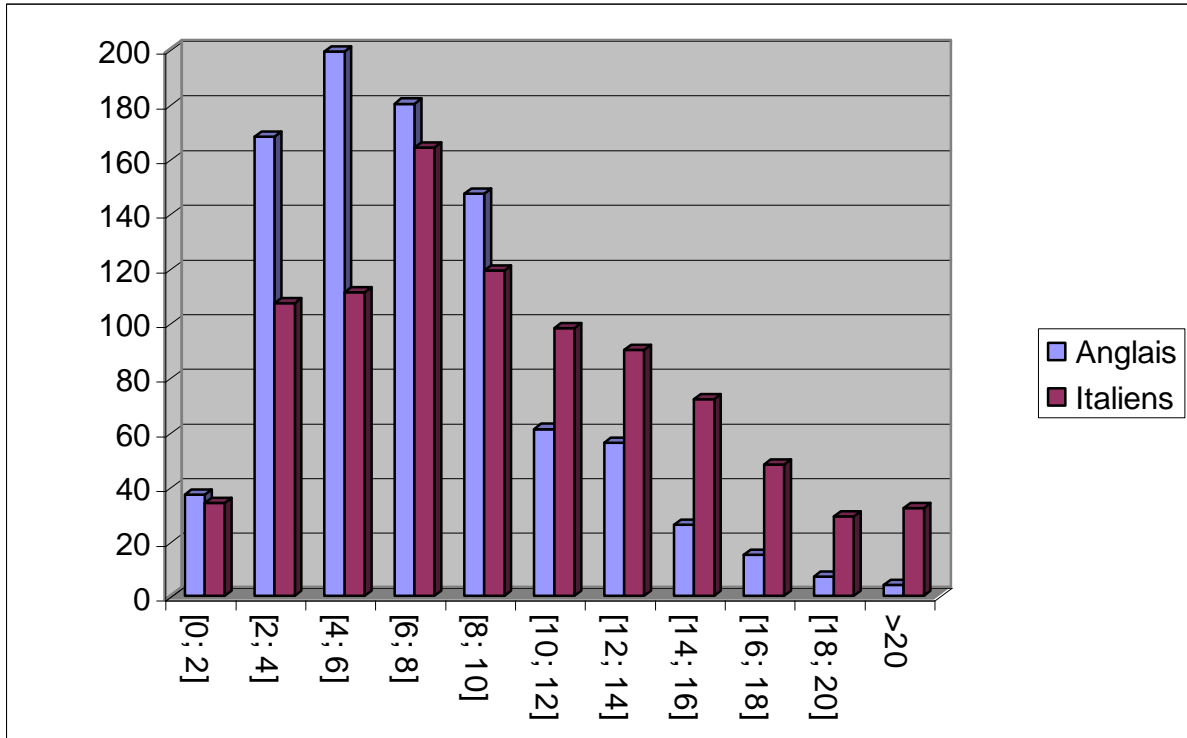


Figure 3 - Comparaison de la distribution des variances des individus

3 - Les styles agrégés moyenne/variance

On pourrait alors penser qu'il existe un lien entre la différence de comportement en matière de moyenne et en matière de variance. Ceux-là mêmes qui donnent des moyennes extrêmes ne jouissent en effet pas du même degré de liberté en matière de dispersion que ceux qui se rassemblent autour de la moyenne⁶.

L'étude de cette hypothèse conduit, non pas à invalider, mais à confirmer les conclusions qui précèdent. En effet, les Anglais se trouvant noter en général plus près de la moyenne, ils disposent a priori d'une plus grande marge de manœuvre que les Italiens, plus nettement polarisés vers les notes positives (un 5 peut être la moyenne d'un 0 et d'un 10, mais un 9 ne peut au mieux qu'être la moyenne d'un 8 et d'un 10). Que leur variance soit en général plus faible ne peut donc aucunement être expliqué par le calage de leur moyenne vers le centre (c'est même l'inverse qui devrait être attendu).

Pour être tout à fait précis dans cette affaire, il est possible d'analyser les variances constatées pour les Anglais et les Italiens *par tranche de moyenne*. Ce calcul inhabituel se justifie par la généralisation de la question suivante: deux moyennes identiques d'un Anglais et d'un Italien sont-elles en général obtenues de la même façon, ou bien par la confrontation de notes modérément contrastées de l'Anglais et de notes plus fortement contrastées de la part de l'Italien?

⁶ E. Dion, How to compare standard deviations free of trend ?, *cahiers du CREA*, décembre 1998.

