

D.T. 2/96  
Décision rendue le 15 février 1996

LOI CANADIENNE SUR LES DROITS DE LA PERSONNE  
(L.R.C., 1985, ch. H-6 (version modifiée))

TRIBUNAL DES DROITS DE LA PERSONNE

ENTRE:

L'ALLIANCE DE LA FONCTION PUBLIQUE DU CANADA

la plaignante

- et -

COMMISSION CANADIENNE DES DROITS DE LA PERSONNE

la Commission

- et -

CONSEIL DU TRÉSOR

l'intimé

DÉCISION DU TRIBUNAL

Tribunal: Donna Gillis, présidente  
Norman Fetterly, membre  
Joanne Cowan-McGuigan, membre

Comparutions: Andrew Raven  
Avocat de l'Alliance de la fonction publique  
du Canada

Rosemary Morgan et René Duval  
Avocats de la Commission canadienne des  
droits de la personne

Duff Friesen, Lubomyr Chabursky et Deborah Smith  
Avocats du Conseil du Trésor

Lieu de l'audience: Ottawa (Ontario)

## TABLE DES MATIERES

I. INTRODUCTION .....	1
II. QUESTION A TRANCHER .....	14
III. LES TEXTES LÉGISLATIFS .....	15
IV. FARDEAU DE LA PREUVE .....	44
V. NORME DE PREUVE .....	53
VI. RÉSUMÉ DES FAITS .....	58
A. LE PLAN WILLIS .....	58
B. LE PROCESSUS WILLIS .....	63
(i) Collecte des données .....	65
(ii) Le questionnaire Willis .....	66
(iii) Coordonnateurs .....	70
(iv) Réviseurs .....	74
C. LE PROCESSUS D'ÉVALUATION .....	83
(i) Comité directeur .....	84
(ii) Comités d'évaluation multiples .....	88
(iii) Processus d'évaluation des questionnaires ..	93
(iv) Formation des comités d'évaluation multiples	97
(v) Évaluations du Comité directeur .....	102
(vi) Évaluations des comités d'évaluation multiples	110
(vii) Recyclage des comités d'évaluation multiples	126
(viii) Récapitulation .....	126
D. TESTS DE FIABILITÉ .....	127
(i) Tests de fiabilité inter-évaluateurs (FIE)	127
(ii) Application des tests de fiabilité inter- évaluateurs (FIE) aux comités d'évaluation multiples ..	141
(iii) Tests de fiabilité inter-comités (FIC) ..	142
(iv) Application des tests FIC aux comités d'évaluation multiples .....	144
(v) Réévaluations Wisner .....	152
E. LA COMMISSION .....	170
(i) L'enquête de la Commission .....	172
(ii) Analyse de M. Sunter .....	208
F. ROLE DES CONSULTANTS DANS LES RÉÉVALUATIONS ...	225
G. OPPORTUNITÉ DE RAJUSTER LES RÉSULTATS - LES EXPERTS	236
VII. DÉCISION ET ANALYSE .....	240
VIII. CONCLUSION .....	275

## I. INTRODUCTION

1. La Commission canadienne des droits de la personne (la «Commission») est constituée en vertu de la Loi canadienne sur les droits de la personne, L.R. (1985), ch. H-6 dans sa forme modifiée (la «Loi») et est une partie à la présente plainte représentant l'intérêt public.
2. La Commission a présenté six témoins qui avaient la compétence pour témoigner à titre d'experts. Le premier témoin que nous avons entendu fut Mme Nan Weiner, une experte en équité salariale et en rémunération. Le deuxième expert à témoigner fut M. Norman D. Willis, un expert en équité salariale et en évaluation d'emplois. Ont ensuite témoigné deux spécialistes de la statistique, M. Richard Shillington, un expert en analyse de données, et M. Alan Sunter, un expert en statistique. La Commission a aussi cité comme témoins deux de ses employés, MM. Paul Durber et James Sadler. M. Durber est un expert en équité salariale, en évaluation d'emplois et en d'autres domaines généraux de l'évaluation d'emplois, et M. Sadler est un expert en équité salariale et en évaluation d'emplois.
3. L'intimé, le Conseil du Trésor (l'«employeur»), est l'employeur des fonctionnaires qui travaillent dans la fonction publique du Canada au sein d'un ministère ou secteur mentionné à l'annexe 1, partie 1 de la Loi sur les relations de travail dans la fonction publique, 1966-1967, ch. 72, art. 1, p. 35, annexe 1 (la «LRTFP»). En plus de M. Willis, l'employeur n'a cité qu'un seul expert à témoigner, M. Fred Owen. M. Owen, un ancien consultant de M. Willis, est un expert en équité salariale et en évaluation d'emplois.
4. La plaignante, l'Alliance de la Fonction publique du Canada (l'«Alliance»), est une «organisation syndicale» au sens de la LRTFP. L'Alliance a été accréditée en vertu de la LRTFP pour agir à titre d'agent négociateur de nombreuses unités de négociation dans la fonction publique fédérale. Elle est le troisième plus gros syndicat du Canada et représente environ 170 000 fonctionnaires, dont 70 pour cent travaillent à l'extérieur de la région de la capitale nationale. L'Alliance est composée de 18 éléments qui sont, exception faite d'un ou deux éléments, composés majoritairement d'hommes. L'unité de négociation comptant le plus de membres que représente l'Alliance est le groupe Commis aux écritures et aux règlements (le «groupe CR») qui comprend environ 50 000 employés. Ce groupe est composé à 80 pour cent de femmes et inclut des fonctionnaires qui exercent un très large éventail de fonctions.
5. L'Alliance a cité quatre témoins experts au cours de cette audience. Le premier a été Mme Pat Armstrong, que le Tribunal a acceptée à titre d'experte en évaluation d'emploi et en équité salariale. L'Alliance

a également fait témoigner M. Eugene Swimmer, en tant qu'expert en économique du travail et en statistique. Le Tribunal a accepté une employée de l'Alliance, Mme Margaret Jaekl, à titre d'experte en équité salariale et en évaluation d'emplois. Une autre personne, Mme Margaret I. Krachun, qui à l'époque de l'audience était employée par l'Alliance, a été acceptée comme profane ayant une certaine expérience de l'évaluation, qu'elle a acquise en tant que membre d'un des comités d'évaluation.

2

6. Au départ, l'affaire dont le Tribunal est saisi a surgi de plaintes déposées par l'Alliance et l'Institut professionnel de la fonction publique du Canada (l'«Institut») alléguant violation de l'article 11 de la Loi. L'Institut a cité un témoin expert, M. Dan Butler, un négociateur de l'Institut. Le Tribunal l'a accepté à titre d'expert pour qu'il exprime l'opinion de l'Institut sur plusieurs questions devant le Tribunal, principalement la méthodologie de rajustement des salaires.

7. Les plaintes relatives aux droits de la personne que le Tribunal a à trancher maintenant ne concernent que les plaintes de l'Alliance. Le Tribunal n'est plus saisi des plaintes de l'Institut, lesquelles ont été réglées à la suite d'une entente négociée entre l'employeur et l'Institut. Le 31 mai 1995, le Tribunal a rendu une ordonnance sur consentement qui donnait effet à ce règlement.

8. En ce qui concerne l'Alliance, il nous reste deux plaintes à trancher. La première, datée du 19 décembre 1984, allègue que l'employeur a traité de façon discriminatoire les fonctionnaires du groupe CR à prédominance féminine, en contravention des articles 7, 10 et 11 de la Loi. Ce n'est que la partie de cette plainte qui est fondée sur l'article 11 qui a été renvoyée au Tribunal pour décision. La plainte déposée au nom des membres du groupe CR met en cause les droits d'environ 50 000 fonctionnaires appartenant à ce groupe.

9. La deuxième plainte, datée du 16 février 1990, allègue que les résultats obtenus à la suite des travaux du Comité mixte syndical-patronal sur la mise en oeuvre de la parité salariale ont montré que les taux de rémunération versés aux fonctionnaires des groupes professionnels à prédominance féminine Commis aux écritures et aux règlements, Secrétariat, sténographie et dactylographie, Traitement des données, Soutien de l'enseignement, Services hospitaliers et Bibliothéconomie contrevenaient à l'article 11 de la Loi. L'Alliance a déposé cette deuxième plainte auprès de la Commission peu de temps après la dissolution du Comité mixte syndical-patronal sur la mise en oeuvre de la parité salariale (dont nous discuterons en détail plus loin). Cette plainte s'appuie sur les données

relatives à l'évaluation des emplois produites par une étude du Comité mixte, données qui, selon l'Alliance, prouvent que les fonctionnaires faisant partie des groupes précités continuent -- en contravention de l'article 11 de la Loi -- de faire l'objet de discrimination salariale, en dépit des paiements unilatéraux annoncés par l'employeur en janvier 1990.

10. Dès le début, la position privilégiée par l'Alliance a consisté à tenter de résoudre les questions de parité salariale en négociant avec l'employeur à la table de négociation. Ce n'est que lorsque ces mesures n'ont pu mener à un correctif que l'Alliance a recouru au mécanisme de plainte prévu par la Loi.

11. Les plaintes relatives aux droits de la personne de l'Alliance ne sont pas les premières plaintes que l'Alliance présente en vertu de la Loi. Elle a déjà déposé une plainte au nom du groupe Bibliothéconomie (le «groupe LS»), et une autre concernant le groupe Services hospitaliers (le «groupe HS»), déposée au nom des fonctionnaires membres des sous-groupes du groupe Services généraux (le «groupe GS») à majorité féminine.

3

12. Dans chacun de ces cas, une indemnisation pécuniaire sous forme de rajustements salariaux a été versée aux fonctionnaires touchées. La plainte concernant le groupe LS a été réglée avec l'entente que la mesure corrective finale attendrait l'issue de l'étude. En ce qui concerne la plainte visant le groupe HS, qui a fait l'objet d'une ordonnance d'un Tribunal antérieur le 15 juillet 1987, il a formellement été convenu par les parties que la plainte fondée sur l'article 11 se serait réglée qu'après le calcul des écarts salariaux une fois l'étude terminée.

13. Chaque fonctionnaire fédéral occupe un poste qui est classifié selon le système de classification de l'employeur. Ce système comprend 69 groupes professionnels dont chacun comporte sa propre norme de classification («système d'évaluation des emplois»).

14. Dans le système de classification, les postes sont classifiés comme appartenant à des groupes professionnels, à des sous-groupes (s'il y a lieu) et à des niveaux. Les groupes professionnels sont désignés par une abréviation de deux lettres, les sous-groupes, par une abréviation de trois lettres. Un poste est l'unité organisationnelle la plus petite et représente un ensemble unique de tâches et de fonctions exécutées par un individu. L'employeur a autant de postes qu'il a d'employés. Par contre, un emploi dans la fonction publique est un groupe de postes qui ont les mêmes fonctions et responsabilités.

15. Les groupes professionnels sont regroupés en six catégories professionnelles, à savoir : (i) la catégorie scientifique et professionnelle; (ii) la catégorie de l'administration et du service extérieur; (iii) la catégorie technique; (iv) la catégorie du soutien administratif; (v) la catégorie de l'exploitation; et (vi) la catégorie de la direction.

16. En mars 1985, le gouvernement a adopté des mesures proactives en vue de mettre en oeuvre les principes de l'équité salariale dans la fonction publique fédérale. Il a invité les syndicats et l'employeur à participer en tant que partenaires à un comité mixte syndical-patronal, de niveau supérieur, sur la mise en oeuvre de la parité salariale (le «Comité mixte»). On a confié au Comité mixte le mandat de préparer un plan de mise en oeuvre détaillé dans le domaine de la parité salariale. Les syndicats, non seulement l'Alliance mais d'autres syndicats aussi, ont accepté l'invitation du gouvernement. L'Alliance, lorsqu'elle a accepté l'invitation, avait adopté une politique cohérente de soutien au principe de l'égalité de rémunération pour fonctions équivalentes. Lors de l'initiative volontaire, trois plaintes fondées sur l'article 11 étaient en suspens devant la Commission.

17. Le plan d'action convenu par le Comité mixte consistait à mener une étude (l'«étude sur la parité salariale») en application de l'article 11 de la Loi afin de déterminer le degré de discrimination salariale fondée sur le sexe et d'élaborer des correctifs propres à corriger ces disparités à la grandeur du système. (pièce HR-11A, onglet 9, annexe B.) La Commission a été invitée à participer à l'étude en tant qu'observatrice aux réunions du Comité mixte et en vue de fournir à ce dernier les interprétations et conseils dont il pourrait avoir besoin

4

(pièce HR-11A, onglet 7). La Commission a gardé en suspens toutes les plaintes fondées sur l'article 11 qui avaient été déposées avant que le Comité mixte amorce son étude. Elle a accepté que toute nouvelle plainte reçue durant l'étude et qui pourrait être touchée par les travaux du Comité soit également gardée en suspens.

18. Le Comité mixte comptait un nombre égale de représentants de l'intimé et des huit différents syndicats. La tâche première du Comité mixte était de définir les paramètres de l'étude. Une exigence essentielle de son travail était que les représentants des deux parties, patronale et syndicale, s'entendent sur le processus qu'on allait suivre durant l'étude. Ni les syndicats ni l'employeur ne devaient agir indépendamment ou prendre des décisions sans l'approbation de l'autre partie au cours de l'exécution

de l'étude. Le Comité mixte a retenu les services de Willis & Associates, un cabinet de consultants ayant son siège social à Seattle, Washington, pour l'aider dans l'étude. La société Willis & Associates a été fondée et est dirigée par M. Norman Willis.

19. Dès le début de ses travaux, le Comité mixte a clairement fait savoir à M. Willis qu'il n'avait aucun pouvoir décisionnel dans la conduite de l'étude. M. Willis devait assister aux réunions et donner son avis à la demande du Comité mixte.

20. A divers stades de ses travaux, le Comité mixte a établi des sous-comités à qui il a demandé de fournir des avis, d'accomplir certaines tâches et de lui faire des recommandations sur des points particuliers. Les membres du Comité devaient s'entendre sur la formation de tout sous-comité. Chaque sous-comité ainsi formé se composait d'un nombre égal de représentants syndicaux et patronaux.

21. A l'automne de 1987, le Comité mixte a établi le Secrétariat de l'étude sur la parité salariale (le «SEPS») pour effectuer le travail administratif associé à l'étude sur la parité salariale. Le Secrétariat était géré par un représentant du Conseil du Trésor, M. Pierre Collard. Le SEPS avait pour mandat de fournir le soutien administratif nécessaire aux multiples comités d'évaluation participant à l'étude et de coordonner toutes les activités de soutien.

22. Après avoir retenu les services de Willis & Associates, le Comité mixte s'est plus tard entendu sur d'autres questions importantes. Il a convenu d'évaluer les postes de groupes professionnels composés majoritairement d'hommes ou de femmes à l'aide d'un plan d'évaluation commun. Il serait alors possible de comparer les salaires payés aux groupes professionnels à prédominance masculine ou féminine exerçant des fonctions équivalentes. Le Comité a convenu que l'étude serait fondée sur des postes déterminés et qu'elle utiliserait un échantillon représentatif de ceux-ci. Une telle étude signifie que chaque poste choisi pour fins d'évaluation fait l'objet d'une évaluation distincte, par opposition à une étude dans laquelle sont sélectionnés pour les besoins de l'évaluation les postes qui représentent le mieux une classification ou un groupe d'emplois. Le Comité mixte a convenu que seuls les postes de groupes professionnels composés majoritairement d'hommes ou de femmes au sens de l'article 13 de

l'Ordonnance sur la parité salariale (l'«Ordonnance») seraient inclus dans l'échantillon représentatif.

23. En mars 1985, s'appuyant sur l'article 13 de l'Ordonnance qui définit en quoi consiste un groupe composé majoritairement de membres d'un sexe, les parties ont convenu qu'il y avait 9 groupes professionnels à prédominance féminine, 53 à prédominance masculine et 8 au sein desquels il n'y avait aucune prédominance de membres d'un sexe. A titre de clarification, nous reproduisons ci-dessous l'article 13 de l'Ordonnance :

13. Pour l'application de l'article 12, un groupe professionnel est composé majoritairement de membres d'un sexe si, dans l'année précédant la date du dépôt de la plainte, le nombre de membres de ce sexe représentait au moins:

- (a) 70 pour cent du groupe professionnel, dans le cas d'un groupe comptant moins de 100 membres;
- (b) 60 pour cent du groupe professionnel, dans le cas d'un groupe comptant de 100 à 500 membres;
- (c) 55 pour cent du groupe professionnel, dans le cas d'un groupe comptant plus de 500 membres.

24. Les neuf groupes professionnels composés majoritairement de femmes que représentent l'Alliance et l'Institut sont les suivants (avec l'abréviation du groupe) :

- .. Commis aux écritures et aux règlements (CR);
- .. Traitement des données (DA);
- .. Soutien de l'enseignement (EU);
- .. Sciences domestiques (HE);
- .. Services hospitaliers (HS);
- .. Bibliothéconomie (LS);
- .. Sciences infirmières (NU);
- .. Ergothérapie et physiothérapie (OP);
- .. Secrétariat, sténographie et dactylographie (ST).

25. Les postes des groupes professionnels sans prédominance féminine ou masculine et de la catégorie de la Direction ont été exclus de l'étude. L'étude menée par le Comité mixte, si elle portait sur l'ensemble de la fonction publique, ne visait toutefois pas à inclure tous les employés qui fournissaient des services à l'administration fédérale. Elle n'incluait pas notamment les employés des sociétés d'État ni ceux des employeurs distincts. A des fins législatives, les employeurs distincts sont identifiés à la partie 2 de la LRTFP comme suit:

- .. Commission de contrôle de l'énergie atomique;
- .. Conseil consultatif canadien sur le statut de la femme;
- .. Service canadien du renseignement de sécurité;
- .. Centre de sécurité des télécommunications, ministère de la

- Défense nationale;
- .. Conseil économique du Canada;
- .. Conseil de recherches médicales;
- .. Office national du film;

6

- .. Conseil national de recherches Canada;
- .. Conseil de recherches en sciences naturelles et en génie;
- .. Commission d'énergie du Nord canadien;
- .. Administration du pipe-line du Nord;
- .. Bureau du Vérificateur général du Canada;
- .. Commission des relations de travail dans la Fonction publique;
- .. Conseil des sciences du Canada;
- .. Conseil de recherches en sciences humaines du Canada;
- .. Le personnel des fonds non publics des Forces canadiennes.

26. L'échantillon qu'on a finalement prélevé était représentatif des postes des groupes professionnels à prédominance féminine en ce qui concerne le groupe et le niveau, et représentatif des groupes professionnels à prédominance masculine en ce qui concerne le groupe. Environ 2 800 postes des groupes professionnels à majorité féminine et 1 500 postes des groupes professionnels à majorité masculine ont finalement été inclus dans l'échantillon. La taille de l'échantillon et sa composition ont été approuvées par Statistique Canada.

27. Le Comité mixte a accepté d'utiliser le plan d'évaluation des emplois Willis, avec certaines modifications, comme étant l'instrument approprié pour évaluer l'échantillon représentatif des postes. Le Comité mixte a également convenu d'utiliser le questionnaire Willis, avec certaines modifications, pour recueillir des informations sur les postes à évaluer. Un sous-comité du Comité mixte a recommandé une stratégie de communication -- qu'on a acceptée -- en vue d'encourager des titulaires choisis à participer à l'étude et à fournir des renseignements sur leur poste. Les informations sur les postes ont ensuite été recueillies de septembre 1987 à janvier 1989.

28. Sur l'avis de M. Willis et à titre d'étape initiale dans le processus d'évaluation, le Comité mixte a établi un Comité directeur de l'évaluation (le «Comité directeur») à qui il a demandé d'évaluer 503 questionnaires sur des postes qui devaient servir de postes-repères et de cadre de référence pour toutes les évaluations qu'effectueraient subséquemment d'autres comités d'évaluation. Le Comité directeur a commencé son important travail en septembre 1987 et l'a achevé en juillet

1988. A cette date, le Comité directeur avait complété 501 évaluations de postes-repères.

29. Après que le Comité directeur eut terminé ses évaluations, quatorze comités d'évaluation (les «comités d'évaluation multiples») ont effectué les évaluations restantes. Les cinq premiers comités d'évaluation multiples ont amorcé leurs travaux en septembre 1988. En avril 1989, ils avaient évalué environ 1 283 postes. En avril 1989, le nombre de ces comités a été porté de cinq à neuf. Les neuf comités comprenaient, en plus de nouveaux membres, certains membres des cinq premiers comités multiples. Ils ont évalué environ 1 400 postes entre avril 1989 et septembre 1989.

30. En mai 1989, le Comité mixte a décidé, à cause de la lenteur avec laquelle progressait l'évaluation des questionnaires, que la taille de l'échantillon devrait être réduite d'environ 882 postes. Il a ensuite

7

convenu de réduire l'échantillon initial de 4 300 postes à environ 3 279 postes. Le Bureau du statisticien en chef du Canada a été avisé de la nature et des motifs de la réduction, qu'il a approuvée. Au bout du compte, le Comité directeur et les 14 comités d'évaluation multiples ont évalué 3 185 postes puisés dans l'échantillon réduit.

31. Les représentants de la Commission ont fait office d'observateurs tout au long des évaluations effectuées par le Comité directeur et les comités d'évaluation multiples. Ils ont assisté aux réunions du Comité mixte et à celles des présidents des comités d'évaluation multiples.

32. Dans l'ensemble, les travaux du Comité mixte ont été marqués de nombreuses lacunes, en grande partie à cause du mode de fonctionnement du Comité. Selon M. Willis, le Comité mixte avait été «mal formé». Au lieu de travailler en équipe, il a adopté le mode de négociation: les syndicats d'un côté, l'employeur de l'autre. Chaque partie exprimait son avis d'une seule voix, d'ajouter M. Willis. Comme le Conseil du Trésor représentait une position unique, cela a obligé les syndicats à se regrouper afin de réagir d'une seule voix. Au lieu d'un comité mixte patronal-syndical travaillant en équipe, le processus a plutôt pris l'allure de négociations patronales-syndicales.

33. Par conséquent, de nombreuses décisions n'ont pu être prises qu'après énormément de temps et d'efforts, et l'entente fut souvent difficilement acquise dans un climat de méfiance. Par exemple, après la première réunion du Comité mixte qui a eu lieu le 16 septembre 1985, il a fallu un an aux parties pour se mettre d'accord sur les attributions et le

plan d'action de l'étude sur la parité salariale -- elles n'y sont parvenues que le 22 septembre 1986.

34. Le temps que prenait l'exécution de l'étude a incité le président de la Commission, à différentes occasions, à tenter de persuader le président du Conseil du Trésor de résoudre les questions sur lesquelles le Comité mixte n'arrivait pas à se mettre d'accord.

35. Un autre problème qui a marqué les travaux du Comité mixte fut l'incapacité des parties patronale et syndicale à régler de façon définitive certains aspects de l'étude. Par exemple, une fois que le Comité directeur eut terminé l'évaluation des postes-repères, le Conseil du Trésor n'a pas entériné sans réserve ces évaluations. Bien qu'il acceptât qu'on effectue le reste des évaluations, le Conseil du Trésor a continué d'entretenir des doutes et a fait savoir son intention d'étudier lui-même la fiabilité des évaluations des postes-repères réalisées par le Comité directeur.

36. Des problèmes ont également surgi durant les évaluations effectuées par les comités d'évaluation multiples. M. Willis a recommandé la dissolution de l'un de ces cinq comités qui avaient été formés au départ. Le Comité mixte a rejeté cette recommandation et n'arrivait pas à se mettre d'accord sur un règlement. De plus, certains des comités d'évaluation multiples ont contesté les évaluations des postes-repères faites par le Comité directeur. Le Comité mixte a établi un sous-comité du Comité directeur (le «mini-comité directeur») pour examiner et discuter ces

8

contestations. Ce dernier n'a pas réussi à en arriver à un consensus de sorte que, au bout du compte, la question n'a jamais été entièrement résolue.

37. L'étude sur la parité salariale devait inclure les quatre phases suivantes:

#### Phase I

Entente concernant le plan d'évaluation commun à utiliser pour déterminer la valeur relative des postes et l'évaluation des postes-repères.

#### Phase II

Entente concernant la méthode statistique à adopter pour l'échantillonnage de postes.

### Phase III

Échantillonnage et évaluation des postes réels à l'aide du plan d'évaluation convenu et des postes-repères.

### Phase IV

Déterminer l'ampleur de la disparité salariale et recommander les mesures correctives à prendre pour éliminer les aspects discriminatoires du système de classification, en conformité avec l'article 11 de la Loi canadienne des droits de la personne. (pièce HR-11A, onglet 9)

38. Tout au long de l'étude sur la parité salariale les tensions entre les parties patronale et syndicale ont persisté et se sont intensifiées. Les parties n'arrivaient pas à s'entendre sur la communication des cotes d'évaluation. Le Comité mixte a convenu de divulguer les données après que les deux tiers des évaluations auraient été achevées. Selon M. Willis, après la communication des cotes du Comité directeur le 13 juillet 1988, les rapports au sein du Comité mixte ont commencé à se détériorer. Il était devenu évident aux yeux de M. Willis que le climat au sein du Comité mixte avait changé. Lorsque les résultats du Comité directeur ont été communiqués aux parties, le système de classification de l'intimé est devenu un sujet de controverse pour ce dernier. Une lettre qu'il a reçue le 18 août 1988 de la coprésidente patronale et dans laquelle ce dernier indiquait que les parties avaient des vues divergentes sur l'objet de l'étude a laissé M. Willis inquiet et perplexe.

39. Au cours des derniers mois de l'étude sur la parité salariale, un litige a surgi entre les parties patronale et syndicale au sujet d'un rapport publié par Willis & Associates concernant la réévaluation qu'un consultant de Willis avait effectuée de 222 évaluations des comités d'évaluation multiples. Cette question n'a jamais été résolue par le

Comité mixte et a finalement entraîné l'échec définitif de l'étude sur la parité salariale.

40. Les parties souhaitaient finir par s'entendre sur une recommandation conjointe qu'elles auraient faite au président du Conseil du

Trésor en vue de la mise en oeuvre de la parité salariale. Elles n'ont jamais réussi à atteindre la phase 4 de l'étude. Après environ quatre ans, en décembre 1989, la partie syndicale s'est temporairement retirée de l'étude. En janvier 1990, le syndicat le plus gros qui prenait part à l'étude, l'Alliance, s'est retiré en permanence de celle-ci.

41. Au début de 1990, le gouvernement du Canada a décidé de mettre en oeuvre unilatéralement des mesures immédiates en vue d'appliquer le principe de l'égalité de rémunération pour fonctions équivalentes dans les groupes professionnels de la fonction publique fédérale à prédominance féminine. Les mesures adoptées par le gouvernement étaient fondées sur les résultats des évaluations issus de l'étude sur la parité salariale avec certains rajustements visant à corriger la partialité fondée sur le sexe qu'avait mise en lumière le rapport controversé de Willis & Associates sur les 222 réévaluations. Ces mesures ont été appelées les «rajustements en cas de disparité salariale» ou «rajustements paritaires» de la fonction publique. Les rajustements paritaires ont été appliqués à trois groupes professionnels composés majoritairement de femmes, à savoir les groupes CR, NU et ST.

42. L'employeur n'a consulté ni la Commission ni l'Alliance, ni aucun autre des syndicats participants, avant de faire ces rajustements de son propre chef. Les parties ont été informées pour la première fois de la décision de l'employeur lorsque le président du Conseil du Trésor en a fait l'annonce le 26 janvier 1990. Les rajustements se traduisaient par des paiements d'environ 317 millions de dollars en traitements rétroactifs remontant au 1er avril 1985 et par des paiements annuels de 76 millions de dollars en rajustements permanents. Le gouvernement a fait remonter au 31 mars 1985 la période de rétroactivité visée, soit le mois au cours duquel le président du Conseil du Trésor a pour la première fois annoncé l'établissement du Comité mixte syndical-patronal à qui était confié le mandat d'étudier comment on éliminerait la discrimination salariale fondée sur le sexe dans la fonction publique fédérale.

43. Après l'échec de l'étude sur la parité salariale, la Commission et l'Alliance ont clairement fait savoir à l'employeur qu'elles entendaient présenter les données produites par l'étude en tant que preuves devant un tribunal des droits de la personne.

44. L'enquête officielle concernant les plaintes fondées sur l'article 11 déposées auprès de la Commission a commencé après l'annonce des rajustements paritaires. Dans le cadre de son enquête, la Commission -- dans le but de respecter pleinement la Loi et l'Ordonnance sur la parité salariale -- a examiné ces rajustements. A l'issue de l'enquête officielle qu'elle a menée, et qui a duré six mois, la Commission a décidé de renvoyer

les plaintes fondées sur l'article 11 à un tribunal. Cette décision a été prise le 16 octobre 1990.

45. Au cours de l'audience qui s'ensuivit, lorsque la Commission a voulu produire en preuve les données de l'étude menée par le Comité mixte, l'employeur s'y est opposé en affirmant que les données étaient inadmissibles puisqu'elles avaient été créées dans le but de résoudre ou d'éviter un litige et qu'on devait donc les traiter comme éléments de preuve privilégiés. Le Tribunal a mené sur cette question un voir-dire qui l'a amené à rejeter l'objection de l'employeur dans une décision datée du 21 août 1992. (Voir la décision du voir-dire pour plus de détails.)

46. L'intimé soutient que les données relatives à l'évaluation des postes produites au cours de l'étude ne sont pas assez fiables pour que le Tribunal puisse s'y appuyer pour trancher les plaintes dont il est saisi. L'employeur n'est pas satisfait de la fiabilité des résultats d'évaluation. Les rajustements paritaires effectués par l'employeur révèlent la mesure dans laquelle l'employeur est prêt à se fier aux résultats des évaluations. La Commission et l'Alliance cherchent à utiliser les données des évaluations comme base pour déterminer la disparité salariale et les rajustements salariaux en vertu de l'article 11 de la Loi.

## II. QUESTION A TRANCHER

47. A la suite d'une initiative proactive prise par l'employeur, la plaignante, ainsi que 13 autres syndicats de la fonction publique, et l'intimé ont entrepris une étude sur la parité salariale sous l'égide du Comité mixte patronal-syndical sur la mise en oeuvre de la parité salariale.

48. L'étude sur la parité salariale a été amorcée en 1985 et s'est prolongée jusqu'en janvier 1990, lorsqu'elle a été arrêtée prématurément d'abord par la plaignante et ensuite par l'intimé. Sur ce laps de temps, la plaignante et l'intimé ont produit des résultats d'évaluation des emplois.

49. Avant le début de l'étude sur la parité salariale, la plaignante avait déposé auprès de la Commission, en vertu de l'article 11 de la Loi, une plainte de discrimination salariale contre l'intimé. Après l'arrêt de l'étude, la plaignante a présenté une deuxième et nouvelle plainte contre l'intimé.

50. La Commission et la plaignante ont l'intention d'utiliser les résultats de l'évaluation des emplois découlant de l'étude sur la parité salariale comme preuve de la valeur du travail accompli par les fonctionnaires, hommes et femmes, dont les emplois sont visés par ces plaintes. La Commission et la plaignante comptent par ailleurs invoquer les résultats de l'évaluation des emplois comme preuve d'une disparité salariale qui, selon les plaintes, contrevient à l'article 11 de la Loi.

51. L'intimé soutient que les résultats de l'évaluation des emplois ne sont pas fiables pour les besoins de la décision. Plus spécifiquement, il maintient que ces résultats sont biaisés, dans la mesure où les questionnaires concernant les postes à prédominance masculine et les questionnaires portant sur les postes à prédominance féminine utilisés pour

11

produire les résultats ont été traités différemment par les responsables des évaluations.

52. Par conséquent, la question à trancher est de savoir si les résultats de l'étude sur la parité salariale sont fiables pour les besoins des plaintes fondées sur l'article 11 renvoyées au Tribunal pour décision.

### III. LES TEXTES LÉGISLATIFS

53. Les plaintes dont nous sommes saisis allèguent la discrimination fondée sur le sexe en contravention de l'article 11 de la Loi. L'article 11 est ainsi libellé :

11. (1) Constitue un acte discriminatoire le fait pour l'employeur d'instaurer ou de pratiquer la disparité salariale entre les hommes et les femmes qui exécutent, dans le même établissement, des fonctions équivalentes.

(2) Les critères permettant d'établir l'équivalence des fonctions exécutées par des salariés dans le même établissement est le dosage des qualifications, d'efforts et de responsabilités nécessaire pour leur exécution, compte tenu des conditions de travail.

[...]

(5) Des considérations fondées sur le sexe ne sauraient motiver la disparité salariale.

54. Les dispositions de l'article 11 de la Loi concernant l'égalité de rémunération pour fonctions équivalentes ont fait l'objet d'une décision de la Cour suprême du Canada dans l'affaire *Syndicat des employés de production du Québec et de l'Acadie c. Canada* (Commission canadienne des droits de la personne), [1989] 2 R.C.S. 879 (C.S.C.). Dans cette affaire, la Cour avait à déterminer si la décision de la Commission canadienne des droits de la personne de rejeter une plainte fondée sur l'alinéa 36(3)b) de la Loi était «légalement soumise» à un processus judiciaire ou quasi judiciaire et si, par conséquent, elle pouvait faire l'objet d'un examen par la Cour d'appel fédérale en vertu de l'article 28 de la Loi sur la Cour fédérale. La majorité de la Cour a conclu que la décision de la Commission ne donnait pas lieu à examen par la Cour d'appel fédérale en vertu de l'article 28 de la Loi sur la Cour fédérale et que, par conséquent, cette décision n'avait pas à être soumise à un processus judiciaire ou quasi judiciaire.

55. Bien que l'interprétation de l'article 11 de la Loi ne fit pas partie intégrante de la décision de la majorité, le juge Sopinka, au nom de la majorité, affirme ce qui suit à la p. 903 :

L'article 11 vise à interdire à un employeur de pratiquer la discrimination «entre les hommes et les femmes» qui exercent des fonctions équivalentes. Il n'a pas pour but de garantir aux employés individuels un salaire égal pour un travail équivalent peu importe leur sexe.

56. A notre avis, selon l'interprétation du juge Sopinka le libellé de l'article 11 interdit à l'employeur de faire des distinctions fondées sur le «sexe» lorsqu'il détermine les salaires qu'il paiera aux hommes et aux femmes de son entreprise accomplissant des fonctions équivalentes. Le paragraphe 11(5) renforce cette disposition en disant clairement que des considérations fondées sur le sexe ne sauraient motiver la disparité salariale. D'autres articles de la Loi portent aussi sur les interdictions fondées sur le sexe. Au paragraphe 3(1) de la Loi, le «sexe» figure parmi les motifs de distinction illicite. L'article 7 déclare qu'il est discriminatoire de refuser d'employer un individu ou de le défavoriser en cours d'emploi pour un motif illicite, par exemple, le sexe. L'article 10 porte qu'il est discriminatoire -- si l'acte est fondé sur un motif de distinction illicite, par exemple, le sexe, et s'il est susceptible d'annihiler les chances d'emploi ou d'avancement d'un individu ou d'une catégorie d'individus -- de fixer ou d'appliquer des lignes de conduite, ou de conclure des ententes touchant le recrutement, les mises en rapport,

l'engagement, les promotions, la formation, l'apprentissage, les mutations ou tout autre aspect d'un emploi présent ou éventuel.

57. La pratique discriminatoire alléguée dans les plaintes dont le Tribunal est saisi est que l'employeur maintient une disparité salariale entre les hommes et les femmes travaillant dans la même entreprise et exécutant des fonctions équivalentes, en contravention de l'article 11. La Loi, si elle interdit les disparités salariales discriminatoires, prévoit cependant certaines exceptions au paragraphe 11(4), qui dit ce qui suit :

11. (4) Ne constitue pas un acte discriminatoire au sens du paragraphe (1) la disparité salariale entre hommes et femmes fondée sur un facteur reconnu comme raisonnable par une ordonnance de la Commission canadienne des droits de la personne en vertu du paragraphe 27(2).

58. La courte histoire législative de l'article 11 montre que le gouvernement du Canada a déclaré en 1976 qu'il allait présenter un projet de loi sur les droits de la personne. Ce projet de loi aurait principalement pour effet d'interdire la discrimination fondée sur la race, la couleur, l'origine nationale ou ethnique, la religion, l'âge, le sexe, l'état matrimonial ou la déficience. Il établirait notamment le principe de l'égalité de rémunération entre les hommes et les femmes lorsqu'ils exercent des fonctions équivalentes. (pièce PIPSC-82).

59. Dans les notes documentaires que le ministre de la Justice de l'époque avait jointes au projet de loi sur les droits de la personne, il est expliqué que le projet de loi inclurait, parmi les motifs de distinction illicite, les pratiques discriminatoires telles que la disparité salariale fondée sur le sexe entre salariés exécutant des fonctions équivalentes. On peut notamment lire ce qui suit à la p. 4 :

[TRADUCTION] Cette disposition vise principalement à éliminer les «ghettos d'emploi» chez les femmes; elle permettrait aux travailleuses accomplissant un genre de travail, de secrétariat par exemple, de voir leur rémunération établie en fonction non seulement de celle des autres secrétaires, mais aussi de celle qui est accordée pour des fonctions équivalentes au sein de l'entreprise.

(pièce PIPSC-82, p.3)

60. En 1977, le gouvernement du Canada a adopté la Loi. L'objet de l'article 11 de la Loi est de faire en sorte que les hommes et les femmes

qui exercent des fonctions de valeur égale puissent toucher une rémunération égale. L'article 11 est entré en vigueur le 1er mars 1978. Le paragraphe 27(2) de la Loi autorise la Commission canadienne des droits de la personne à prendre une ordonnance précisant les modalités d'application de la Loi. Depuis la promulgation de la Loi en 1978, la Commission a deux fois pris une Ordonnance. La première ordonnance prise en application de la Loi visait à faciliter l'interprétation de l'article 11 de la Loi et fut promulguée le 18 septembre 1978. Elle a été révoquée par l'Ordonnance de 1986, datée du 18 novembre 1986 et publiée dans la Gazette du Canada en décembre 1986.

61. L'Ordonnance de 1986 décrit les modalités d'application de la Loi et les facteurs reconnus raisonnables pour justifier la disparité salariale entre les hommes et les femmes qui exécutent dans le même établissement des fonctions équivalentes. L'Ordonnance de 1986 prescrit dix facteurs justifiant un écart salarial entre les hommes et les femmes qui exercent des fonctions de valeur égale. Aucune de ces exceptions n'entre en ligne de compte dans les présentes plaintes.

62. L'opinion dissidente dans *Syndicat*, supra, est utile parce que le juge y aborde certains des éléments préalables à l'établissement du bien-fondé d'une plainte fondée sur l'article 11 de la Loi. L'opinion dissidente est celle du juge L'Heureux-Dubé qui, dans ses observations, se reporte à des arrêts antérieurs de la Cour suprême du Canada, à savoir *Robichaud c. Canada (Conseil du Trésor)*, [1987] 2 R.C.S. 84 et *Compagnie des chemins de fer nationaux du Canada c. Canada (Commission canadienne des droits de la personne)*, [1987] 1 R.C.S. 1114 (sub nomine : *Action Travail des Femmes*). Dans ces affaires, la Cour avait à examiner des plaintes présentées en vertu des articles 7 et 10 de la Loi respectivement. Dans les deux cas, la Cour a clairement affirmé que l'intention ne constituait pas une condition préalable à la prononciation d'un jugement de discrimination préjudiciable en vertu de la Loi. Le juge L'Heureux-Dubé signale que la portée de la protection accordée par l'article 11 diffère de celle qui est prévue aux articles 7 et 10; elle fait les observations suivantes à la p. 925 :

L'intention n'étant pas un élément essentiel de la discrimination préjudiciable, il est loisible au plaignant d'étayer son argumentation fondée sur les art. 7 et 10 en présentant le même type de preuve que celle qui a été produite en l'espèce. Ainsi, le recours à des données statistiques démontrant l'existence d'une ségrégation professionnelle est un outil des plus précieux dans la découverte d'une discrimination préjudiciable. L'article 11, cependant, diffère des art. 7 et 10. La portée de sa protection est en effet délimitée par le concept d'«équivalence»: cette disposition n'empêche pas un employeur de rémunérer différemment des

emplois non équivalents. Dans le cadre particulier de cette disposition, la discrimination salariale présuppose l'équivalence, dans le même établissement, des fonctions exécutées par des femmes et par des hommes. En conséquence, pour que soit accueillie

14

une plainte fondée sur l'art. 11, il faut établir l'équivalence des fonctions à l'égard desquelles est alléguée une situation de disparité salariale discriminatoire.

63. Le juge L'Heureux-Dubé estime qu'un plaignant peut étayer son argumentation sur les articles 7 et 10 sans inclure dans sa preuve l'élément d'intention. Pour reprendre ses mots, le recours à des données statistiques est «un outil des plus précieux dans la découverte d'une discrimination préjudiciable».

64. Bien que la formulation du principe de «l'égalité de rémunération pour fonctions équivalentes» soit simple et directe, de souligner le juge L'Heureux-Dubé, son application aux termes de l'article 11 de la Loi soulève des difficultés considérables. Ce n'est qu'en apparence que le concept est simple, maintient-elle. Une des difficultés réside dans la notion d'égalité qui, à son avis, ne devrait pas être interprétée de façon technique ou restrictive. A propos du concept d'égalité elle fait les observations suivantes aux pp. 926-27 :

L'interdiction dont est l'objet la discrimination salariale s'inscrit dans un régime législatif plus large visant à éliminer tout acte discriminatoire et à promouvoir l'égalité dans l'emploi. C'est dans ce cadre élargi que l'art. 11 s'attaque au problème de la sous-évaluation du travail exécuté par des femmes. Puisque cet objectif va au-delà de l'interdiction évidente de payer un salaire inférieur pour des fonctions strictement identiques, la notion d'égalité contenue à l'art. 11 ne devrait pas recevoir une interprétation formaliste ou restrictive.

65. Une autre difficulté, selon le juge L'Heureux-Dubé, persiste dans le concept de valeur. A la p. 928 elle affirme :

Les notions de «qualifications», d'«effort», de «responsabilités» et de «conditions de travail» que l'on retrouve dans la Loi, ainsi que dans les Ordonnances sur l'égalité de rémunération qui l'accompagnent, relèvent du langage technique. Traditionnellement, on les mesure à l'aide de plans industriels d'évaluation des emplois.

66. Le paragraphe 11(2) définit, en termes généraux, la manière dont on doit mesurer la valeur du travail et établit quatre critères, à savoir les qualifications, les efforts, les responsabilités et les conditions de travail. Ces critères sont définis plus en détail à l'article 3 de l'Ordonnance, le document d'interprétation de l'article 11.