Cette analyse, à notre sens plus puissante que toutes celles qui précèdent, est particulièrement édifiante. Elle montre très clairement que quelle que soit la tranche de moyenne dans laquelle on se situe, la variance d'un Italien est en moyenne nettement supérieure à celle d'un Anglais (Figure 4). Cela nous paraît prouver que l'exagération et la prudence sont deux caractéristiques fondamentales et permanentes des styles de réponse italien et anglais.

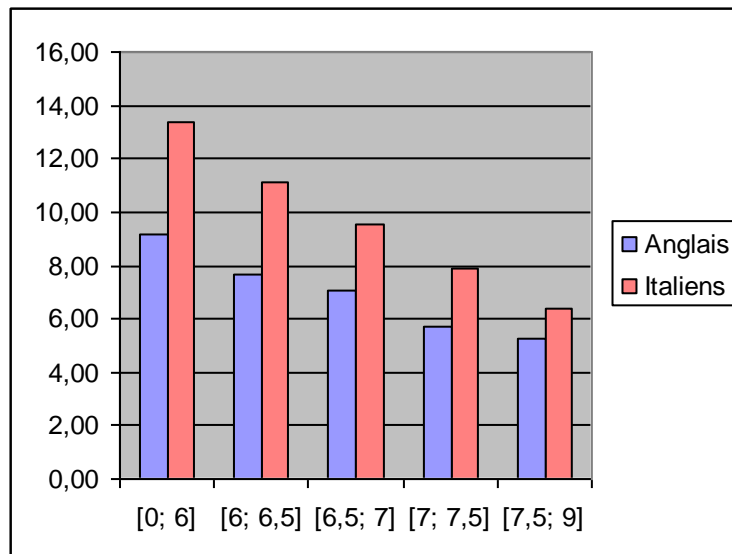


Figure 4 - Comparaison des variances par catégorie de moyennes

4 - Les conséquences en termes de risques dans l'interprétation des données

En quoi les biais mis en évidence ici peuvent-ils se révéler perturbants pour l'analyse des données collectées? Prenons un exemple simple. Soit la question du degré d'accord avec l'affirmation : "Pour avoir une alimentation saine, il suffit de faire trois repas par jour". Les résultats, donnés dans le tableau 1, montrent qu'une proportion voisine d'Anglais et d'Italiens sont d'accord avec l'affirmation (75% et 83%). Seulement, les Italiens l'expriment plus fortement que les Anglais. Ils sont en effet bien plus nombreux à répondre "tout à fait d'accord" plutôt que "plutôt d'accord" (50% contre 32%), et c'est évidemment l'inverse chez les Anglais (20% contre 56%).

Pour avoir une alimentation saine, il suffit de faire trois repas par jour:	Anglais	Italiens	Total
Pas du tout d'accord	14 1,56%	59 6,56%	73 4,06%
Plutôt pas d'accord	206 22,97%	98 10,89%	304 16,92%
Plutôt d'accord	502 55,96%	292 32,44%	794 44,18%
Tout à fait d'accord	175 19,51%	451 50,11%	626 34,84%
Total	897 49,92%	900 50,08%	1797 100%

Tableau 1 - Réponses à la question de l'importance des trois repas quotidiens

La question est alors de savoir si le surcroît d'adhésion des Italiens relève d'une différence de fond (les Italiens pensant réellement que les trois repas par jour sont plus importants) ou bien d'une simple différence de style de réponse, les Anglais pensant tout comme les Italiens que ces repas sont importants mais l'exprimant avec plus de réserve.

L'argument décisif qui pourrait faire pencher pour la seconde hypothèse est que de l'autre côté du spectre d'évaluation, ce sont également les Italiens qui sont plus nombreux à exprimer un net désaccord avec la proposition.

Alors, s'agit-il là d'un simple paradoxe du point de vue de la question posée, ou plutôt d'une confirmation que sur ce point comme sur d'autres peut-être, les styles culturels de réponse ont pris l'ascendant sur le fond du problème? Nous avons souligné dès le début de l'analyse qu'il est bien difficile d'apporter une réponse définitive à cette question. Mais la fréquence avec laquelle les configurations de ce type ont été observées dans l'enquête OCHA invite en tout cas à la plus grande prudence sur le sujet.