67. Le juge L'Heureux-Dubé fait remarquer que c'est plus qu'une coïncidence si on trouve les même quatre termes dans la contrepartie américaine de la LCDP, et que ces termes indiquent qu'il est possible d'utiliser des plans d'évaluation des emplois pour déterminer si on a affaire à des emplois équivalents au sens de l'article 11. Toutefois, elle estime que le recours à un plan d'évaluation des emplois n'est pas nécessairement le seul moyen d'appliquer les dispositions de l'article 11. Selon la Commission, dont le point de vue a été exprimé par le témoin M. Durber, dans une plainte fondée sur l'article 11 il est possible d'établir

15

l'équivalence des fonctions au moyen d'un plan d'évaluation des emplois, mais on peut aussi le faire par des méthodes moins formelles.

68. Le Tribunal a entendu le témoignage d'experts venus dire que l'objet d'un plan d'évaluation des emplois consistait, eu égard à une plainte fondée sur l'article 11, à déterminer la valeur relative des emplois au sein d'une organisation. Il s'agit d'un processus systématique qui définit et établit d'abord les facteurs qui se rapportent aux quatre critères énoncés au paragraphe 11(2) de la Loi. Les facteurs sont pondérés les uns en fonction des autres afin d'en préciser l'importance relative. Avant de pouvoir élaborer une hiérarchie, il faut passer par diverses étapes, à savoir : recueillir de l'information sur les emplois, définir les emplois qui feront l'objet de l'évaluation, évaluer chaque emploi et attribuer des points à chaque critère d'évaluation.

69. Le juge L'Heureux-Dubé commente ainsi (à la p. 931) le recours aux plans d'évaluation des emplois et le nombre d'étapes en cause :

Les étapes d'un plan d'évaluation d'emploi comportent toutes une certaine dose de subjectivité. En effet, les croyances sociales qui ont abouti à une sous-évaluation du travail des femmes peuvent être une source de préjugés dans l'élaboration et l'application de ces méthodes. Par exemple, les renseignements que fournissent eux-mêmes les employés au sujet de la nature de leur emploi peuvent présenter certaines caractéristiques qui, en raison de valeurs sous-jacentes, pourraient être négligées lors de l'évaluation. Il peut ainsi y avoir confusion entre les caractéristiques pertinentes aux fins de

l'évaluation et certaines notions stéréotypées de ce que l'on perçoit comme des traits inhérents à la femme.

70. Les spécialistes de la parité salariale qui ont témoigné à l'audience ont fait des observations semblables. Bien qu'on puisse exercer un certain contrôle sur les procédures d'évaluation des emplois, jusqu'à un certain point il s'agit toujours d'un processus foncièrement subjectif. La valeur attribuée à chaque emploi est matière d'opinion et varie selon le jugement des évaluateurs. Selon M. Willis, l'expert en parité salariale et le consultant auprès du Comité mixte syndical-patronal sur la mise en oeuvre de la parité salariale, un tel processus peut comporter des erreurs de jugement aléatoires et systématiques.

71. Dans son témoignage, M. Willis a expliqué qu'il était normal que des erreurs aléatoires se glissent dans une entreprise de la taille de l'étude menée par le Comité mixte, erreurs qui peuvent résulter d'un manque d'information sur les emplois, d'hypothèses concernant des aspects particuliers des emplois, du manque d'uniformité dans l'application du plan Willis (le plan d'évaluation des emplois), ou simplement d'interprétations divergentes des données sur les emplois. M. Willis a fait remarquer que si les écarts aléatoires étaient normaux et avaient tendance à s'annuler les uns les autres, les divergences structurelles ne sont pas du domaine du prévisible et ne s'annulent pas les unes les autres. Ces divergences ou erreurs systématiques de jugement, selon M. Willis, sont la preuve d'un parti pris de la part des évaluateurs et devraient être évitées.

16

72. Mme Weiner, une des nombreuses spécialistes de la parité salariale qui ont témoigné devant le Tribunal, a qualifié de discrimination systémique la discrimination salariale dont il est question à l'article 11 de la Loi. Elle décrit dans les termes suivants (volume 6, p. 875) l'aspect involontaire de la discrimination systémique :

[TRADUCTION] La discrimination systémique est involontaire, impersonnelle et intégrée aux systèmes permanents qu'on dénomme souvent des systèmes «exempts de discrimination» puisqu'ils n'ont jamais été conçus pour avoir un effet discriminatoire.

Elle poursuit ainsi à la p. 877 du volume 6 :

[TRADUCTION] La discrimination systémique s'exerce à l'intérieur de systèmes. Elle se manifeste sans cesse dans les manuels de politiques et puisque personne n'a conçu ces systèmes pour qu'ils

aient un effet discriminatoire, il est beaucoup plus difficile de relever cette discrimination.

73. Selon Mme Weiner, cette discrimination émane des pratiques et processus d'un employeur concernant la rémunération plutôt que d'actions individuelles.

74. La décision que la Cour suprême du Canada a rendue dans CN, supra, éclaire l'interprétation de la discrimination systémique. Dans cet arrêt, la Cour a maintenu une ordonnance d'un tribunal canadien des droits de la personne imposant à la Compagnie des chemins de fer nationaux du Canada un programme spécial d'équité en matière d'emploi. En maintenant l'ordonnance remédiate, le juge en chef Dickson (ce qu'il était à l'époque) a souligné à la p. 1134 l'esprit dans lequel on devait interpréter les codes et lois régissant les droits de la personne :

La législation sur les droits de la personne vise notamment à favoriser l'essor des droits individuels d'importance vitale, lesquels sont susceptibles d'être mis à exécution, en dernière analyse, devant une cour de justice. Je reconnais qu'en interprétant la Loi, les termes qu'elle utilise doivent recevoir leur sens ordinaire, mais il est tout aussi important de reconnaître et de donner effet pleinement aux droits qui y sont énoncés. On ne devrait pas chercher par toutes sortes de façons à les minimiser ou à diminuer leur effet.

75. Le juge en chef Dickson s'est penché sur la raison d'être et les objectifs de la législation sur les droits de la personne et sur l'attitude générale qui doit animer la Cour dans l'interprétation qu'elle fait de ces textes de loi, à savoir fournir une interprétation propre à favoriser les grands objectifs de la législation. Il s'est reporté à l'arrêt rendu par la Cour suprême dans *Commission ontarienne des droits de la personne c. Simpsons-Sears Ltd.*, [1985] 2 R.C.S. 536, qui a reconnu que la législation sur les droits de la personne vise non seulement la discrimination intentionnelle, mais aussi la discrimination involontaire, et qu'elle interdit la «discrimination par suite d'un effet préjudiciable».

76. La Cour suprême du Canada, dans CN, supra, a reconnu qu'il fallait distinguer la discrimination systémique vue dans le contexte de l'équité en matière d'emploi de la question de l'égalité de rémunération pour fonctions équivalentes abordée par Mme Weiner dans ses observations sur l'article 11 de la Loi. La Cour suprême reconnaît que le paragraphe 15(1) et par extension l'alinéa 41(2)a) de la Loi canadienne sur les droits

de la personne de 1976-1977 dans sa forme modifiée en 1985 avaient pour objet de résoudre le problème de la discrimination systémique. Le juge en chef Dickson, à la p. 1139, a décrit ainsi la discrimination systémique :

En d'autres termes, la discrimination systémique en matière d'emploi, c'est la discrimination qui résulte simplement de l'application des méthodes établies de recrutement, d'embauche et de promotion, dont ni l'une ni l'autre n'a été nécessairement conçue pour promouvoir la discrimination. La discrimination est alors renforcée par l'exclusion même du groupe désavantagé, du fait que l'exclusion favorise la conviction, tant à l'intérieur qu'à l'extérieur du groupe, qu'elle résulte de forces «naturelles», par exemple que les femmes «ne peuvent tout simplement pas faire le travail» [...]. Pour combattre la discrimination systémique, il est essentiel de créer un climat dans lequel tant les pratiques que les attitudes négatives peuvent être contestées et découragées. Le tribunal a tenté d'atteindre cet objectif par son ordonnance de «mesures spéciales temporaires».

77. Dans sa décision, le juge en chef Dickson a souligné que l'ordonnance du tribunal, que la Cour était appelée à examiner alors, avait pour but de mettre en oeuvre un programme d'équité en matière d'emploi qui n'était pas simplement compensatoire mais aussi prospectif, en ce sens qu'il visait à produire des avantages destinés à «améliorer la situation professionnelle du groupe à l'avenir». De plus, le juge en chef a soutenu qu'un tel programme était conçu pour rompre le cycle vicieux de la discrimination systémique que subissaient les femmes sur le plan de l'emploi. A son avis, l'objet de la législation, l'alinéa 41(2)a) notamment, était de chercher à éliminer les barrières insidieuses qui viendraient se dresser devant les futurs postulants d'emplois, en l'occurrence les femmes, obstacles qui, par le passé, ont pris la forme de pratiques d'emploi injustes que leurs prédécesseurs ont dû subir comme groupe. En revanche, la législation ne vise pas tant à indemniser les victimes passées de discrimination ni même d'ouvrir de nouveaux horizons à des individus particuliers qui en auraient été privés par le passé.

78. Le juge en chef a conclu que l'objectif recherché n'était pas d'indemniser les victimes du passé, voire de fournir de nouvelles possibilités à des individus particuliers à qui on aurait refusé injustement des emplois ou des promotions par le passé, mais de tâcher de faire en sorte qu'à l'avenir les postulants à un emploi faisant partie du groupe touché n'aient pas à faire face aux mêmes barrières insidieuses qui ont arrêté leurs prédécesseurs.

79. Dans cette affaire, le juge en chef était du même avis que le juge McGuigan, le juge dissident de la Cour d'appel fédérale, qui avait

conclu que l'alinéa 41(2)a de la Loi (maintenant l'alinéa 53(2)a) était conçu pour permettre aux tribunaux des droits de la personne d'empêcher

qu'à l'avenir des groupes protégés identifiables ne soient victimes de discrimination en matière d'emploi. Par ailleurs, de l'avis du juge en chef on ne saurait -- dans un programme d'équité en matière d'emploi -- dissocier radicalement les notions de «réparation» et de «prévention». De plus, le terme «prévention» a une acception large, a-t-il dit, et il est souvent nécessaire de se reporter aux pratiques passées de discrimination afin d'élaborer des stratégies pour l'avenir.

80. A notre avis, l'article 11 ne reconnaît pas spécifiquement le phénomène que nous qualifions de «discrimination systémique» et il ne constitue pas un instrument bien conçu pour rompre le cycle de la discrimination. Il faut situer dans leur contexte les observations que le juge en chef Dickson a faites dans l'affaire CN, supra. Dans cette affaire, un Tribunal avait rendu une ordonnance en vertu de l'alinéa 41(2)a, maintenant l'alinéa 53(3)a), obligeant la Compagnie des chemins de fer nationaux à adopter un programme spécial d'équité en matière d'emploi à l'égard du groupe touché de femmes qui voulaient occuper des emplois manuels, ordonnance qui était en appel. La plainte invoquait des pratiques d'emploi discriminatoires et fut tranchée en 1986.

81. La description de la discrimination systémique que le juge en chef Dickson propose dans l'affaire CN, supra, constitue, à notre avis, le genre de discrimination involontaire que l'article 11 a été conçu pour éliminer.

82. Selon l'opinion des experts, la discrimination systémique n'a pas d'origine ni d'objet précis : elle se développe simplement avec le temps. Il s'agit d'un phénomène tenant de l'attitude qui tend à sous-estimer le travail des femmes et qui établit donc des distinctions fondées sur le sexe au détriment d'un individu ou groupe. La recherche a montré que le groupe le plus touché par ce genre de discrimination était les femmes et que leurs salaires, par rapport à ceux des hommes, étaient inférieurs. Ce qui est à l'origine de ce genre de discrimination, ce sont les attitudes, les croyances et les mentalités concernant le travail traditionnellement accompli par les hommes et le travail traditionnellement accompli par les femmes.

83. L'avocate de la Commission soutient que l'article 11 fait partie d'un régime législatif qui interdit la discrimination systémique fondée sur le sexe et la disparité salariale entre les groupes composés

majoritairement d'hommes et majoritairement de femmes accomplissant des fonctions équivalentes. Toujours selon elle, l'article 11 vise à remédier à la sous-évaluation historique du travail des femmes et à éliminer les situations de disparité salariale discriminatoire fondée sur le sexe. L'avocate maintient qu'il y a preuve de ce genre de discrimination s'il existe un écart entre les salaires versés aux groupes professionnels composés majoritairement d'hommes et ceux composés majoritairement de femmes exécutant un travail équivalent. (Volume 218, p. 28424).

84. Il importe à ce stade-ci de comprendre ce qu'on entend par «disparité salariale» à l'article 11 de la Loi. Le Tribunal a eu le bénéfice d'entendre le témoignage expert de Mme Armstrong, spécialiste de l'évaluation d'emploi et de l'équité salariale, qui a décrit la disparité

19

générale qui existe actuellement entre les taux de rémunération versés aux femmes par rapport à ceux qu'on paie aux hommes. Selon Mme Armstrong, on ne saurait comprendre la disparité salariale si on ne saisit pas les facteurs sous-jacents qui ont pu y contribuer. Il se peut fort bien que certains facteurs légitimes et inchangeables soient à l'origine, dans une certaine mesure, de cet écart.

85. Un écart salarial n'est pas quelque chose de clairement délimité. Le Tribunal reconnaît que les disparités salariales entre les hommes et les femmes peuvent être fonction des exigences professionnelles qui font que, pour l'employeur, certains emplois ont une valeur intrinsèque plus grande que d'autres. De tels écarts contrastent avec les disparités qui reposent entièrement sur le sexe et ce sont ces dernières, estime le Tribunal, que l'article 11 vise à éliminer.

86. L'article 11 incorpore dans son libellé le concept d'égalité de rémunération pour fonctions équivalentes. Selon Mme Weiner, deux questions surgissent lorsqu'on invoque ce concept dans le contexte de l'évaluation des emplois à partir des mêmes critères : premièrement, la question de savoir ce qu'on entend par «fonctions équivalentes» et, deuxièmement, la définition d'«égalité de rémunération». Aux yeux de Mme Weiner, le concept d'«égalité de rémunération pour fonctions équivalentes» est synonyme d'«équité salariale».

87. La preuve qui a été produite devant le Tribunal montre que la législation en matière d'équité salariale est axée sur une tendance qui admet une discrimination systémique à l'égard des emplois occupés majoritairement par des femmes. Certaines provinces ont adopté des lois en matière d'équité salariale afin de remédier aux pratiques salariales

discriminatoires en cernant et en fermant l'écart au moyen de plans d'équité salariale. Ce type de législation est «proactif» puisque, d'expliquer Mme Weiner, son objet est de fournir un cadre pour redresser la discrimination salariale, plutôt que de blâmer les employeurs ou les syndicats pour la discrimination salariale historique. La différence entre une loi proactive et l'article 11 est que l'article 11 est axé sur le dépôt de plaintes, c'est-à-dire qu'un plaignant allègue la discrimination contre un groupe de comparaison donné. Puisque, aux termes de l'article 11, la discrimination salariale peut s'exercer au détriment des emplois occupés majoritairement par des femmes comme de ceux occupés majoritairement par des hommes, on peut présumer, selon Mme Weiner, qu'une plainte pourrait alléguer la discrimination à l'endroit des hommes.

88. Bien que les dispositions de l'article 11 soient étayées sur le principe de l'égalité de rémunération pour fonctions équivalentes, qu'on dénomme souvent l'«équité salariale», on tend aujourd'hui à donner un sens proactif à cette expression. En fait, il y a une différence importante entre le principe enchâssé dans l'article 11, qui repose sur le dépôt de plaintes, et l'approche proactive du problème de la disparité salariale que les experts du domaine acceptent aujourd'hui et dénomment l'«équité salariale». Les observations que Mme Weiner a faites dans son témoignage devant le Tribunal sont instructives et illustrent le problème auquel on doit faire face lorsqu'on applique les principes de la Loi, notamment

20

l'article 11, en vue de remédier à la discrimination systémique sur le plan du travail. Voici ce qu'elle dit à la p. 2124 du volume 16 :

[TRADUCTION] Je conviens avec vous que la loi de la commission des droits de la personne, incluant l'article 11, a été rédigée dans l'optique du dépôt de plaintes. Je pense que ce fut une erreur. Mais ça, on ne le savait pas en 1977 au moment de sa rédaction. Et en réalité, si cela est tout à fait convenable dans le cas de nombreuses questions qui relèvent de la Loi sur les droits de la personne, on ne peut en dire autant de la discrimination systémique ou de quelque chose d'aussi compliqué que le processus de fixation des salaires.

Alors je pense que vous avez raison, il y a, à mes yeux, une anomalie si on considère que le législateur a inclus une méthodologie qui correspondait à la façon dont nous nous imaginions que la discrimination se manifestait dans les années 1970 ainsi qu'une certaine pensée proactive à propos d'un autre problème, mais sans reconnaître que la question de l'équivalence

des fonctions et de l'équité salariale constituait un problème systémique qui se prêtait mal à une approche fondée sur le dépôt de plaintes. [c'est nous qui soulignons]

89. Mme Weiner, qui est -- avec M. Morley Gunderson -- coauteure de *Pay Equity: Issues, Options and Experiences*, résume à la fin du chapitre 8, aux pp. 127-28, leur conclusion sur la législation fédérale :

[TRADUCTION] L'équité salariale est une idée dont on peut dire qu'elle est mûre si on considère que depuis 1985 huit juridictions canadiennes ont adopté des dispositions en matière d'équité salariale. Au cours des dix années précédentes, deux juridictions seulement avaient adopté une législation sur l'équité salariale. Contrairement à la plupart des lois qui ont été adoptées par la suite, ces deux premières lois -- passées par le gouvernement fédéral et le Québec -- reposaient sur la notion de plainte. Le fait que la plupart des lois récentes prévoient un mécanisme d'exécution émanant de l'employeur montre bien l'incapacité des premières lois à régler les problèmes systémiques tels que celui de l'équité salariale.

90. Certains changements ont vu le jour par le biais du mouvement politique qui, aux États-Unis, a poussé à l'adoption de plans inspirés de la théorie de l'équivalence des fonctions («comparable worth») qui, en retour, ont créé un cadre à l'intérieur duquel ont été mises en évidence et jugées dignes de rémunération des qualifications jusque-là invisibles ou non reconnues qui avaient été associées historiquement au travail des femmes et des groupes minoritaires. Au Canada, le mouvement parallèle d'équité salariale a vu l'adoption de lois provinciales conçues pour corriger la discrimination systémique en matière salariale et la rémunération du travail accompli par les membres de groupes professionnels composés majoritairement de femmes. Il convient de signaler ici le préambule de la Loi de 1987 sur l'équité salariale (Ontario), qui dit qu'on doit recourir à des mesures de promotion sociale pour redresser la discrimination salariale systémique. Cependant, l'histoire législative de l'article 11 ne révèle pas la même motivation politique à l'égard de cette

loi, ce qui vaut aussi pour les autres lois adoptées au Manitoba, en Ontario, à l'Île-du-Prince-Édouard, en Nouvelle-Écosse et au Nouveau-Brunswick.

91. Au niveau provincial, la législation vise à corriger la discrimination systémique et prévoit un délai et une procédure pour

réaliser l'équité salariale. L'approche retenue y est orientée vers l'avenir et, bien qu'on reconnaisse les injustices passées, les redressements sont axés sur la réalisation de l'équité sur le plan tant de l'emploi que de la rémunération. En revanche, l'article 11 de la Loi repose sur la présentation de plaintes et ne dit rien sur les moyens à mettre en oeuvre pour atteindre l'égalité de rémunération pour fonctions équivalentes. Bien que l'Ordonnance prise en application de la Loi apporte plus de détails sur les quatre éléments essentiels du paragraphe 11(2) -- c'est-à-dire les qualifications, les efforts, les responsabilités et les conditions de travail -- et qu'elle définisse la méthode d'évaluation, qui est un employé, qu'est-ce qu'un groupe, etc., elle n'établit pas de programme ni ne décrit de méthodologie pour réaliser l'objectif d'éliminer la «discrimination systémique». C'est un phénomène que ne mentionnent expressément ni la Loi ni l'Ordonnance.

92. Mme Weiner, pour revenir à son témoignage, a par ailleurs fait le commentaire suivant à propos de l'article 11 (volume 16, à la p. 2115) : [TRADUCTION] «[L]a législation ne reconnaît pas que l'égalité de rémunération pour fonctions équivalentes était un problème systémique qui se prêtait mal à une mentalité où l'on doit se plaindre d'une situation.» C'est là que réside la difficulté concernant l'article 11, qui ne cadre pas entièrement avec l'évolution et l'application des principes de l'équité salariale (ou de la «comparable worth») au cours des deux dernières décennies. Néanmoins, compte tenu de la nature générale et de l'esprit de la loi, qui est de combattre la discrimination systémique, il est nécessaire d'adopter le raisonnement que le juge en chef Dickson a exprimé ainsi à la p. 1139 de l'affaire CN, supra :

[...] il est essentiel de créer un climat dans lequel tant les pratiques que les attitudes négatives peuvent être contestées et découragées.

93. Le processus de fixation des salaires dans la fonction publique fédérale est un processus hautement complexe qui s'est développé sur plusieurs décennies, chacun apportant de nouvelles tendances et réalisations, dont la plus marquante fut l'introduction de la négociation collective dans les années 1960. L'avènement de la négociation collective a amené la négociation des conventions collectives proprement dites qui, en retour, influent sur la détermination des taux de rémunération. En général, la classification des emplois au sein de la fonction publique fédérale a été déterminée par un processus d'évaluation des emplois; toutefois, aucun processus unique n'a jamais été prescrit et il en a résulté une structure de classification de groupes professionnels multiples sans plan commun d'évaluation des emplois. Les taux de rémunération ont été fixés au moyen de ce processus, avec l'aide d'études du marché de l'emploi effectuées en grande partie, jusqu'en 1992, par le Bureau de

recherches sur les traitements. Le système de classification est manifestement en voie de réforme depuis 1990.

94. Des preuves ont été produites comme quoi le gouvernement du Canada s'est engagé à simplifier le système de classification des postes dans la fonction publique au moyen d'une initiative appelée Fonction publique 2000. Dans le cadre de FP 2000, le gouvernement s'est notamment engagé à rémunérer les fonctionnaires équitablement, sans discrimination fondée sur le sexe et suivant le principe de la parité salariale pour fonctions équivalentes. On est à mettre en oeuvre un nouveau système de classification pour respecter ces engagements.

95. D'après la preuve documentaire présentée, un groupe de travail est à examiner et développer cette initiative; jusqu'ici il a produit, en novembre 1992, le texte préliminaire d'une brochure qui sert de guide aux employés du secteur public afin qu'ils puissent établir des descriptions de travail exemptes de discrimination sexuelle.

96. Le témoignage des experts révèle que les systèmes de rémunération qui s'appuient sur les études du marché de l'emploi peuvent entraîner des disparités salariales pour des emplois réputés être équivalents. La recherche montre que le marché reflète la tendance historique à verser des salaires inférieurs pour le travail accompli principalement par des femmes. En général, les taux du marché sont établis en recourant à des systèmes traditionnels d'évaluation des emplois, lesquels perpétuent le problème de la sous-estimation du travail des femmes du fait que ces systèmes n'ont pas été conçus pour saisir les compétences associées au travail des femmes.

97. Les spécialistes de l'équité salariale expliquent que les systèmes et pratiques de rémunération courants sont empreints de discrimination sexuelle. Historiquement ces systèmes et pratiques ont sous-estimé le travail des femmes. Puisque l'article 11 a pour objet d'éliminer de la rémunération la discrimination fondée sur le sexe, à partir de la valeur intrinsèque d'un emploi, tout système qu'on utilisera pour évaluer les emplois doit être conçu pour éliminer les facteurs qui contribuent à un parti pris contre l'autre sexe et pour inclure les facteurs qui saisiront les compétences associées au travail de femmes qui, par le passé, ont été négligées.

98. Nous sommes d'avis que l'article 11 a une fonction de redressement des iniquités salariales qui surgissent entre des emplois jugés équivalents par un processus d'évaluation. L'iniquité salariale, ou l'écart salarial qui en résulte, est la différence entre le taux de

rémunération qu'un employeur verse à des employés de sexe masculin et celui qu'il accorde à des employés de sexe féminin dont les fonctions sont équivalentes. Il ne s'agit pas de la disparité salariale globale mentionnée par Mme Armstrong et qui, à ses yeux, équivaut à l'écart entre les salaires versés aux hommes et aux femmes, mais dont la cause peut être liée à d'autres facteurs en plus de la discrimination fondée sur le sexe. La réponse que Mme Armstrong a donnée à la question suivante est informative (volume 179, à la p. 22879, ligne 2-7) :

[TRADUCTION]

Si l'équité salariale était réalisée dans toutes les professions, dans tous les emplois, la disparité salariale disparaîtrait-elle?

23

Le témoin: La disparité salariale globale ne disparaîtrait probablement pas entièrement, non. Il pourrait toujours subsister un écart.

99. Nous devons nous assurer que les plaintes cherchent à redresser un écart salarial imputable à la discrimination sexuelle et non à d'autres facteurs. Il semble évident que l'existence d'une disparité salariale n'est pas en soi une preuve de discrimination. Soutenir le contraire réduirait à rien tout le processus d'évaluation des emplois, dont l'objet est de comparer les emplois selon un plan ou système de cotation du travail où l'on applique les critères prescrits par le paragraphe 11(1) de la Loi.

100. Nous concluons par ailleurs que l'article 11 vise à éliminer l'inégalité économique créée par la discrimination salariale fondée sur le sexe. Cette discrimination est involontaire, comme le montre clairement la décision que le juge en chef Dickson a rendue dans l'affaire CN, supra. Il s'agit néanmoins d'une forme de discrimination subtile qui s'est glissée dans les pratiques d'emploi au fil des ans depuis que les femmes ont intégré la population active. Nous reconnaissons, après avoir entendu le témoignage d'expert de Mme Weiner, de Mme Armstrong et de M. Willis, que la discrimination systémique se manifeste dans les systèmes et finit par s'intégrer aux pratiques de fixation des salaires des organisations, et que la classification des emplois peut résulter de la discrimination systémique. Comme la discrimination systémique fait partie d'un système qui n'a jamais été conçu pour être discriminatoire, Mme Weiner estime que la parité salariale ne peut être réalisée instantanément ou rapidement.

101. En tant que disposition remédiate, l'article 11 permet de corriger les disparités salariales faisant partie des pratiques de

rémunération des employeurs. Selon M. Willis, l'article 11 n'est pas une mesure réelle d'équité salariale puisqu'il vise les disparités salariales. Par contre, selon le témoin, c'est probablement une première étape dans la voie de la véritable équité salariale, puisque les dispositions de cet article obligent les employeurs à porter les salaires payés aux femmes au même niveau que les salaires qu'ils versent aux hommes. M. Willis a déclaré ce qui suit au volume 29, p. 3760, ligne 22, jusqu'à la p. 3761, ligne 9 :

[TRADUCTION] Le concept d'équité salariale se rapporte à la rémunération sans discrimination fondée sur le sexe; autrement dit, la rémunération basée sur la valeur intrinsèque d'un emploi plutôt que sur la valeur de l'emploi sur le marché.

Je reconnais qu'il y a une école de pensée qui dit : Laissez tomber l'évaluation des emplois, donnez-nous simplement l'argent.

Mais pour arriver logiquement à établir la valeur intrinsèque d'un emploi, il faut appliquer un plan d'évaluation des emplois; c'est-à-dire que l'évaluation de l'emploi permet de le décomposer et d'examiner le niveau des qualifications, des efforts et des responsabilités qui sont requis, ainsi que les conditions de travail qui s'y rattachent.

102. Si les employeurs ont recours à des systèmes d'évaluation des emplois qui ont un parti pris en faveur du travail des hommes, on pourra le

constater dans l'écart entre les salaires payés aux hommes et aux femmes exerçant des fonctions considérées comme équivalentes. Les systèmes d'évaluation des emplois qui traditionnellement favorisent le travail des hommes sous-estiment les qualifications exigées pour les emplois désignés un travail de femmes ainsi que la nature de ces emplois. L'évaluation traditionnelle des emplois est le plus souvent conçue pour valoriser les caractéristiques du travail des hommes. En revanche, l'évaluation d'emploi qui s'inspire de l'équité salariale a pour but de recourir aux systèmes qui éliminent de l'évaluation du travail la discrimination fondée sur le sexe.

103. A ce stade-ci, il convient de rappeler certaines des circonstances qui ont mené à la situation dont le Tribunal est saisi. Dans le cadre des travaux du Comité mixte syndical-patronal sur la mise en oeuvre de la parité salariale, les parties se sont engagées dans une étude proactive dans l'intention de mettre en oeuvre le principe de l'égalité de rémunération pour fonctions équivalentes tel qu'il est incorporé à

l'article 11 de la Loi. Elles ont retenu les services d'un expert en équité salariale, M. Willis, pour les aider à réaliser cette étude et ont eu recours au plan Willis d'évaluation des emplois pour évaluer des emplois choisis dans les groupes professionnels de la fonction publique fédérale composés majoritairement d'hommes ou de femmes. Le Comité mixte n'a jamais mené sa tâche à terme. Les données produites par l'étude font maintenant partie de la preuve produite devant le Tribunal. La Commission s'est appuyée sur ces données dans son enquête concernant les plaintes, et ces données servent maintenant de preuve de la violation de l'article 11 de la Loi. La Commission et l'Alliance demandent au Tribunal d'accepter les cotations numériques comme preuve de la valeur du travail. Ces résultats, soutiennent-elles, peuvent servir à établir l'équivalence des fonctions et sont la preuve d'une disparité salariale. L'intimé, pour sa part, fait valoir que les résultats ne sont pas fiables. On nous demande de déterminer si on peut se fier aux résultats.

104. Nous avons précédemment fait allusion aux étapes que comportent un plan d'évaluation des emplois et dont le juge L'Heureux-Dubé a fait état dans l'arrêt *Syndicat*, supra. Le Tribunal a entendu de longs témoignages sur le processus Willis qui incorpore les étapes typiques de l'évaluation d'emploi. La fiabilité des résultats, qui est la question que nous devons trancher, concerne surtout une étape importante du plan, à savoir son application par les évaluateurs à qui on avait confié l'analyse de l'information sur les emplois et la cotation de chacun des facteurs compris dans le plan Willis.

105. M. Willis et Mme Weiner conviennent que l'évaluation des emplois, si on veut qu'elle soit efficace pour éliminer la discrimination, doit s'effectuer de façon systématique. En même temps, il faut comprendre que l'évaluation d'emploi est un processus foncièrement subjectif. La question de savoir ce qui constitue de la «discrimination» ou un «parti pris» est complexe et elle est au cœur des arguments qui nous ont été présentés. L'avocate de la Commission se reporte à l'Ordonnance de 1986 sur la parité salariale adoptée en vertu des dispositions de la Loi, et plus particulièrement à l'alinéa 9(a), qui est ainsi libellé

9. Lorsque l'employeur a recours à une méthode d'évaluation pour établir l'équivalence des fonctions exécutées par des employés dans le même établissement, cette méthode est utilisée dans les enquêtes portant sur les plaintes dénonçant une situation de disparité salariale si elle:
  - (a) est exempte de toute partialité fondée sur le sexe;

106. L'avocate de la Commission, appuyée par l'Alliance, formule dans les termes suivants la question de la fiabilité des résultats sur laquelle le Tribunal est appelé à se prononcer :

Existe-t-il une variation systématique de la façon dont on a traité les questionnaires des hommes et des femmes (dans le processus d'évaluation) qui a été causée par la discrimination fondée sur le sexe ou qu'on peut attribuer à un tel parti pris? [c'est nous qui soulignons]

107. L'avocat de l'intimé maintient qu'on doit accorder une interprétation large à l'expression «partialité fondée sur le sexe» figurant à l'alinéa 9(a) de l'Ordonnance. Selon lui, toute partialité qui se traduit par un traitement différent des questionnaires des hommes et des femmes constitue une partialité fondée sur le sexe. Il fonde cet argument sur l'interprétation du témoignage de M. Willis, qui a décrit comme suit la partialité à la p. 208 du volume 26937, lignes 11 à 16 :

26

[TRADUCTION]

R. La «partialité» signifie simplement qu'on a traité différemment les emplois à prédominance masculine et ceux à prédominance féminine, que ce soit conscient ou non, l'écart dans le traitement représentant l'ampleur de la partialité.

108. D'après M. Willis, il est possible qu'une partialité soit indirectement fondée sur le sexe, situation qu'il qualifie plutôt de préférence fondée sur le sexe. Il a donné à titre d'exemple l'évaluateur qui aurait une préférence pour les personnes qui portent une chemise bleue ou qui ont un col bleu. Par exemple, on appelle «cols bleus» les gens exerçant un travail manuel, de signaler M. Willis. Si une préférence pour les cols bleus fait en sorte qu'un évaluateur cote plus favorablement les emplois manuels, cela n'est peut-être pas la manifestation d'une partialité fondée sur le sexe mais un parti pris en faveur des cols bleus. Selon M. Willis, cela donnerait le même résultat que s'il s'était agi d'une partialité directe fondée sur le sexe.

109. L'intimé fait valoir que la partialité fondée sur le sexe peut inclure des attitudes conscientes ou inconscientes envers l'autre sexe. A son avis, une partialité peut aussi être reliée à une caractéristique quelconque qui n'est pas le sexe de la personne proprement dit, mais qui y est indirectement reliée. L'avocat de l'intimé maintient que l'article 11 est conçu pour redresser les deux genres de partialité. Il soutient également que si une préférence pour les «cols bleus» donne lieu à un

traitement différent des questionnaires des hommes et des femmes, il faut alors parler de «partialité fondée sur le sexe» au sens de l'alinéa 9(a) de l'Ordonnance. L'avocat insiste sur le fait que la question à laquelle doit répondre le Tribunal n'est pas celle posée par la Commission mais plutôt la suivante :

Y a-t-il indication d'une tendance à traiter différemment les questionnaires des hommes et des femmes?

110. Il y a lieu de signaler que ni la Commission ni l'Alliance n'éprouvent de difficulté à accorder une interprétation large à l'expression «partialité fondée sur le sexe»; nous reproduisons à cet égard les observations suivantes de l'avocate de la Commission (volume 230, à la p. 30583, lignes 14 à 25) :

[TRADUCTION] Ce sont des choses très différentes. Toutes ces autres choses comme -- je veux dire, elles renvoient à quelque chose qu'on appelle le «travail sale». Je ne sais pas si vous avez une préférence pour les gens qui font un travail dur à l'extérieur. Si cela devait être relié au sexe et devait influencer sur la perception qu'ont les gens de ce travail et sur l'évaluation des emplois, et si, en fin de compte, cela avait un effet sur les emplois qui était relié au sexe, alors personne ne peut contester le fait qu'il s'agirait d'une partialité fondée sur le sexe en contravention de l'article 9 de l'Ordonnance sur la parité salariale et de ce fait de l'article 11 de la Loi.

111. En formulant la question que doit trancher le Tribunal, l'avocat de l'intimé fait valoir que sa formulation ne requiert pas de facteur de

causalité pour expliquer le traitement différent des questionnaires des hommes et des femmes. Le désaccord entre les parties ne tient pas au fait que l'on voudrait ou non donner une large interprétation à l'expression «partialité fondée sur le sexe». Il découle plutôt de la question de savoir si l'article 11 exige l'établissement d'une cause lorsque l'on constate que les questionnaires des hommes et des femmes ont été traités différemment ou si, en revanche, il s'agit simplement de constater la différence de traitement sans qu'il soit nécessaire d'en préciser la cause. A l'appui du point de vue défendu par l'intimé, elles s'en remettent au sens de la partialité qui, à leur avis, ne nécessite pas l'établissement d'un lien de cause à effet aux termes de l'article 11 de la Loi.

112. Les parties divergent d'opinion sur l'analyse que la Commission a faite du processus d'évaluation des emplois et des données statistiques et sur les conclusions auxquelles elle est parvenue. Le litige porte sur les arguments de la Commission et de l'Alliance selon lesquels certaines des différences de traitement des questionnaires des hommes et des femmes entre les comités et les consultants ne tiennent pas à une partialité fondée sur le sexe ou reliée au sexe, mais s'expliquent plutôt par une «partialité fondée sur les valeurs». La Commission et l'Alliance s'appuient sur le témoignage de l'expert en statistique Sunter, dont les analyses, soutiennent-elles, démontrent que la partialité fondée sur les valeurs explique une partie, sinon la totalité, des différences de traitement entre le comité et le consultant. D'après le témoin Sunter, l'effet de la partialité fondée sur les valeurs a l'apparence d'une partialité fondée sur le sexe, et la différence de traitement entre les comités et les consultants peut aussi s'expliquer par l'une et l'autre formes de partialité.

113. Pour rejeter la nécessité de connaître la cause du traitement différent des questionnaires des hommes et des femmes, l'avocat de l'intimé invoque le témoignage de M. Willis. Ce dernier a maintes fois affirmé durant l'audience qu'une fois le processus d'évaluation terminé, il n'est nullement nécessaire d'explorer les raisons qui expliqueraient pourquoi les comités et les consultants ont produit des évaluations divergentes. Ce témoignage est bien résumé dans la lettre en date du 19 mai 1994 que M. Willis a envoyée à l'avocat de l'intimé, lettre dans laquelle il s'étend plus longuement sur les écarts dans l'évaluation des emplois. Voici le texte de cette lettre :

[TRADUCTION] Les disparités dans l'évaluation représentent un manque d'uniformité dans l'application du système d'évaluation. Par conséquent, il y a lieu de s'inquiéter de ces disparités et d'y regarder de plus près pour déterminer si elles ont amené les évaluateurs à traiter différemment les divers genres d'emplois.

Le pourquoi de ces disparités est important durant le travail des comités. Il peut être utile d'en comprendre les raisons afin d'adapter la formation des membres en conséquence. Par contre, une fois la phase de l'évaluation terminée, les raisons de toute disparité ne revêtent plus d'importance réelle. Ce qu'il importe de savoir, c'est si les évaluations traduisent une tendance quelconque à porter un jugement à partir d'un parti pris.  
[c'est nous qui soulignons]

(pièce R-164)

114. Au cours de l'audience, en plus de sa définition de la partialité en tant que différence dans le traitement des emplois occupés par les hommes et les femmes, M. Willis a exprimé son avis sur le sens à donner à la partialité fondée sur le sexe dans une étude sur l'équité salariale. Il dit ceci au volume 80, à la p. 9737, lignes 13 à 18 :

[TRADUCTION] Dans le contexte de l'étude sur l'équité salariale, la partialité fondée sur le sexe a trait à la mesure dans laquelle les emplois traditionnellement occupés par les membres d'un sexe sont rémunérés plus favorablement que ceux traditionnellement occupés par les membres de l'autre sexe.

115. M. Willis parle de partialité fondé sur le sexe en terme de traitement différent et de rémunération différente. Pour mieux comprendre ses définitions de la partialité, il est utile de se reporter à la théorie du traitement disparate dont il est fait état dans la décision *American Federation of State, County and Municipal Employees, ASL-CIO et al. v. State of Washington et al.*, Nos. 84-3569, 84-3590, 770 s. 2d 1401 (1985) United States Courts of Appeal, 9th Circuit.

116. Les demandeurs dans l'affaire américaine ont allégué que l'État de Washington avait exercé de la discrimination sexuelle en matière de rémunération, en contravention de l'alinéa 703(a) du Title VII de la Civil Rights Act of 1964, 42 U.S.C. Le Tribunal fédéral de première instance pour le district de l'ouest de Washington avait tranché en faveur de la catégorie d'employés de l'État dont au moins 70 pour cent étaient des femmes, et l'État en avait appelé de la décision à la Cour d'appel (9th Circuit). Un fait important mérite d'être signalé dans la décision du Tribunal fédéral de première instance, à savoir que M. Willis avait mené une étude en 1974 pour examiner et relever les disparités salariales parmi les catégories d'emploi occupées majoritairement par des hommes en comparaison avec les catégories d'emploi occupées majoritairement par des femmes, et ce, à partir de la valeur des emplois. Dans le rapport Willis de 1974 admis en preuve, l'auteur concluait -- en se basant sur le contenu des 121 classifications d'emploi évaluées -- que les catégories féminines étaient en général moins bien rémunérées que les catégories masculines pour des emplois de valeur comparable, et que la disparité était de l'ordre de 20 pour cent. L'étude de M. Willis avait considéré les postes occupés par les hommes et par les femmes comme étant de valeur comparable. La notion de valeur comparable telle que la définit l'État, selon le Tribunal fédéral, signifie l'égalité de rémunération pour des postes auxquels se rattachaient des responsabilités, un jugement et des connaissances semblables.

117. En première instance, le tribunal avait conclu à la violation du Titre VII en s'appuyant sur les théories américaines de l'effet disparate et du traitement disparate. Comme le tribunal l'a expliqué dans son jugement, le Titre VII interdit deux types de discrimination en matière d'emploi : (i) le traitement défavorable et volontaire d'employés à partir de critères exclus; et (ii) les pratiques ayant un effet discriminatoire, à

29

savoir les pratiques en apparence neutres qui ont un effet discriminatoire et qui ne sont pas justifiées par des impératifs commerciaux.

118. Il en a été appelé de la décision du Tribunal de première instance devant le juge Kennedy de la Cour d'appel des États-Unis qui, après avoir examiné les allégations de traitement disparate, a conclu que les syndicats n'avaient pas réussi à établir *prima facie* le bien-fondé de leurs accusations de discrimination sexuelle selon la prépondérance de la preuve. Dans ses motifs, le juge Kennedy formule le commentaire suivant sur l'étude Willis (à la p. 1408) :

[TRADUCTION] Nous rejetons également la prétention de l'ASFCME selon laquelle l'État de Washington, ayant commandé l'étude Willis, s'était engagé à mettre en oeuvre un nouveau régime de rémunération fondé sur la valeur comparable telle que définie dans l'étude. La question de savoir si la valeur comparable est une approche viable de la rémunération des employés est matière à débat [...] En supposant, toutefois, qu'il puisse s'agir -- comme d'autres études sur l'évaluation des emplois -- d'un outil de diagnostic utile, nous rejetons une règle qui pénaliserait plutôt que récompenserait les employeurs pour l'effort et l'innovation dont ils auraient fait preuve en entreprenant une telle étude.

119. Tel qu'il est signalé dans la décision du juge Kennedy, selon la théorie du traitement disparate, l'intention ou le motif de l'employeur qui a adopté une politique contestée est un élément essentiel de responsabilité lorsqu'il est question de violation du Titre VII. Pour établir la responsabilité, le demandeur doit montrer que l'employeur a choisi une politique particulière à cause de son effet sur les membres d'un groupe protégé; il ne suffit pas qu'il allègue, selon cette théorie, que l'employeur était simplement conscient des conséquences négatives que la politique aurait sur un groupe protégé.