Appendice

	0	1	2	3	4	5	6	7	8	9	10
Anglais	4%	2%	4%	4%	5%	15%	8%	12%	19%	8%	19%
Italiens	10%	1%	2%	2%	3%	12%	8%	11%	17%	6%	28%
Français	8%	1%	4%	5%	4%	16%	7%	11%	19%	6%	18%
Américains	4%	1%	3%	4%	4%	14%	7%	11%	17%	8%	25%
Suisses	7%	1%	4%	4%	4%	15%	7%	9%	18%	6%	23%
Allemands	8%	2%	4%	6%	5%	16%	6%	9%	17%	6%	21%

Tableau 2 - Fréquence d'utilisation des différentes notes par nationalité

	Variances moyennes entre bornes				
	[0;6]	[6;6,5]	[6,5;7]	[7;7,5]	[7,5;9]
Suisses	11,34	9,91	8,34	7,08	5,03
Allemands	10,81	9,00	9,06	7,71	6,03
Anglais	9,19	7,63	7,10	5,70	5,29
Italiens	13,37	11,09	9,58	7,88	6,41
Américains	9,01	8,03	7,96	7,15	5,72
Français	10,04	8,29	6,98	6,92	6,00

Tableau 3 - Variances des individus par tranche de moyenne et par nationalité

	Variance
Suisses	8,91
Allemands	9,34
Anglais	7,13
Italiens	9,46
Américains	7,32
Français	8,35

Tableau 4 - Variance moyenne des individus par nationalité

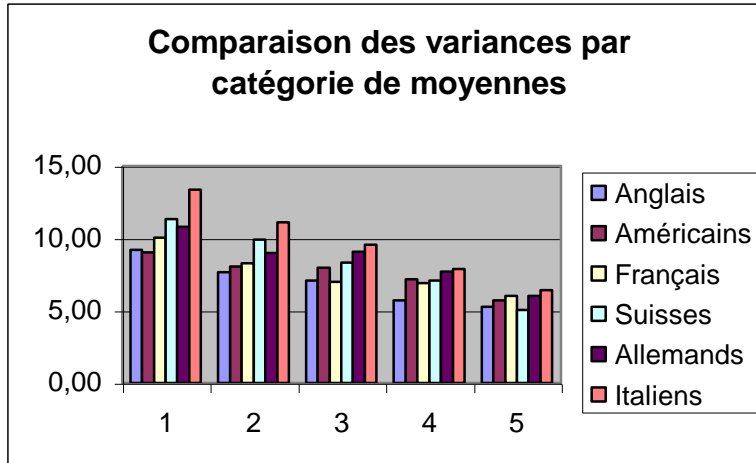


Figure 5 - Comparaison des variances par catégorie de moyennes pour toutes les nationalités

Dans la plupart des enquêtes internationales, l'un des intérêts majeurs de la démarche, la comparaison entre différentes cultures, se heurte à un problème méthodologique récurrent: l'impossibilité où l'on se trouve d'attribuer sans ambiguïté les différences constatées au fond (le sujet étudié) ou à la forme (la méthode d'investigation retenue). Au-delà de la simple question de la traduction en effet, il existe ce qu'on pourrait appeler une forme culturellement dominante du jeu des questions et des réponses, qui biaise ou teinte les données recueillies, et dont on ne peut se débarrasser tout à fait.

Le cas de la différence de "style culturel de réponse" entre les Anglais et les Italiens est le plus exemplaire. Le cliché habituel est que les Italiens font preuve d'exagération ou d'outrance dans leurs réponses (en moyenne et quelles que soient les questions posées si l'on part du principe qu'il s'agit d'un trait culturel dominant) alors que les Anglais émettent de préférence des jugements nuancés, voire équivoques.

Cette vision stéréotypée est confirmée par l'analyse objective des données de l'enquête OCHA: on y observe en particulier qu'à l'ensemble des 16 questions posées sous forme d'échelle d'évaluation numérique de 0 à 10, la variance moyenne est de 9,46 pour les Italiens (n=904) et de seulement 7,13 pour les Anglais (n=900; $p < 0,001$). Cette analyse globale est confirmée par un ensemble de mesures plus détaillées qui montrent que quelle que soit la tranche de moyenne dans laquelle on se situe, la variance d'un Italien est en moyenne nettement supérieure à celle d'un Anglais, ce qui tend à prouver que l'exagération et la prudence sont deux caractéristiques fondamentales et permanentes des styles de réponse italien et anglais.

Les biais mis en évidence peuvent se révéler perturbants pour l'analyse des données collectées en raison de leur fréquence d'apparition élevée dans l'enquête OCHA, apparition qui empêche d'attribuer clairement certaines différences constatées à une divergence d'opinion ou à une simple variation dans les styles de réponses des personnes interrogées.