120. La Cour d'appel des États-Unis avait à trouver la preuve d'un élément intentionnel requise dans une affaire reposant sur la théorie du traitement disparate, contrairement à ce que qu'exige l'article 11

concernant la discrimination systémique, qui est une forme de discrimination involontaire. A notre avis, il faut situer les définitions de la partialité avancées par M. Willis dans le contexte de cette jurisprudence américaine qui, comme nous venons de le montrer, s'appuie sur des dispositions législatives différentes de la Loi et sur la nécessité de prouver l'intention, ce qui n'est pas le cas ici.

121. En ce qui concerne l'alinéa 9(a) de l'Ordonnance, supra, nous remarquons qu'il prévoit, notamment, qu'un employeur peut utiliser une méthode d'évaluation pour établir l'équivalence des fonctions si celle-ci est exempte de toute partialité fondée sur le sexe. Par contraste et en revenant à la question que nous devons trancher, c'est l'application de la méthode qui nous intéresse. A cet égard, il est utile de se reporter aux observations de Mme Weiner, plus particulièrement à l'affirmation suivante qu'elle a faite :

[TRADUCTION] Même si j'ai mentionné la partialité fondée sur le sexe dans l'évaluation des emplois et la partialité fondée sur le sexe dans

30

les systèmes d'évaluation, j'estime que c'est la partialité fondée sur le sexe dans l'application du système qui importe au premier point. Si vous donnez aux gens de l'information exempte de partialité ainsi qu'un système d'évaluation exempt lui aussi de partialité, ils peuvent toujours introduire la partialité fondée sur le sexe en l'appliquant.

122. Les parties s'accordent pour dire que le système d'évaluation, c'est-à-dire le plan Willis d'évaluation des emplois, est exempt de partialité selon toute norme raisonnable. D'après M. Willis, s'il y a une tendance à traiter différemment les questionnaires des hommes et des femmes, cela prouve qu'il s'exerce un parti pris systématique dans l'application de l'évaluation des emplois. Pour déterminer si les résultats sont empreints de partialité, M. Willis n'était pas prêt à donner son opinion en s'appuyant uniquement sur ses observations du processus suivi par le comité; il a plutôt choisi de s'appuyer sur l'analyse statistique des données. Selon M. Willis, il y a deux façons de déterminer si l'application du plan est empreinte de partialité : (i) les observations des consultants qui ont participé au processus; et (ii) l'analyse statistique.

123. Selon M. Willis, sur le strict plan des faits, lui et ses consultants qui étaient présents durant l'évaluation des emplois ont pu

découvrir et identifier la partialité directe fondée sur le sexe. Ils ont observé comment les évaluateurs répondaient aux questions qu'on leur posait sur les raisons pour lesquelles ils avaient évalué les emplois d'une façon particulière. Les consultants ne permettaient pas à un évaluateur de défendre son évaluation en invoquant des opinions ou conclusions.

124. La partialité indirecte, d'expliquer M. Willis, est cependant plus difficile à déceler. Habituellement la seule façon de la détecter consiste à faire une évaluation statistique des résultats pour voir s'il se dégage une tendance dans les cotations. Comme les évaluateurs ne sont pas d'habitude conscients de ces préjugés, ils n'ont pas l'impression de porter des jugements partiels fondés sur le sexe. Les évaluateurs attribueront les cotations de façon inégale selon qu'il s'agira d'emplois occupés par des hommes ou des femmes, et les emplois remplis par des hommes ou des femmes recevront systématiquement des cotes élevées ou faibles. En règle générale, une analyse statistique révèle ce genre de tendance si les préjugés ont indirectement influé sur le processus.

125. Tant la question posée par l'intimé que celle de la Commission, à notre avis, empêchent le Tribunal d'évaluer entièrement la question de la fiabilité. La fiabilité n'est pas purement statistique et les questions telles qu'elles sont posées limitent notre examen de la preuve à des mesures statistiques.

126. Il importe de ne pas perdre de vue le fait que les résultats ont été produits par un processus d'évaluation des emplois supervisé par le Comité mixte syndical-patronal sur la mise en oeuvre de la parité salariale, appuyé par les conseils et avis d'un expert en équité salariale. M. Willis a témoigné avoir recommandé au Comité mixte certaines mesures de protection destinées à assurer des résultats cohérents et fiables. Parmi ces mesures mentionnons les suivantes : une information fiable sur les

emplois, des comités d'évaluation équilibrés, la sélection et la cotation de postes-repères, des exercices de récapitulation, la formation des participants, des contrôles de qualité sous la forme de tests subis par les évaluateurs ainsi que la participation de comités et de consultants.

127. M. Willis a soutenu tout au long de l'audience que pour analyser les résultats il avait besoin d'un statisticien. A ses yeux, le rôle de ce dernier consistait à examiner les données et à faire le point sur l'ampleur de tout problème éventuel.

128. M. Willis a déclaré qu'il ne saurait défendre la fiabilité des résultats en s'appuyant uniquement sur le processus. A la question de savoir s'il examinerait le processus sans tenir compte des résultats, il a répondu ce qui suit (volume 78, à la p. 9570, lignes 12 à 22) :

[TRADUCTION]

R. [...] si toutes mes recommandations avaient été suivies, si j'avais été convaincu que les méthodes suivies étaient toutes solides, alors il est fort probable que j'aurais pu m'en tenir aux résultats de l'étude sans procéder à des tests.

Cela ne s'est pas produit. Je n'ai pas à ce jour approuvé les résultats de l'étude. Mais en dernière analyse ces tests vont dire quelque chose à quelqu'un, à moi ou à quelqu'un d'autre, à savoir si l'étude était solide ou pas.

129. Le témoin expert Weiner a affirmé que ce qui importe dans une évaluation d'emploi, c'est que le processus soit systématique, qu'il comporte une série d'étapes. En matière de parité salariale, l'évaluation des emplois devrait à son avis avoir pour objet d'appliquer le processus de façon juste à tous les emplois. Nonobstant l'avis de M. Willis, qui se préoccupe surtout des résultats plutôt que du processus, le présent Tribunal doit pouvoir évaluer les mesures d'autocontrôle s'exerçant durant le processus, et il doit pouvoir le faire d'un point de vue qui n'est pas strictement statistique. En effet, il faut que nous puissions analyser la fiabilité en évaluant comment les comités d'évaluation des emplois ont appliqué le plan Willis.

130. Nous avons le droit de tenir compte de l'ensemble de la Loi, y compris l'Ordonnance et les règlements pris en application de celle-ci, pour nous aider à interpréter le paragraphe 11(1). Voir Driedger, *The Construction of Statutes*, chapitre 11, 3e édition, par Ruth Sullivan. Nous sommes d'avis que la causalité est implicite dans les dispositions de la législation, compte tenu du stade d'évolution du principe de l'égalité de rémunération pour fonctions équivalentes à l'époque, du fait que la Loi est axée sur la nécessité de présenter des plaintes et de l'importance que la partialité fondée sur le sexe occupe dans le libellé de ses articles pertinents.

131. La disparité salariale, pour faire l'objet d'un redressement en vertu de l'article 11, doit avoir été causée par la discrimination fondée sur le sexe. L'alinéa 9(a) de l'Ordonnance est subordonné à la loi habilitante, la Loi canadienne sur les droits de la personne, et il est

autorisé par le paragraphe 27(2) de cette loi. Il existe une présomption en faveur de la validité des règlements à la lumière de leur loi habilitante. Le savant auteur de *Interpretation of Legislation in Canada*, 2e édition, Pierre André-Côté, fait l'observation suivante à la p. 310 de son ouvrage :

[TRADUCTION] Finalement il faut signaler que les règlements ne sont pas seulement réputés demeurer *intra vires*, mais aussi être formellement cohérents par rapport à la loi habilitante.

132. De plus, l'article 16 de la Loi de l'interprétation (Canada) dispose ce qui suit :

Les termes figurant dans les règlements d'application d'un texte ont le même sens que dans celui-ci.

133. Pour les besoins de l'article 11 de la Loi, nous ne croyons pas qu'il soit nécessaire de faire une distinction entre la discrimination fondée sur le sexe et la préférence fondée sur le sexe. Nous convenons avec les parties que l'expression «partialité fondée sur le sexe» contenue à l'alinéa 9(a) de l'Ordonnance devrait comprendre tout parti pris dans le contexte de l'évaluation des emplois qui a pour conséquence de favoriser un sexe au détriment de l'autre. De plus, nous sommes du même avis que M. Willis et l'intimé lorsqu'ils soutiennent qu'il n'est pas nécessaire de déterminer pourquoi un évaluateur donné est enclin à manifester de la partialité. Cependant, nous estimons nécessaire d'examiner les écarts entre les évaluations des comités et celles des consultants du point de vue statistique et en fonction du processus suivi afin de déterminer s'il y a eu partialité ou non.

134. Le Tribunal est par ailleurs d'avis que la causalité est implicite dans la Loi ainsi que dans l'Ordonnance. Il faut donc prouver que tout écart dans la façon de traiter les emplois occupés par des hommes par rapport à ceux qu'occupent des femmes est fondé sur le sexe. Ce point de vue cadre avec les opinions exprimées par M. Willis, puisque ce dernier ne parle pas simplement d'un traitement différent, mais d'un traitement différent qui [TRADUCTION] «influe sur un sexe ou sur l'autre» (volume 38, p. 4794) et d'une partialité qui «favorise» un sexe (volume 38, p. 4792). C'est l'aspect du traitement lié au sexe qui préoccupe M. Willis et qui préoccupe aussi le Tribunal.

135. Par conséquent, le Tribunal s'intéresse à l'aspect qui est fondé sur le sexe et, selon notre interprétation de la Loi, la question à trancher est la suivante :

Y a-t-il indication d'un traitement différent des questionnaires des hommes et des femmes, dans le processus d'évaluation, qui fut causé par la partialité fondée sur le sexe ou reliée au sexe, ou qu'on pourrait attribuer à un tel parti pris?

136. Nous allons maintenant examiner la question de savoir si le traitement des questionnaires des hommes et des femmes est empreint de partialité fondée sur le sexe. Notre enquête englobera la preuve relative

33

au processus qui a produit ces résultats ainsi que la preuve statistique présentée à l'audience.

#### IV. FARDEAU DE LA PREUVE

137. Dans l'affaire dont le Tribunal est saisi les questions en litige, en raison de la longueur et de la complexité de la preuve, ont été débattues en étapes par les parties. Dans un premier temps, il s'est agit d'établir jusqu'à quel point on pouvait se fier aux résultats issus de l'étude menée par le Comité mixte syndical-patronal sur la mise en oeuvre de la parité salariale. D'une part, on affirme que les résultats -- à la lumière d'une norme raisonnable -- sont fiables et exempts de parti pris contre l'autre sexe. D'autre part, on soutient qu'il n'est pas possible de se fier aux résultats puisqu'ils sont empreints de discrimination fondée sur le sexe au point de ne pas permettre au Tribunal de statuer sur le litige.

138. L'expression «fardeau de la preuve» décrit l'obligation qui incombe à l'une ou l'autre partie d'établir une prétention ou d'établir les faits relatifs à une question particulière. Voir M. N. Howard, Phipson on Evidence, 14e édition, (Londres; Sweet & Maxwell, 1990), par. 4-01.

139. Dans *Miller v. Minister of Pensions*, [1947] 2 All E.R. 372, (K.B.), Lord Denning, à la p. 374, définit comme suit le degré de probabilité qui se rattache au fardeau de la preuve dans une cause civile :

[TRADUCTION] Ce degré est bien établi. Il doit s'agir d'un degré de probabilité raisonnable mais pas aussi élevé que ce qu'on exige dans une affaire criminelle. Si la preuve est telle que le tribunal puisse affirmer «Cela nous apparaît plus probable que le contraire», alors la partie à qui incombe le fardeau de la preuve s'en est acquittée; en revanche, si les probabilités sont égales, elle n'a pas réussi à le faire.

140. L'avocate de la Commission, dans son exposé introductif, a concédé que le fardeau d'établir *prima facie* le bien-fondé de la plainte incombait à l'Alliance et à la Commission (volume 218, à la p.28337).

141. Cette concession de la part de l'avocate de la Commission reconnaît simplement la règle de preuve énoncée fréquemment par les cours et contenue dans les traités sur le sujet.

142. Selon le point de vue exprimé par le juge Sopinka dans *The Law of Evidence in Canada*, (Toronto: Butterworth, 1992), une preuve *prima facie* n'entraîne pas obligatoirement une décision spécifique à moins qu'une règle de droit particulière exige une telle conclusion. Après avoir examiné et analysé plusieurs décisions de la Cour suprême du Canada dans lesquelles les juges diffèrent d'opinion, les auteurs affirment qu'une preuve *prima facie* permet simplement à la cour de débouter l'intimé en l'absence de preuves contraires. Les auteurs citent en y souscrivant le passage suivant de l'arrêt *R. c. Girvin* (1911), 45 R.C.S., 167 (C.S.C.), à la p. 169 :

34

[TRADUCTION] J'ai toujours compris que la règle, dans une cause criminelle, n'obligeait pas la Couronne à faire plus que produire une preuve qui, si elle n'est pas réfutée et si on la croit, suffit à établir *prima facie* une prétention sur laquelle un jury pourrait fonder un verdict. [c'est nous qui soulignons]

143. Ce passage a récemment été adopté dans *R. c. Mezzo* [1986], 1 R.C.S. 802 (C.S.C.) et les savants auteurs concluent à la p. 73 que :

[TRADUCTION] [L]es expressions «preuve *prima facie*» et «prétention établie *prima facie*» sont sans signification si l'auteur n'en précise pas le sens. Si on veut être clair et concis, il est préférable [...] d'expliquer l'effet probant qui résultera de l'établissement de certains faits plutôt que d'utiliser à tort et à travers ces locutions latines.

144. Comme le sens qu'il faut accorder en l'espèce à l'expression «fardeau de la preuve» ne semble pas faire l'unanimité, nous nous reportons une fois de plus à Phipson on Evidence, supra. Selon le savant auteur, cette expression a trois sens :

- (i) Le fardeau de persuasion, le fardeau de la preuve en tant que question de droit, c'est-à-dire le fardeau de prouver une prétention selon la prépondérance de la preuve;

- (ii) La fardeau de présentation, le fardeau de produire la preuve;
- (iii) Le fardeau d'établir l'admissibilité de la preuve.

145. Dans une cause civile, la fardeau de persuasion, aussi appelée «charge ou fardeau ultime», incombe à la partie qui fait valoir en substance le côté affirmatif ou positif d'une question et qui est fixée au début du procès ou de l'audience par l'état des conclusions, c'est-à-dire les plaintes effectuées en vertu de la loi. Il est par ailleurs établi en tant que point de droit que le fardeau demeure inchangé tout au long de l'audience -- il reste exactement là où les plaintes l'avaient placé; aussi se déplace-t-il rarement, sauf dans des circonstances spéciales.

146. Le fardeau ultime surgit normalement après que la preuve a été présentée et il s'agit alors de savoir si le juge des faits a été convaincu du bien-fondé de la prétention ou de la cause selon la norme de preuve civile ou criminelle. Le fardeau ultime, cependant, surgit ordinairement après qu'une partie a d'abord satisfait un fardeau de présentation par rapport au fait ou à la question en cause. Voir *The Law of Evidence in Canada*, supra, à la p. 58.

147. En d'autres termes, le fardeau ultime n'influe pas sur le processus décisionnel si le juge des faits en litige peut tirer une conclusion déterminée en s'appuyant sur la preuve. Si, toutefois, la preuve laisse celui-ci dans l'incertitude, le fardeau ultime est appliqué afin que la question soit tranchée d'après la prépondérance des probabilités. Voir aussi *The Law of Evidence in Canada*, supra, à la p. 60, qui citent un passage d'une décision rendue par le Conseil privé dans

*Robins v. National Trust Company*, [1972] 2 D.L.R. 97, qui dit notamment ce qui suit :

[TRADUCTION] Mais le fardeau en tant que facteur déterminant de toute l'affaire ne peut surgir que si le tribunal n'arrive pas à trancher le litige parce que les éléments de preuve pour et contre s'équilibrent. C'est alors que le fardeau de la preuve sera déterminant.

148. On peut comparer ce passage aux observations que le juge McIntyre a faites dans *Commission ontarienne des droits de la personne c. Simpsons-Sears* [1985] 2 R.C.S. 536, à la p. 558 :

Mais, faute de mieux en pratique, on a jugé nécessaire, pour assurer une solution claire dans toute instance judiciaire, d'attribuer le fardeau de la preuve à l'une ou l'autre partie, pour les départager.

149. Le fardeau de présentation, par contre, peut se déplacer constamment tout au long de l'audience, selon la prépondérance qui est accordée à l'un ou l'autre niveau de preuve. Le fardeau de la preuve, en ce sens, incombe à la partie qui serait déboutée si aucune preuve n'était produite, ou si aucun autre élément de preuve, selon le cas, n'était produit par l'un ou l'autre côté. En matière civile le fardeau de la preuve peut être satisfait par tout élément de preuve qui suffit à établir les allégations de la partie plaignante jusqu'à preuve contraire. En l'espèce, il revient au Tribunal de décider en droit si la preuve produite est suffisante pour satisfaire au fardeau de présentation, autrement dit, pour établir *prima facie* la prétention de la plaignante. Voir *Phipson on Evidence*, supra, au par. 4-10(b).

150. Le fardeau de la preuve dans toute affaire particulière dépend de circonstances qui ont donné lieu à l'action. En général, selon *Phipson* la règle qui s'applique est la suivante : [TRADUCTION] «la partie qui invoque la loi devrait être celle qui établit en premier le bien-fondé de son action.» Cette règle repose sur le bon sens; il va de soi, par ailleurs, qu'il est plus difficile d'établir le volet négatif d'une prétention que son aspect affirmatif. Voir *Robins v. National Trust Co.*, supra; *Constantine Ligne v. Imperial Smelting Corp.*, [1942] A.C. 154, à la p. 174, Lord Maugham.

151. Dans son exposé oral, l'avocate de la Commission affirme ce qui suit (volume 218, p. 28349, ligne 25, jusqu'à la p. 28350, ligne 11) :

[TRADUCTION] S'il est créé un processus que les spécialistes considèrent comme le meilleur pour mettre en évidence la discrimination fondée sur le sexe [...] alors il n'y a aucune raison de chercher plus loin. En pareil cas, l'existence d'un processus fiable est établie *prima facie* et, en l'absence de preuve contraire, cela suffirait pour conclure à la fiabilité du processus. [c'est nous qui soulignons]

152. Cette affirmation plutôt générale de l'avocate est étayée sur les arrêts *Farnquist v. Blackett-Galway Insurance Ltd.* (1969), 72 W.W.R. 161 (C.A. de l'Alberta) (juge d'appel Allen), aux pages 172-73, et *OPSEU v.*

Ontario (Ministry of Community and Social Services) (1986), 15 O.A.C. 78 (C.D.), à la p. 79, qui portent sur la preuve à établir selon la prépondérance des probabilités.

153. Il n'est pas clair ce que l'avocate veut dire par «S'il est créé un processus». Par souci de clarification, si l'avocate se réfère aux procédures et structures mises en place par M. Willis et par le Comité mixte syndical-patronal sur la mise en oeuvre de la parité salariale pour l'évaluation des emplois, la preuve a établi qu'en général les structures étaient conformes aux exigences de la Loi et de l'Ordonnance ainsi qu'aux principes d'équité salariale reconnus par les spécialistes.

154. En supposant que le processus englobe non seulement les procédures et les structures, notamment le plan d'évaluation, la formation des évaluateurs, les questionnaires et la collecte de l'information, mais également -- ce qui est plus important, selon M. Weiner -- l'application du système d'évaluation sans discrimination fondée sur le sexe, alors on pourrait accepter l'affirmation de l'avocate de la Commission comme étant correcte. Mais l'avocate, au volume 218, p. 28357, ligne 22, jusqu'à la p. 28358, ligne 7, poursuit son argumentation ainsi :

[TRADUCTION] L'objet de tout cela est de faire le lien avec le déplacement du le fardeau de la preuve qui incombe à l'intimé. Une fois que les plaignants ont établi prima facie le bien-fondé de leur prétention et que le fardeau de la preuve incombe dorénavant à l'intimé, il revient à ceux-ci de prouver selon la prépondérance des probabilités [...] que la discrimination fondée sur le sexe ou l'allégation qu'ils avanceront, quelle qu'elle soit, est en fait la cause de la réalité constatée et, par conséquent, constitue la raison pour laquelle on ne peut se fier aux résultats. [c'est nous qui soulignons]

155. Le déplacement du fardeau de la preuve mentionné par l'avocate dans le passage précité ne libère pas la partie qui soutient la position affirmative, en l'occurrence la Commission et l'Alliance, de l'obligation de produire la preuve établissant que le Tribunal peut et doit se fier aux résultats de l'étude pour trancher le litige. Si, à la lumière de l'ensemble de la preuve, y compris la preuve anecdotique et statistique, le Tribunal peut en arriver à une conclusion déterminée, il ne sera pas nécessaire, à notre avis, d'invoquer le fardeau ultime pour trancher l'affaire.

156. Dans son exposé introductif, l'avocate de la Commission a admis que «le processus» suivi par le Comité mixte était vicié, mais pas au point d'entacher les résultats de nullité. Les témoignages anecdotiques des participants à l'étude, les consultants associés à M. Willis et certains

des évaluateurs nous amènent à mettre en doute l'impartialité de certains des évaluateurs ainsi que le fonctionnement de certains comités. Certains incidents troublants ont mis les consultants mal à l'aise vis-à-vis du processus et de son application. De plus, il y avait des divergences entre les comités et les consultants à propos des exercices de réévaluation menés par les consultants à divers stades de l'étude. Les analyses des résultats, en retour, ont entraîné une critique des données de la part d'experts en statistique. Cette courte description de certains des

problèmes qui ont surgi durant l'étude et qui ont pu influencer sur les cotations des évaluateurs n'est pas exhaustive.

157. Les problèmes concernant la fiabilité des résultats, qu'ils découlent du processus d'évaluation ou des analyses statistiques des réévaluations, seront abordés dans une partie subséquente de la présente décision.

158. L'employeur, le Conseil du Trésor, fait valoir que si le processus Willis a bien fonctionné, alors l'Alliance et la Commission ont établi *prima facie* la fiabilité du processus et nul n'est besoin d'examiner les résultats pour trouver d'autres preuves de fiabilité. Si, en revanche, le processus n'a pas bien fonctionné, alors la charge de la preuve, selon l'employeur, incombe toujours à l'Alliance et à la Commission, qui doivent produire «d'autres preuves» montrant que les résultats sont fiables et solides.

159. L'avocat de l'intimé a décrit ce qu'il entendait par «d'autres preuves», à savoir des analyses statistiques effectuées par des spécialistes de la statistique démontrant qu'il y a une structure systématique dans les disparités. Mais l'avocat soutient qu'en [TRADUCTION] «attaquant la crédibilité et l'utilité de ces disparités» la Commission et plus particulièrement l'Alliance se trouvent sans point de référence pour comparer les évaluations des comités et les réévaluations accomplies par les consultants, dont la crédibilité et l'impartialité ont été attaquées par l'Alliance.

160. Selon l'avocat de l'intimé, dans l'éventualité où le processus n'avait pas bien fonctionné et si la Commission et l'Alliance étaient empêchées de s'appuyer sur les analyses statistiques, celles-ci se trouveraient devant rien et elles n'auraient donc pas réussi à établir *prima facie* le bien-fondé de leur prétention.

161. Le Tribunal ne saurait toutefois accepter la proposition avancée par l'avocat de l'intimé. A notre avis, l'alternative n'est pas simplement d'accepter ou de rejeter l'une ou l'autre des possibilités qui nous sont présentées, voire les deux.

162. Dans le cadre de l'étude du Comité mixte syndical-patronal sur la mise en oeuvre de la parité salariale et sous la direction de M. Willis, l'approche adoptée par le Comité mixte a consisté à utiliser des tests statistiques pour valider le processus. L'ensemble du processus Willis est quelque chose de complexe qui inclut non seulement une évaluation des emplois, mais aussi de nombreuses étapes dont l'une consiste à valider les résultats en mesurant le coefficient d'objectivité des évaluations d'un évaluateur et d'un comité à l'autre. Certains des tests utilisés s'appuient sur des analyses statistiques. Ces tests font partie intégrante du processus que M. Willis utilise dans ses vastes études sur l'équité salariale, et ils ont manifestement été utilisés dans l'étude du Comité mixte. D'autres tests significatifs ont été effectués, notamment une réévaluation de 222 postes par le consultant de M. Willis, M. Jay Wisner, qui a appliqué des tests statistiques aux résultats de la réévaluation. Le seul test statistique qui n'a pas eu lieu durant le processus comme tel a

38

consisté en une autre réévaluation de 300 postes menée par la Commission dans le cadre de l'enquête qu'elle a faite à l'issue de l'étude.

163. L'analyse statistique effectuée par la Commission combine les réévaluations qui ont eu lieu durant l'étude menée par le Comité mixte et celles qui ont été exécutées par la suite. La combinaison de ces réévaluations, à notre avis, ne crée pas de cadre artificiel eu égard à la preuve relative au processus ou la preuve relative aux analyses statistiques. Ni les mesures concernant le processus ni les mesures statistiques ne peuvent être considérées sans référence aucune à l'une ou l'autre catégorie de mesures : en effet, elles sont interdépendantes en ce sens qu'on ne saurait comprendre les unes sans comprendre les autres.

164. Par conséquent, nous sommes fondés à examiner l'ensemble de la preuve et à la soupeser à la lumière de toutes les circonstances. Nous examinerons de façon très détaillée le témoignage des participants à l'étude, le témoignage expert des consultants, ainsi que le témoignage expert des statisticiens et d'autres personnes ayant participé à l'étude. Notre décision tiendra donc compte de tous les éléments de preuve qui nous ont été présentés au cours de l'audience.

165. A notre avis, les éléments qui sont requis pour que, en l'occurrence, il soit satisfait à la charge de présentation, sont les suivants; ils reposent sur les dispositions du par. 11 de la Loi, sur l'Ordonnance afférente à celle-ci et sur l'énoncé des prétentions des parties :

- (i) Les groupes plaignants sont composés majoritairement de femmes au sens de l'Ordonnance sur la parité salariale;
- (ii) Les groupes de comparaison sont composés majoritairement d'hommes au sens de l'Ordonnance sur la parité salariale;
- (iii) L'évaluation de l'équivalence des fonctions est fiable;
- (iv) La comparaison des salaires versés pour des fonctions équivalentes produit un écart salarial.

166. Tel qu'il a été mentionné précédemment, à ce stade-ci le Tribunal examinera le troisième élément, à savoir la question suivante : les emplois échantillonnés dans l'étude du Comité mixte ont-ils été évalués comme il faut de façon à ce qu'on obtienne des résultats fiables? Par ailleurs, il y a lieu de signaler que les parties, y compris l'intimé, le Conseil du Trésor, ont convenu que le plan Willis était en fait un plan d'évaluation non sexiste qui convenait à l'étude réalisée par le Comité mixte et qui saisissait les critères qui, aux termes du paragraphe 11(2) de la Loi doivent être mesurés.

167. L'avocate de la Commission, lorsqu'elle a abordé le troisième élément concernant la fiabilité des évaluations, a énuméré plusieurs considérations qui devaient entrer en ligne de compte, à savoir : le plan permet d'effectuer des comparaisons entre les professions; le processus a été conçu pour obtenir de l'information raisonnablement fiable sur les

emplois; il y avait en place des procédures additionnelles permettant de recueillir une information complète sur les emplois; le plan a en fait été appliqué de façon raisonnablement uniforme par les multiples comités; il y avait une certaine cohérence dans l'information sur les emplois; il y avait une certaine cohérence dans les résultats; et les données sur les salaires étaient raisonnablement fiables.

168. Ces considérations, pensons-nous, sont appropriées et seront utiles pour évaluer la preuve; aussi le Tribunal entend-il les appliquer

pour évaluer la preuve, anecdotique et statistique, dans les sections suivantes de la décision.

## V. NORME DE PREUVE

169. La «norme de preuve» détermine le degré de probabilité qui doit être établi par la preuve pour que la partie à qui incombe le fardeau de la preuve puisse être réputée avoir prouvé sa prétention ou une question qui s'y rattache.

170. Il y a deux niveaux de probabilité, selon que la cause à juger est de nature criminelle, auquel cas la preuve doit être établie au-delà de tout doute raisonnable, ou de nature civile, auquel cas le demandeur doit établir le bien-fondé de sa prétention, ou une question s'y rattachant, d'après la prépondérance des probabilités, ce qui revient à dire qu'il y a une plus grande probabilité que la conclusion avancée par le demandeur [TRADUCTION] «soit en substance la plus probable des interprétations possibles des faits». Voir les motifs du juge Duff dans *Clark c. Treking*, [1921] 61 Can. R.C.S. 608, à la p. 616.

171. La norme que le Tribunal a appliquée dans *Haldimand-Norfolk* (29 mai 1991), 0001-8 P.E.H.T. pour interpréter le paragraphe 5(1) de la Loi sur l'équité salariale de l'Ontario figure au paragraphe 24 de cette décision, que nous reproduisons ci-après :

[TRADUCTION]

24. Après avoir examiné soigneusement la preuve et les arguments présentés en l'espèce, nous concluons qu'il incombe aux parties de faire en sorte que la collecte d'information sur le contenu des emplois réponde aux exigences de la Loi de façon à préciser exactement l'habileté, l'effort et la responsabilité qu'exige normalement l'accomplissement de ce travail, ainsi que les conditions dans lesquelles il est normalement effectué, et ce, tant pour les catégories d'emplois à prédominance féminine dans l'établissement que pour les catégories d'emplois à prédominance masculine avec lesquelles la comparaison sera effectuée. Non seulement est-ce là une condition nécessaire de tout système de comparaison exempt de discrimination fondée sur le sexe, mais nous sommes également d'avis que l'article 5 de la Loi impose une norme d'exactitude, autrement dit, les qualifications, les efforts, les responsabilités et les conditions de travail doivent être enregistrés et cotés avec exactitude. [c'est nous qui soulignons]

172. Le paragraphe 5(1) de la Loi sur l'équité salariale de l'Ontario est ainsi libellé :

40

5 (1) Pour l'application de la présente loi, le critère qui sert à déterminer la valeur du travail se fonde sur l'habileté, l'effort et la responsabilité qu'exige normalement l'accomplissement de ce travail, ainsi que sur les conditions dans lesquelles il est normalement effectué.

173. Comme tel, le paragraphe 5(1) n'impose aucune norme particulière à laquelle les parties doivent satisfaire pour répondre aux critères.

174. Par conséquent, la décision que le Tribunal a rendue dans l'affaire Haldimand-Norfolk, pour autant qu'elle porte sur la norme à laquelle il faille satisfaire concernant la collecte d'information sur les emplois, constitue l'interprétation du Tribunal du paragraphe 5(1) de la Loi sur l'équité salariale de l'Ontario. Il y a lieu de signaler que les questions en litige dans cette affaire étaient de savoir si l'employeur avait adopté un système de comparaison exempt de discrimination fondée sur le sexe et s'il avait ou non négocié de bonne foi avec ses employés. La question de la fiabilité des résultats qui nous intéresse ici n'a pas été traitée directement. Cette question a trait au processus. Le processus nécessite une norme en fonction de laquelle on puisse évaluer la collecte de données sur les emplois et une norme en fonction de laquelle on puisse porter un jugement sur les procédures ayant servi à l'évaluation de ces données.

175. La question que nous devons examiner concerne notamment le format du questionnaire, les méthodes utilisées pour recueillir l'information sur les emplois, les procédures de suivi et mesures de protection, la composition et le fonctionnement des comités, l'application du plan d'évaluation des emplois et la vérification des résultats des comités au moyen d'analyses statistiques.

176. La Commission et l'Alliance ont invoqué l'application de la norme du raisonnable à l'appréciation de l'information sur les emplois et de l'évaluation des emplois. La norme du raisonnable devrait également s'appliquer à l'évaluation des «dommages-intérêts» qui, on le suppose, signifient la mesure des dommages indirects auxquels les plaignants ont droit en vertu de la Loi canadienne sur les droits de la personne.

177. L'avocat de l'intimé, dans ses arguments oraux sur le fardeau de la preuve, affirme ce qui suit (volume 226, p. 29671, lignes 16 à 24) :

[TRADUCTION] Voici un autre point sur la charge de la preuve. L'employeur soutient que la norme qu'on doit appliquer pour évaluer la fiabilité, la norme d'évaluation du processus qui permet de décider si on peut s'y fier, est la norme du raisonnable. Le processus a-t-il bien fonctionné? Il ne s'agit pas de savoir si le processus a marché à perfection ni si l'information sur les emplois était parfaite. L'employeur n'a jamais exigé la perfection.

178. En commentant l'affaire Haldimand-Norfolk, l'avocat fait également les observations suivantes à la page 29761, ligne 25, jusqu'à la page 29762, ligne 13, du volume 226 :

41

[TRADUCTION] Si je mets cela en contraste avec l'affaire Haldimand-Norfolk dans laquelle [...] le Tribunal a dit que nous voulons une norme d'exactitude [...] L'exactitude, c'est comme s'ils recherchaient la perfection. Je ne sais pas... Quoi qu'il en soit, ce que l'employeur maintient à cet égard, c'est qu'il faut nous demander si le processus a bien fonctionné, et non s'il a fonctionné à perfection.

179. Alors, au bout du compte, les parties ont elles-mêmes revendiqué la norme du raisonnable. Selon l'avocat de l'intimé, si on applique cette norme aux résultats de l'étude du Comité mixte, le Tribunal doit conclure que les résultats ne sont pas assez fiables pour étayer une décision favorable.

180. Quelle norme le Tribunal doit-il appliquer pour évaluer la fiabilité des résultats? Il faut examiner le concept du raisonnable en se demandant ce qu'on veut atteindre par l'équité salariale ou la valeur comparable et comment cela doit normalement se réaliser. Il y a, en outre, des considérations pratiques qui entrent en ligne de compte, à savoir quels effets cela aura-t-il, au lieu de travail, sur les parties en cause.

181. Tout au long de son témoignage, M. Willis, un expert reconnu dans son domaine, a souligné que la réalisation de l'équité salariale ou du principe de l'égalité de rémunération pour fonctions équivalentes -- comme entre des emplois occupés majoritairement par des hommes et des emplois occupés majoritairement par des femmes -- n'est pas une affaire scientifique, mathématique ou statistique. Il s'agit au contraire d'un «art» qui fait appel à la fois aux compétences analytiques, à la compréhension, à l'intuition et finalement à une évaluation subjective de l'emploi dans le cadre du plan, mais en respectant, ce faisant, la rigueur que le plan impose.

182. Le savant juge Fudge de l'Osgoode Law School, dans un article intitulé «Legal Standard for Gender Neutrality under the Pay Equity Act (Ontario): Achieving the Impossible?», fait le commentaire suivant à propos de la norme légale en fonction de laquelle on doit juger si un système de comparaison des emplois est exempt de discrimination sexuelle :

[TRADUCTION] [...] à ce jour il n'existe pas de méthode concluante qui permette de démontrer qu'un système de comparaison des emplois est non sexiste ou exempt de discrimination fondée sur le sexe. C'est pourquoi le Tribunal de l'équité salariale devrait adopter la norme du raisonnable à l'égard de la question du non-sexisme.

183. Elle expose ensuite les critères minimums pour élaborer un système d'évaluation des emplois non sexiste. Il n'est pas nécessaire d'examiner ces critères pour nos besoins puisque toutes les parties à l'enquête ont convenu que le plan Willis adopté au début de l'étude répondait aux critères minimums et était donc exempt de discrimination sexuelle. Par contre, les observations que l'auteur fait à la p. 20 peuvent être utiles en l'espèce :

[TRADUCTION] [...] il n'existe pas de méthode concluante qui permette de démontrer qu'un système de comparaison des emplois est non

42

sexiste ou exempt de discrimination fondée sur le sexe. C'est pourquoi le Tribunal de l'équité salariale devrait adopter la norme du raisonnable à l'égard de la question du non-sexisme.

184. Au sujet du processus proprement dit d'évaluation des emplois, l'auteur affirme ce qui suit :

[TRADUCTION] Même si la conception du système de comparaison des emplois évite scrupuleusement toute discrimination fondée sur le sexe, un parti pris contre l'autre sexe peut se glisser dans la cotation comme telle des emplois. Autrement dit, le système d'évaluation des emplois peut être juste, mais son application peut être empreinte de discrimination.

185. Le juge Fudge décrit ensuite l'utilisation des comités d'évaluation des emplois qui, s'ils sont bien constitués et s'ils suivent des procédures clairement définies, sont en mesure de réduire le risque de parti pris.

186. Nous nous reportons par ailleurs à l'observation de Mme Armstrong selon laquelle dans l'ensemble la disparité salariale ne disparaîtrait pas entièrement si l'équité salariale était réalisée dans tous les emplois.

187. Ce que font ressortir clairement ces observations ainsi que la nature du sujet, c'est que l'égalité de rémunération pour fonctions équivalentes est un but à rechercher qu'il est impossible de mesurer de façon précise et qui ne doit pas faire l'objet d'une norme d'exactitude absolue. De plus, le non-sexisme entendu dans son sens absolu est probablement irréalisable dans un monde imparfait, et l'on devrait donc se satisfaire de résultats raisonnablement satisfaisants fondés sur ce qui, selon le bon sens, constitue un règlement juste et équitable de tout écart discriminatoire entre les salaires payés aux hommes et ceux versés aux femmes pour des fonctions équivalentes.

## VI. RÉSUMÉ DES FAITS

### A. LE PLAN WILLIS

188. Durant l'étude sur la parité salariale, l'évaluation du travail s'est effectuée au moyen d'un plan d'évaluation des emplois.

189. Une des premières tâches du Comité mixte était de choisir un plan d'évaluation des emplois. Le Comité mixte a créé un sous-comité pour examiner les divers plans d'évaluation des emplois et lui faire des recommandations. Ce sous-comité, après avoir examiné plusieurs plans, a recommandé le plan Willis, conçu par M. Willis, avec quelques modifications mineures pour qu'il réponde mieux aux critères de la Loi. A la suite de consultations avec des représentants du Comité mixte, M. Willis a accepté d'apporter des changements au plan, notamment des modifications au tableau des conditions de travail.

190. La Commission a elle aussi examiné le plan Willis et exprimé des réserves à l'égard du traitement des «efforts» relativement aux conditions de travail. Elle s'inquiétait aussi de la façon dont le plan traitait de l'«obligation de rendre compte». M. Willis a accepté de modifier certains aspects du plan afin qu'on évalue l'effort physique et l'effort intellectuel dans les conditions de travail. Il a aussi accepté de modifier la façon dont le plan traitait l'obligation de rendre compte.

191. Tous les participants, y compris la Commission, ont paru satisfaits des modifications. M. Paul Durber, directeur de la Direction de la parité salariale à la Commission canadienne des droits de la personne et

expert en la matière de parité salariale, a présenté un témoignage d'expert pour expliquer de quelle façon le plan Willis satisfaisait aux exigences de l'article 11 de la Loi et à l'Ordonnance. M. Durber a affirmé qu'un élément essentiel du plan d'évaluation des emplois, pour les besoins de la parité salariale, était qu'il soit exempt de partialité fondée sur le sexe. Selon M. Durber, il n'y a rien dans le plan Willis qui, à première vue, indiquerait une partialité fondée sur le sexe et il n'y a rien dans le plan qui ferait qu'il serait difficile de mesurer le travail traditionnellement exécuté par les hommes par comparaison avec celui accompli par les femmes.

192. Le plan Willis est de conception complexe. M. Willis a élaboré ce plan en 1974 après avoir travaillé pendant trois ans pour le cabinet de consultants Hay & Associates. Il utilise un format matriciel qui permet d'évaluer les quatre facteurs: qualifications, efforts, responsabilités et conditions de travail, en les subdivisant en sous-facteurs. La conception matricielle permet d'évaluer jusqu'à trois sous-facteurs sur un même barème d'évaluation avec un total de quatre barèmes d'évaluation. Un barème d'évaluation présente les critères utilisés dans le plan Willis. Dans certains cas, un facteur est imbriqué dans un autre. Par exemple, l'habileté en communications interpersonnelles est mesurée à l'intérieur des niveaux de l'habileté en gestion; ainsi, la cote qu'on attribue à l'habileté en gestion influe sur le nombre de points accordés à chacun des niveaux de l'habileté en communications interpersonnelles.

44

193. Le plan Willis est un système de cotation numérique selon lequel on attribue des points à chacun des facteurs d'un emploi. La somme des points donne la cote globale de chaque emploi. Le plan Willis est conçu géométriquement. M. Willis a choisi un écart de 15 % entre chaque niveau du plan. Il trouve que ce pourcentage constitue un écart discernable dans les définitions sémantiques des différents niveaux des barèmes. Si les écarts étaient trop faibles, a-t-il signalé, les évaluateurs seraient incapables de faire un choix.

194. La détermination de la valeur relative de chaque emploi est fonction du nombre de points qui sont disponibles pour chaque facteur dans le plan Willis, et il y a un nombre infini de points disponibles. Le nombre relatif de points qui sont disponibles pour chaque facteur contribue à la détermination de la valeur relative des différents postes (volume 77, à la p. 9377).

195. Mme Nan Weiner, présidente de NJ Weiner Consulting, Inc. et consultante spécialisée en parité salariale et en équité en matière

d'emploi, a été acceptée par le Tribunal à titre d'experte en équité salariale et en rémunération. On lui a demandé son avis sur le plan Willis, qu'elle a appelé un système. Bien qu'elle n'eût pas travaillé avec le plan Willis, elle a affirmé que rien n'indiquait que celui-ci sous-estimerait d'une façon ou d'une autre les emplois féminins. A son avis, une des faiblesses du plan Willis serait peut-être que, vu l'ampleur et la diversité de la fonction publique fédérale, quatre niveaux d'habileté en communications interpersonnelles, ce n'est tout simplement pas suffisant pour faire les différenciations que requièrent certains postes. A son avis, le plan Willis attribue plus de points aux facteurs connaissances et qualifications, ainsi qu'à l'obligation de rendre compte ou aux responsabilités qu'aux facteurs efforts et conditions de travail; à certains égards, il favorise par ailleurs les cols blancs par opposition aux cols bleus. Selon Mme Weiner, il est important que les évaluateurs qui utilisent le système fassent en sorte, par leurs discussions, que les emplois manuels soient mesurés équitablement.

196. A son avis, ce n'est pas le caractère distinctif du travail mais la façon dont le système est adapté par l'utilisateur qui importe. Sur ce point, elle fait remarquer ce qui suit (volume 11, à la p. 1564, lignes 20-23) :

[TRADUCTION]

LE TÉMOIN : Ce qui importe, c'est que vous demandiez aux gens qui utilisent effectivement le système ce qu'ils ont fait pour s'assurer que le système était appliqué équitablement dans le cas des cols bleus.

197. M. Willis a témoigné que, dans le plan Willis modifié courant, les points ne figurent pas sur les barèmes. Les évaluateurs ne savent donc jamais ce que sont les points. Ils évaluent à l'aide de plus et de moins. C'est un logiciel qui détermine les points. Selon M. Willis, cela libère l'évaluateur, qui n'a pas besoin de connaître les rapports entre les différents emplois du point de vue des cotes.

198. Deux éléments émergent de la conception d'un plan d'évaluation des emplois dans le contexte de l'équité salariale. Premièrement, un plan d'équité salariale doit pouvoir saisir et évaluer comme il faut le travail accompli par les femmes et par les hommes. Deuxièmement, il faut attribuer une «pondération» aux divers facteurs du plan. Par exemple, pour la pondération des facteurs, le plan Willis accorde beaucoup plus de points

aux connaissances et aux qualifications qu'aux conditions de travail et à l'effort physique.

199. M. Willis a décrit le mécanisme de pondération de son plan. Il a témoigné que les pondérations avaient été validées à l'aide des taux de rémunération du marché. Mme Weiner a critiqué le fait d'utiliser le marché pour valider les pondérations, car les influences qu'exerce le marché sur les salaires dans le cas des emplois à prédominance féminine sont incompatibles avec la parité salariale. En clair, cela veut dire que lorsque les systèmes d'évaluation des emplois sous-estiment le travail traditionnellement accompli par les femmes, la situation est amplifiée sur le marché. Par conséquent, à son avis, le marché reflète la sous-estimation du travail des femmes.

200. La dernière validation officielle des pondérations utilisées dans le plan Willis a eu lieu en 1985. Tôt dans l'étude, M. Willis a accepté, à la demande du Comité mixte, d'effectuer une étude de validation des pondérations à cause de craintes exprimées par la Commission. M. Willis ne croyait pas qu'il était nécessaire de valider de nouveau les pondérations, car il utilisait le système de façon constante et n'avait aucune preuve empirique indiquant que les pondérations des facteurs étaient inappropriées. Les représentants de la partie patronale ont finalement décidé de ne pas faire l'étude de validation à cause de considérations financières.

201. Au cours des longues audiences tenues par le Tribunal, l'employeur a contesté la pondération du plan Willis et sa validité en tant qu'outil pour évaluer les emplois d'une façon exempte de partialité fondée sur le sexe. Le Tribunal a entendu une preuve considérable sur cette question. Ce n'est que lorsque les arguments écrits de l'avocat de l'intimé ont été disponibles que le Tribunal et les autres parties ont été avisés que l'employeur reconnaissait le plan Willis comme plan d'évaluation approprié et acceptable pour les besoins de l'étude. De plus, les parties ont alors convenu que le plan Willis répondait aux exigences de l'article 11 de la Loi et qu'il constituait un instrument approprié pour l'examen de ces plaintes aux termes de l'article 11. Nous nous reportons aux observations écrites de l'employeur, plus particulièrement au passage suivant (p. 11, par. 41) :

[TRADUCTION]

41. Néanmoins, pour les besoins de la présente affaire, l'employeur reconnaît que le plan Willis était un plan approprié à utiliser pour évaluer les emplois dans la fonction publique fédérale. Par conséquent, le Tribunal n'a pas à décider s'il est valable de pondérer le plan Willis.

202. De plus, dans sa plaidoirie, l'avocat de l'intimé a fait valoir ce qui suit (volume 218, à la p. 28453, lignes 4-11) :

[TRADUCTION]

Alors, ayant traité de tous ces points, nous affirmons ceci : Pour les besoins de ce litige, l'employeur reconnaît que le plan Willis était un plan approprié à utiliser pour évaluer les emplois dans la fonction publique fédérale. Ce que j'ai donc voulu dire clairement dans mon argumentation, c'est que nous ne soulevons aucune question relativement au plan Willis.

203. La Commission est tenue d'assurer le Tribunal que le plan Willis satisfait aux exigences de l'article 11 de la Loi ainsi qu'à l'article 9 de l'Ordonnance. A cet égard, M. Durber a assuré le Tribunal du bien-fondé de l'avis de la Commission, lequel confirme essentiellement que le plan Willis répond aux exigences de la Loi.

204. Durant leur plaidoirie, toutes les parties ont reconnu que le plan Willis constituait un outil approprié pour traiter les plaintes dont nous sommes saisis. Par conséquent, le Tribunal est convaincu et déclare de fait que le plan Willis est un outil qui répond aux exigences de la Loi et de l'Ordonnance en vue de l'évaluation des emplois qui sert de base à cette décision.

205. Le plan Willis offre un outil pour évaluer la valeur relative du travail. Comme tel il ne fournit toutefois pas de méthodologie pour déterminer quel est l'écart salarial entre les postes féminins et les postes masculins. Pour déterminer s'il y a disparité salariale, il faut comparer les évaluations des emplois masculins et féminins. Le système comme tel ne peut déterminer cet écart sans cette étape additionnelle.

## B. LE PROCESSUS WILLIS

206. M. Willis a mis environ 24 ans pour mettre au point le processus Willis, années pendant lesquelles il a travaillé à titre de consultant indépendant dans le domaine de l'évaluation des emplois aux fins de l'équité salariale. Le Tribunal a entendu de nombreux témoignages sur la mise en oeuvre du processus Willis dans l'étude sur la parité salariale. Plus précisément, la preuve a porté sur les évaluations qui ont débuté à l'automne de 1987 et pris fin à l'automne de 1989. Dans l'évaluation de la fiabilité, il nous paraît indiqué d'examiner chaque aspect du processus

afin de déterminer s'il a atteint ou non l'objectif visé, soit éviter la partialité fondée sur le sexe.

207. Le processus Willis sert à examiner, à coter et à évaluer les emplois. Les participants à l'étude avaient pour tâche de mesurer le contenu de chaque poste examiné et d'attribuer une cote reflétant le travail total de chaque poste ou emploi.

208. Bien que M. Willis ait témoigné que le plan d'évaluation devait être un instrument solide, il a aussi insisté pour dire que le processus à l'intérieur duquel le plan était utilisé était encore plus important que le

47

plan lui-même. Selon M. Willis, tout ce qui était fait dans le contexte du processus visait principalement à éviter les évaluations qui pourraient traduire des relations traditionnelles ou des attitudes stéréotypées. Le processus était conçu pour éviter tout ce qui pouvait être considéré comme de la partialité fondée sur le sexe. M. Willis a maintenu tout au long de l'étude que la vigilance durant le stade de l'évaluation revêtait une importance primordiale, et il a continuellement renforcé la nécessité d'obtenir des évaluations objectives, justes et équitables pour tous les postes.

209. A titre d'information, le processus sur lequel le Comité mixte s'est finalement mis d'accord n'était pas le choix préféré de M. Willis. Ce dernier a initialement recommandé au Comité mixte une proposition qui exposait les processus et procédures devant servir à l'étude. A cause de considérations financières dont devait tenir compte le côté patronal, sa proposition a été rejetée. M. Willis a alors préparé une proposition modifiée. Le Comité mixte a accepté la version modifiée après qu'un certain nombre de décisions sur l'opportunité de «faire ou faire faire» ont été prises relativement à certains aspects du processus Willis.

210. M. Willis a apporté les modifications suivantes à la phase de la collecte des données :

- (i) au lieu que ce soit ses consultants qui montrent aux employés choisis comment remplir les questionnaires, il a accepté que des fonctionnaires fédéraux soient formés pour donner cette formation;
- (ii) au lieu que ce soit ses consultants qui revoient les questionnaires remplis, il a accepté de former une équipe de

fonctionnaires fédéraux pour accomplir cette tâche sous la direction d'un consultant qui superviserait le processus;

(iii) au lieu que ce soit les consultants qui fassent les entrevues en personne avec des titulaires choisis, il a accepté de former une équipe de fonctionnaires fédéraux qui se chargeraient de ces entrevues au besoin; et

(iv) dans le dernier stade de l'étude, M. Willis a réduit le temps et l'effort que lui et ses consultants consacraient à la phase de collecte des données de l'étude.

211. M. Willis estimait que ces modifications donneraient lieu à une étude qui serait d'une qualité suffisante pour répondre aux exigences de la Loi; pour s'en assurer, il a mis en place un certain nombre de mesures de protection ou procédés d'appoint qui feraient en sorte qu'on obtienne des informations complètes et exactes sur les emplois.

212. Il nous paraît utile de séparer chaque étape du processus Willis et d'examiner la preuve qui s'applique à chaque étape, ce qui nous aidera à trancher la question de la fiabilité et à déterminer l'efficacité des mesures de protection retenues.

#### 48

##### (i) Collecte des données

213. M. Willis a témoigné que la collecte des données était une étape critique et très importante dans une étude de ce genre.

214. Une des sources d'information est la description de poste. Après avoir examiné un échantillon de descriptions de poste de la fonction publique fédérale, M. Willis a conclu que celles-ci étaient périmées et insuffisantes pour les besoins de l'étude.

215. Une deuxième source est le questionnaire à questions fermées, qui est semblable au questionnaire à choix multiples. Pour construire un tel questionnaire et bien le structurer, il faut connaître l'emploi en profondeur et de façon détaillée. Une grande connaissance du travail est nécessaire afin d'énumérer des choix sensés au répondant. Cette tâche est plus simple dans un petit établissement où la diversité des emplois est moins grande.

216. Le questionnaire à questions fermées a comme avantage d'exiger moins du titulaire du point de vue de la compréhension de l'étendue et du

contenu de l'emploi. C'est plutôt sur la connaissance et la compréhension de la personne qui construit le questionnaire qu'on doit miser. En revanche, si la personne qui structure le questionnaire ne connaît pas toute l'étendue de travail en cause ou si elle n'est pas consciente de toutes les considérations qui doivent entrer en ligne de compte dans un questionnaire servant à des fins de parité salariale, alors il se peut que le questionnaire incorpore une partialité fondamentale (volume 180, à la p. 22971). Il est relativement facile pour les employés de remplir un questionnaire à questions fermées, mais ce genre de questionnaire comporte une lacune fondamentale, selon M. Willis, car il permet à l'employé de porter des jugements de valeur sur son travail au lieu d'obtenir de lui des informations concrètes.

217. Une troisième source est un questionnaire à questions ouvertes, qui est plus difficile à remplir qu'un questionnaire à questions fermées. M. Willis préfère un questionnaire à questions ouvertes. Le témoin Armstrong a expliqué qu'un questionnaire à questions ouvertes devient surtout utile lorsqu'on doit recueillir des informations sur un large éventail d'emplois très différents à évaluer. Un questionnaire à questions ouvertes est aussi plus utile avec un effectif instruit, ce qui est le cas de la plupart des fonctionnaires (volume 180, à la p. 22971).

218. M. Willis a témoigné avoir construit son questionnaire de façon à obtenir des informations complètes, définitives et exactes sur les emplois (volume 68, à la p. 8542). Les données recueillies dans l'étude sur la parité salariale ont été obtenues par un questionnaire à questions ouvertes (le «questionnaire Willis»).

219. La quatrième et dernière source de données consisterait à confier à un groupe de travail d'analystes professionnels de postes la tâche d'interroger chaque employé et de préparer le document contenant les informations voulues. M. Willis a employé cette méthode à quelques reprises, mais il a affirmé qu'il ne serait «pas pratique» de procéder

ainsi dans le contexte de la fonction publique fédérale. (volume 29, à la p. 3696).

(ii) Le questionnaire Willis

220. M. Willis a discuté des avantages et des désavantages des questionnaires à questions ouvertes et fermées dans le contexte d'une étude d'envergure telle l'étude sur la parité salariale. Il préférerait un questionnaire à questions ouvertes ou libres par opposition à un

questionnaire à réponses fermées ou préétablies parce que son objectif était d'empêcher les titulaires de porter des jugements de valeur sur leur propre travail. C'est ce qui se produit lorsqu'on utilise un questionnaire à réponses fermées. M. Willis a fait la remarque suivante dans son témoignage (volume 65, à la p. 8084, lignes 12-14) :

[TRADUCTION]

Je pense que la chose importante, c'est que ce doit être l'évaluateur qui porte ce jugement de valeur, il ne faut pas laisser l'employé le faire.

221. Le questionnaire Willis a été utilisé dans de nombreuses autres études sur la parité salariale réalisées par Willis & Associates tant au Canada qu'aux États-Unis. Le Comité mixte a formé un sous-comité pour finaliser le format et le contenu du questionnaire qui, après modification, a été approuvé par le Comité mixte. Un guide a été annexé à chaque questionnaire pour aider l'employé à le remplir. Ce guide a aussi été modifié par le Comité mixte afin qu'il reflète le contexte de la fonction publique fédérale.

222. En résumant sa participation à la conception du questionnaire, M. Willis affirme ce qui suit (volume 60, de la p. 7429, ligne 18, jusqu'à la p. 7430, ligne 6) :

[TRADUCTION]

R. Je dirais que le questionnaire a été développé sur un certain nombre d'années. C'est probablement le questionnaire qui a été le plus travaillé, puisqu'il remonte à 1974. Nous avons essayé au fil des ans de le modifier pour qu'il soit plus facile à remplir, mais en même temps il demeure un questionnaire à questions ouvertes, ce qui me paraît nécessaire.

Le produit final, bien sûr, résulte des modifications qu'un sous-comité du Comité mixte patronal-syndical y a apportées. Je pense que nous pourrions difficilement espérer avoir un meilleur questionnaire que celui-là pour une étude de ce genre.

223. M. Willis a pris part au processus de modification des questionnaires et du guide et il a approuvé tous les changements apportés. Il a témoigné qu'il était satisfait du questionnaire et du guide qui avaient été utilisés dans l'étude sur la parité salariale (volume 62, à la p. 7654).

224. Il est prévu sur le questionnaire un espace où le superviseur du titulaire peut faire des commentaires. Le sous-comité du questionnaire avait discuté de cette partie du questionnaire et y avait apporté des changements. Selon M. Willis, ces changements étaient mineurs et il était satisfait des questions dans leur forme finale. Les questions adressées au superviseur étaient formulées comme suit :

[TRADUCTION]

Réviser attentivement le questionnaire complété, mais sans modifier ou éliminer quelle que partie que ce soit de la réponse initiale. Veuillez répondre aux questions ci-dessous. Nous vous invitons à consulter votre superviseur sur ce sujet.

1. Que considérez-vous être la plus importante tâche de ce poste et pourquoi? (Référez à la question III).
2. Commentez l'exactitude et l'état complet des réponses de l'employé
3. Veuillez signer à la page 34.

**IMPORTANT:** Le supérieur immédiat devrait réviser les réponses avec l'employé lorsqu'il y a divergences significatives d'opinion.

(Pièce HR-34)

225. M. Willis a expliqué que si on vérifie ainsi les informations sur les emplois, c'est afin de tenir compte de deux situations : premièrement, la tendance qu'ont certains employés de surestimer jusqu'à un certain point leur emploi lorsqu'il n'y a pas de superviseur pour revoir les informations; deuxièmement, et c'est là le point le plus important, souvent le superviseur a des renseignements additionnels que l'employé a pu oublier de mentionner et qui sont utiles pour évaluer le poste.

226. Un des problèmes relevés par M. Willis était celui d'obtenir de bons renseignements concernant les «postes de niveau professionnel complexes» (volume 68, à la p. 8544). Plus l'emploi est complexe et requiert un niveau de connaissances élevé, a-t-il expliqué, plus il fait appel à la compréhension et à l'interprétation de principes et de théories, d'où la difficulté accrue qu'éprouvent les titulaires à décrire leur travail. Dans un poste de haut niveau, il est plus difficile pour l'employé de documenter et de décrire son travail d'une façon que

l'évaluateur puisse comprendre. En revanche, un poste de préposé à l'entretien très simple où l'on applique des procédures spécifiques peut être documenté assez facilement.

227. Un autre problème lié à la collecte des données est qu'il faut accorder suffisamment de temps aux employés pour remplir le questionnaire. Chaque titulaire doit avoir assez de temps pour remplir le questionnaire, ce qui est fonction de la capacité de chacun de décrire son emploi par écrit. Ce n'est pas uniquement le temps qui est important ici, mais également l'effort et le soin que chaque titulaire doit y mettre.

51

228. Selon l'avis expert de M. Willis, le questionnaire était un bon outil pour obtenir des renseignements concrets et à jour sur les emplois. En évaluant la capacité du questionnaire Willis de recueillir suffisamment d'informations pour les comités d'évaluation, nous retenons les remarques suivantes de M. Willis (volume 62, à la p. 7686, lignes 16-22) :

[TRADUCTION]

Q. Si vous vous reportez aux fois où les consultants siégeaient, êtes-vous persuadé que, lorsque est venu le temps d'évaluer les questionnaires, ceux-ci contenaient assez d'informations sur les emplois pour qu'on puisse les évaluer selon le plan Willis?

R. Oui.

229. Le Comité mixte a compris, dès le début de l'étude, la nécessité de communiquer à des employés choisis l'importance de l'étude ainsi que l'importance de fournir des informations complètes sur les postes. C'est pourquoi il a créé un sous-comité des communications qu'il a chargé d'élaborer une stratégie de communication mettant l'accent sur la nécessité de fournir des renseignements complets et précis et de retourner promptement les questionnaires distribués à cette fin.

230. La stratégie de communication comportait les mesures suivantes : (i) l'ajout d'un feuillet au chèque de paye expliquant l'objet de l'étude sur la parité salariale et assurant aux employés que leur niveau de classification ne serait pas touché; (ii) l'envoi de lettres aux employés à qui on demandait de remplir un questionnaire; (iii) la préparation d'un vidéo à l'intention des employés désignés comme réviseurs et devant servir à la formation des titulaires appelés à remplir les questionnaires; et (iv) la préparation de documents de formation pour les coordonnateurs.

231. Le Comité mixte a assuré les fonctionnaires que leur participation n'aurait aucune conséquence négative sur leur carrière. Il leur a également garanti que les informations fournies ne serviraient pas à d'autres fins que celle de l'étude sur la parité salariale. Aux titulaires faisant partie des groupes professionnels à prédominance masculine il a dit que, si un écart salarial était trouvé, leur salaire ne serait pas touché.

232. Afin de contrer les problèmes éventuels associés aux questionnaires à questions ouvertes, M. Willis a mis en place un système d'autocontrôle, sous la forme de mesures de protection, afin d'assurer que les informations recueillies seraient complètes, définitives, exactes et à jour. Nous allons maintenant décrire ces mesures.

(iii) Coordonnateurs

233. M. Willis avait initialement proposé que ses consultants montrent aux titulaires comment remplir le questionnaire Willis, mais il a accepté la décision du Comité mixte de recourir à des coordonnateurs faisant office de formateurs, décision que -- vu les motifs de rentabilité invoqués -- lui a paru valable.

52

234. Les coordonnateurs avaient notamment les fonctions suivantes : montrer aux titulaires comment remplir les questionnaires, mener des séances d'information pour expliquer la nature et l'objet de l'étude, répondre aux questions des employés, distribuer et expliquer le questionnaire Willis, aider au besoin les employés à remplir le questionnaire et coordonner la collecte des données.

235. Les coordonnateurs ont été désignés comme nationaux ou régionaux selon leur mandat et leur lieu de travail. Le processus et les critères de sélection que l'Alliance a appliqués pour nommer les coordonnateurs ont été décrits en détail par un témoin de l'Alliance, Mme Elizabeth Millar, chef de la Classification et de l'égalité de rémunération, Direction de la négociation collective. De même, le processus de sélection de l'Institut a été décrit par Mme Kathryn Brookfield, chef de la Section de la recherche. L'employeur, par contre, n'a présenté aucune preuve à propos du processus et des critères qu'il avait utilisés afin de choisir ses employés pour cette fonction.

236. Des séances de formation pour les coordonnateurs ont eu lieu au cours des mois de septembre et d'octobre 1987. Des documents -- imprimés, diapositives et vidéos -- ont été remis aux coordonnateurs pour les aider à former les titulaires. Le programme de formation des coordonnateurs a duré

environ une demi-journée. D'autres exercices se sont greffés à la formation des coordonnateurs, notamment des séances où ceux-ci ont pu s'exercer à gagner l'appui des personnes qui pouvaient hésiter à remplir le questionnaire, à apprendre à faire face à des difficultés linguistiques et à faire les arrangements voulus pour retenir au besoin les services d'un interprète. Toute la formation relative au plan Willis a été donnée par un consultant de M. Willis.

237. M. Willis a exprimé l'avis suivant sur la qualité de la formation donnée aux coordonnateurs (volume 62, à la p. 7657, lignes 15-21) :

[TRADUCTION]

Q. Pour ce qui est de la formation, vous avez participé à la formation des coordonnateurs, vous ou vos consultants?

R. Oui.

Q. Et vous étiez satisfait de la formation qu'on a donnée aux coordonnateurs?

R. Oui.

238. Après la formation, on a assigné à chaque coordonnateur un certain nombre de titulaires à former. C'est le coordonnateur qui devait décider de la date à laquelle donner cette formation, bien que M. Willis souhaitât que les titulaires reçoivent leur formation le plus tôt possible. M. Willis a par ailleurs souligné aux coordonnateurs l'importance de faire remplir les questionnaires par les titulaires le plus tôt possible après qu'ils eurent reçu leur formation. Il préférait que les questionnaires soient remplis dans les deux semaines suivant la formation des employés.

Selon M. Willis, les titulaires auraient besoin de quatre à huit heures pour bien remplir le questionnaire.

239. A la suite de la formation des coordonnateurs qui a été achevée en octobre 1987, environ les deux tiers des questionnaires avaient été reçus en février 1988 et presque les trois quarts en mars 1988. Le sous-comité de l'administration établi par le Comité mixte a consacré beaucoup de temps à étudier le nombre de questionnaires reçus et les moyens d'obtenir tous les questionnaires restants. Le taux de retour final des questionnaires a été de 95 %. Quelques questionnaires ont continué de rentrer durant l'été et l'automne de 1988.

240. Dans son témoignage, Mme Brookfield a expliqué que de nombreux coordonnateurs de l'Institut ont commencé à donner de la formation peu de temps après avoir eux-mêmes été formés. La formation des employés a pris énormément de temps parce que les coordonnateurs avaient un grand nombre d'employés à former et que les employés n'étaient pas au même lieu de travail. A cause de ces facteurs, il a fallu étaler les séances de formation afin que les coordonnateurs puissent rencontrer les différents groupes d'employés. Mme Brookfield a en outre fait remarquer que certains des titulaires ne pouvaient se libérer de leur travail en même temps, autre facteur qui a eu pour effet d'allonger la période nécessaire à la formation.

241. Mme Brookfield a aussi exprimé le point de vue de l'Institut sur la qualité de la formation reçue par les coordonnateurs (volume 168, à la p. 21007, lignes 5-16) :

[TRADUCTION]

R. Ils m'ont dit franchement que plus ils en faisaient, plus ils avaient l'impression de s'améliorer, et qu'ils avaient reçu de l'«input» des séances de formation précédentes au sujet des questions que les employés auraient et qu'ils y répondraient à ce moment-là. Mais ensuite, après cela, ils pouvaient penser à plus d'informations ou à une autre façon de résoudre la question et ils intégraient cela dans leur séance de formation suivante, d'entrée de jeu peut-être; ou encore, s'il n'y avait pas de questions, ils pourraient soulever des possibilités ou des choses qu'ils avaient glanées au cours d'autres séances de formation.

242. L'Alliance avait beaucoup plus de coordonnateurs que l'Institut, soit environ une centaine. Mme Margaret Jaekl, agente de classification et d'égalité de rémunération à la Direction de la négociation collective, à l'Alliance, a fait le point sur l'efficacité de la formation reçue par les coordonnateurs à partir des commentaires que les coordonnateurs de l'Alliance lui avaient faits (volume 200, de la p. 25831, ligne 25, jusqu'à la p. 25833, ligne 8) :

[TRADUCTION]

Q. Avez-vous reçu du feedback des coordonnateurs sur la façon dont ils avaient l'impression d'être perçus, premièrement, par

l'employeur et, deuxièmement, par les personnes à qui ils montraient comment remplir le plan?

R. Oui. Nous avons eu des réunions périodiques avec l'ensemble de ce que nous appelions nos coordonnateurs nationaux. Chaque élément avait un coordonnateur national et ils avaient pour leur part de nombreux coordonnateurs régionaux aussi.

[...]

R. Le feedback que nous recevions en général, c'était qu'ils travaillaient bien avec leur homologue patronal. Les gens comprenaient leurs exposés. En général, les gens remplissaient leur questionnaire et le retournaient. Certaines personnes avaient des questions et, en général, ils [les coordonnateurs] étaient satisfaits de pouvoir leur répondre.

243. Le Comité mixte a demandé la coopération de l'employeur pour qu'il accorde aux employés le temps nécessaire à la formation. D'après la preuve non contredite, il y a eu une bonne coopération de la part de l'employeur, qui a accordé aux employés choisis suffisamment de temps durant les heures normales de travail afin qu'ils puissent assister à la séance de formation et remplir le questionnaire. Les titulaires ont bénéficié d'un congé payé pour remplir le questionnaire, ce qui, dans certains cas, a pu prendre jusqu'à huit heures.

244. Selon M. Willis, moins il s'écoule de temps entre le moment où les coordonnateurs reçoivent leur formation et le moment où ils forment eux-mêmes les titulaires, plus la formation est efficace. D'après son expérience, la qualité des questionnaires est meilleure lorsque les titulaires le remplissent peu de temps après avoir été formés. Plus il s'écoule de temps, moins bons sont les questionnaires. En l'occurrence, M. Willis n'a pu préciser quand la qualité avait commencé à décliner. Il a constaté que la qualité des informations variait d'un questionnaire à l'autre. Les questionnaires remplis au début étaient de meilleure qualité, a-t-il remarqué. Les employés de la Défense nationale ont retourné leurs questionnaires exactement dans le temps voulu. Ces employés ont rempli les questionnaires comme ils étaient censés le faire et leurs questionnaires étaient «excellents». M. Willis et ses consultants ont remarqué que la qualité «allait en diminuant» plus les employés tardaient à retourner leurs questionnaires.

245. Il y a peu de preuves au sujet des dates spécifiques auxquelles les séances de formation coordonnateurs-titulaires ont eu lieu. Une partie du retard est attribuable au fait que les superviseurs ont mis du temps à lire, à commenter et à signer les questionnaires de leurs employés.

Certains d'entre eux ont attendu que tous leurs employés aient rempli leurs questionnaires, puis ils les ont signés et retournés en bloc. M. Willis a admis qu'il était impossible de savoir si un employé avait, en fait, rempli le questionnaire dans les 10 ou 14 jours après l'avoir reçu ou plus tard. Il a signalé qu'il y avait eu au total 1 258 substitutions de titulaires et que cela représentait 837 questionnaires.

246. La preuve a montré que les informations provenant des fonctionnaires de sexe féminin sont rentrées plus tôt et étaient de meilleure qualité que les renseignements reçus des fonctionnaires de sexe masculin. De plus, les questionnaires des titulaires occupant des postes techniques et professionnels de haut niveau ont été retournés plus tard et contenaient des informations inférieures à ce que l'on a constaté dans les questionnaires provenant des employés de bureau et des gens de métier.

(iv) Réviseurs

247. Lorsque les questionnaires ont commencé à rentrer, M. Willis a demandé à l'une de ses consultantes, Mme Jan Drury, de choisir les meilleurs questionnaires qu'évaluerait le Comité directeur. Mme Drury a fait part à M. Willis de ses réserves sur la qualité générale des questionnaires. À la suite de ses observations, M. Willis a mis en place une procédure d'appoint pour qu'on puisse obtenir des renseignements additionnels. Ce travail a été confié à un groupe de travail de réviseurs nommés par le Comité mixte. La fonction première de ces réviseurs était d'examiner les questionnaires reçus pour vérifier s'ils comportaient des lacunes ou des contradictions quelconques.

248. Selon M. Willis, il est absolument nécessaire, dans le processus Willis, de procéder à un examen préalable des questionnaires reçus. Initialement, M. Willis avait recommandé que cette fonction soit confiée aux consultants. Normalement, il demande à ses consultants d'examiner les questionnaires remplis. Le Comité mixte a décidé de former des fonctionnaires fédéraux pour accomplir cette tâche. Cela a amené le Comité à prendre une autre décision à savoir s'il fallait «faire soi-même ou faire faire». Les réviseurs ont exécuté ce travail tout au long de l'étude.

249. Les réviseurs ont été formés par Mme Drury. Ils ont reçu une formation plus poussée que les coordonnateurs parce qu'ils devaient bien connaître le plan Willis pour pouvoir déterminer si les questionnaires avaient été bien remplis.

250. Ainsi, le côté patronal et le côté syndical ont chacun nommé des personnes qui feraient fonction de réviseurs. Environ 55 individus ont été nommés en cette qualité. On leur a notamment assigné certaines tâches techniques à l'égard de chaque questionnaire, par exemple, retirer toute allusion au genre et à la classification. Après avoir repéré les questionnaires nécessitant des informations additionnelles ou des éclaircissements, le réviseur devait rédiger les questions à poser aux titulaires afin de combler les lacunes. Il voyait en outre à obtenir d'autres renseignements concrets relativement à la terminologie technique contenue dans les questionnaires afin de présenter cette information en un langage que comprendrait plus facilement un comité d'évaluation.

251. Mme Drury a supervisé le travail des réviseurs jusqu'en mars 1988. Elle a examiné les questions et notes qu'ils avaient rédigées pour chaque examen des questionnaires évalués par le Comité directeur. Subséquemment, Mme Diane Saxberg, pour la partie syndicale, et M. Doug Edwards, pour la partie patronale, ont été nommés chefs réviseurs le 7 mars

56

1988. Les chefs réviseurs étaient chargés de réviser le brouillon des questions préparées par les réviseurs.

252. Les réviseurs interrogeaient les titulaires afin d'obtenir les informations manquantes. Un fort pourcentage de ces entrevues de suivi ont été faites par téléphone, un nombre limité d'entre elles, moins d'une douzaine, ayant été réalisées en personne. Dans certains cas, plusieurs appels téléphoniques étaient nécessaires, dont certains étaient très longs. Le réviseur rédigeait ensuite les réponses et les annexait au questionnaire avant de le transmettre à un comité d'évaluation. On appelait «notes du réviseur» les réponses ainsi rédigées.

253. M. Willis voulait que les réviseurs relèvent dans les questionnaires les points qui avaient pu échapper à l'attention du titulaire ou avoir été oubliés, ou qu'ils détectent les contradictions entre les réponses du titulaire et les commentaires de son superviseur. On avait en outre donné instruction aux réviseurs d'être alertes aux expressions d'opinion ou conclusions non étayées sur les faits.

254. Les réviseurs n'ont trouvé qu'une «poignée» de cas où il y avait désaccord entre le superviseur et le titulaire. Mme Saxberg a témoigné que, dans ces situations, après avoir parlé aux deux individus elle constatait que, dans la plupart des cas, c'était davantage sur la forme que sur le fond -- c'est-à-dire la nature des fonctions -- qu'ils ne s'entendaient pas.

255. M. Willis a expliqué que, d'après son expérience, environ la moitié des entrevues nécessaires pouvaient être menées par téléphone; dans les autres cas, il fallait rencontrer le titulaire en personne pour obtenir des informations plus concrètes, surtout lorsqu'il s'agissait d'emplois techniques et professionnels de haut niveau.

256. M. Willis a fait remarquer que le nombre de fois où il faille obtenir plus d'informations pour compléter un questionnaire, que ce soit dans 80 % ou 30 % des cas, n'a pas vraiment d'effet sur la qualité du questionnaire. A son avis, ce sont les informations supplémentaires obtenues et transmises au comité qui comptent.

257. La preuve a montré que certains réviseurs avaient précédemment agi à titre d'évaluateurs au sein de comités d'évaluation et avaient été repérés comme «marginiaux», c'est-à-dire qu'ils cotaient différemment des autres membres du comité. M. Willis a défini ces individus comme des personnes qui, dans un comité d'évaluation, divergeaient du reste du comité dans son ensemble et donnaient des cotes supérieures à certains genres d'emploi ou inférieures à certains autres genres d'emploi par comparaison avec leurs collègues membres du comité (volume 29, à la p. 3793).

258. M. Willis a expliqué qu'une façon de vérifier la validité dans les cas où le réviseur est aussi un évaluateur marginal consiste à examiner les questions qu'ils rédigent et les réponses qu'ils donnent, puis à déterminer si la réponse répond réellement à la question. Pour sa part, il n'a relevé aucun cas où ces individus n'auraient pas noté les réponses aux questions soulevées.

57

259. Le Tribunal a entendu le témoignage de trois personnes ayant fait fonction de réviseur. A propos de l'efficacité des entrevues téléphoniques que les réviseurs faisaient pour obtenir des informations, Mme Christine Netherton a déclaré ce qui suit (volume 173, à la p. 21919, lignes 5-20) :

[TRADUCTION]

R. [...] parfois cela prenait la moitié du temps pour obtenir l'information, mais très souvent il fallait expliquer le but de l'étude et ils disaient «Oh, j'ai rempli ça» ainsi de suite. Donc il y avait beaucoup de bavardage pour établir des relations cordiales avec eux. Et nous tâchions de ne pas pousser dans le dos des gens.

Je pense que, dans l'ensemble, nous avons reçu les informations voulues. Nous pouvions aussi compter sur les réponses des autres réviseurs.

Mais il arrivait que quelqu'un n'aimait pas parler. Je parle de l'impression que j'ai eue, moi. Je ne dis pas que c'était parfait à 100 %. Mais l'impression générale, c'est que dans la majorité des cas on obtenait de bonnes informations par téléphone.

260. Une autre réviseuse qui a témoigné, Mme Mary Crich, a déclaré qu'elle n'avait pas souvent relevé de contradictions entre les titulaires et les superviseurs dans les résumés des fonctions et responsabilités. De ses entrevues téléphoniques elle a conclu que les employés appréciaient avoir la chance de parler à quelqu'un de leur emploi.

261. MM. Willis et Durber ont tous deux été interrogés sur la compétence et l'habileté des réviseurs. M. Durber connaissait bien un certain nombre des réviseurs à cause de sa longue expérience dans la fonction publique fédérale. Il les a qualifiés d'«évaluateurs d'emplois et d'analystes professionnels». A son avis, ils étaient en général plus compétents pour accomplir les tâches qui leur étaient assignées en leur qualité de réviseurs que d'autres individus n'ayant pas leurs antécédents (volume 164, aux pp. 20505-7).

262. A propos du travail des réviseurs, M. Willis a dit ce qui suit dans son témoignage (volume 65, à la p. 8136, lignes 2-7) :

[TRADUCTION]

Q. Mais à votre connaissance les réviseurs ont-ils eu de la difficulté à comprendre leur travail après la formation?

R. Je ne pense pas qu'aucun d'entre eux n'ait éprouvé quelque difficulté que ce soit. Du moins on ne m'en a pas parlé.

263. Après le processus de révision, les questionnaires ainsi que les notes des réviseurs étaient transmis à un comité d'évaluation. Si un évaluateur du comité avait besoin de plus d'informations, le comité rédigeait des questions et les remettait au réviseur pour qu'il obtienne

les renseignements voulus du titulaire. Les renseignements obtenus par le réviseur étaient ensuite retournés par écrit au comité d'évaluation en question.

264. Sous la direction de M. Durber, la Commission a examiné les questionnaires en vue d'en déterminer la qualité. Au cours de l'audience, la Commission a produit un rapport intitulé «An Examination of the Quality of Questionnaire Information used by the Federal Equal Pay Study (pièce HR-245). Le rapport, qui examine la qualité des informations contenues dans les questionnaires, a été préparé par la Direction de la parité salariale de la Commission à la demande de M. Durber, qui a été chargé d'enquêter sur ces plaintes. Une chercheuse expérimentée détenant une maîtrise en études canadiennes de l'Université Carleton a été mandatée pour étudier un échantillon représentatif des évaluations. Cet échantillon comprenait 651 questionnaires, 63 portant sur des postes-repères et 588 sur des postes autres que les postes-repères. La tâche de la chercheuse consistait à vérifier si les informations recueillies au moyen des questionnaires dans le cadre de l'étude sur la parité salariale étaient complètes et exactes, en apparence tout au moins. La chercheuse a été supervisée étroitement par M. Durber qui, pour sa part, a étudié 36 dossiers marqués par celle-ci; M. Durber a conclu que chaque dossier était satisfaisant.

265. Dans son rapport, la chercheuse a signalé que la lisibilité des descriptions contenues dans les questionnaires était bonne et que les questions ouvertes avaient fourni la latitude voulue tant pour les groupes à prédominance masculine que pour ceux à majorité féminine. De nombreux titulaires ont apporté plus de précisions sur leurs fonctions en ajoutant des pages à cette partie du questionnaire.

266. Dans le rapport de la Commission, il est par ailleurs mentionné que plus de 99 % des questionnaires portaient la signature du superviseur et que, dans 96 % des cas, ce dernier avait inscrit des commentaires. Des informations contradictoires ont été relevées dans environ 9 % des questionnaires. Dans les cas où les superviseurs avaient fourni des renseignements contradictoires, il a été possible de concilier les données pour 95 % des questionnaires au moyen des entrevues subséquentes menées par les réviseurs.

267. M. Durber a exprimé ses propres attentes concernant la qualité des informations contenues dans les questionnaires (volume 158, de la p. 19761, ligne 23, jusqu'à la p. 19762, ligne 3) :

[TRADUCTION]

Je peux seulement dire que, d'après mon expérience dans la fonction publique -- et juste pour situer mes attentes dans un contexte --, ce que j'ai vu était de loin supérieur à ce que j'avais vu dans les descriptions de postes et les dossiers des

emplois, ainsi que dans les présentations et même dans les situations de griefs.

59

268. Lorsqu'il a contre-interrogé M. Willis, l'avocat de l'intimé lui a demandé si les mesures de protection ou d'appoint mises en place par le Comité mixte pour contrer certains problèmes liés à la collecte des données avaient atteint l'objectif qu'il souhaitait. M. Willis a répondu ce qui suit (volume 78, de la p. 9543, ligne 3, jusqu'à la p. 9546, ligne 1) :

[TRADUCTION]

Q. Ces mesures de protection -- et elles sont toutes décrites dans vos propositions initiales -- avaient trait tant à la collecte d'informations qu'à l'évaluation, n'est-ce pas?

R. Oui.

Q. Ai-je raison de dire que la presque totalité sinon la totalité des mesures qui ont été mises en place -- et il y en a eu beaucoup -- n'ont pas été efficaces pour atteindre ce que vous souhaitiez?

R. Je pense qu'on peut dire que l'efficacité des mesures a varié.

Q. Et le degré d'efficacité, n'est-ce pas, a été décevant tout au mieux?

R. Oui.

Q. Cela tient en partie au fait que lorsque l'on considère les informations qui ont été fournies aux comités -- les cinq premiers et les neuf subséquents -- une fois les mesures d'appoint appliquées, les données étaient plus faibles que ç'aurait dû être le cas. Etes-vous d'accord?

R. Je ne suis pas sûr de ce que vous voulez dire quand vous dites «plus faibles que ç'aurait dû être le cas».

Q. Plus faibles que ce qu'il aurait été souhaitable d'obtenir pour une bonne évaluation.

R. J'étais d'avis -- et je l'ai fait savoir au Comité mixte -- que la qualité des informations n'était pas aussi élevée que je l'aurais souhaité. Quoi qu'il en soit, j'estimais que dans l'ensemble les informations recueillies étaient satisfaisantes pour nos besoins.

Q. Je comprends, mais vous nous avez également dit que c'était plus faible que ce que vous obteniez normalement dans d'autres études.

R. Oui.

Q. C'est que, à cause des lacunes dans les mesures de protection nous devons nous interroger, en quelque sorte, au sujet

60

de certaines des informations effectivement recueillies et qui sont là, parce qu'il est ressorti de nos discussions que les renseignements n'avaient pas été écrits par un évaluateur d'emplois expérimenté et que certaines inscriptions avaient été faites par des évaluateurs marginaux; toute cette discussion que nous avons eue... Etes-vous de mon avis?

R. Etes-vous en train de dire que certaines des informations n'étaient peut-être pas exactes?

Q. Je ne dis pas qu'elles étaient inexactes. Nous ignorons si elles étaient exactes ou non. Notre niveau de confiance à l'égard des informations est inférieur à ce que nous voudrions parce que les informations, jusqu'à un certain point, ont été écrites par des gens qui n'ont pas l'expérience de ce genre de rédaction, qu'elles ont été révisées par des gens inexpérimentés en révision, que les entrevues ont été menées par des personnes qui ne sont pas des analystes de postes professionnels. C'est cela que je veux dire.

R. Je pense avoir dit au Comité mixte patronal-syndical, ou au mini-comité mixte tout au moins, que nous nous attendions à trouver une disparité plus grande parce que les informations étaient un peu plus faibles que ce que nous souhaitions.

Q. Là où je veux en venir, finalement, c'est que vous dites bien «faibles». Je vous demande si, comme moi, vous estimez que

même avec ce que vous avons, nous ne pouvons avoir qu'un niveau de confiance réduit en ce qui concerne l'exactitude de ces données.

R. Je ne sais pas si j'irais jusque là. Bien sûr, nous aurions souhaité avoir des questionnaires plus complets et contenant davantage de renseignements concrets, des faits. Ce sont des choses qui amènent toujours les évaluateurs à supposer certaines choses, d'où le risque que les disparités soient plus grandes.

269. En dépit du fait que, dans les autres études auxquelles il avait pris part, il avait pu compter sur une information de meilleure qualité, M. Willis a maintenu tout au long de l'audience que, en l'occurrence, la qualité des informations était suffisante pour les besoins de l'étude.

270. Lorsqu'il a été contre-interrogé par l'avocat de l'intimé, M. Willis a de nouveau donné son opinion sur la qualité des informations (volume 69, de la p. 8612, ligne 22, jusqu'à la p. 8615, ligne 15) :

[TRADUCTION]

Q. Alors ce que vous dites, c'est que même après toutes les mesures d'appoint qui ont été appliquées, les informations laissaient encore à désirer dans une large mesure?

R. Je dirais que les informations étaient tout au mieux satisfaisantes, mais non supérieures.

61

Q. Avez-vous en tête une échelle quelconque -- vous me faites penser à nos évaluations de rendement. On peut avoir la cote «satisfaisant», «entièrement satisfaisant» et «supérieur». Est-ce l'échelle que vous employez?

R. Disons que j'avais des réserves au sujet de la qualité. Des entrevues téléphoniques et des entrevues menées par des personnes sans formation professionnelle ne sauraient remplacer complètement un questionnaire bien rempli au départ. Dans l'ensemble je dirais que la qualité était suffisante pour nos besoins, surtout si on considère le grand nombre d'évaluations, mais si nous avions eu affaire, par exemple, aux Services de toxicomanie qui ne comptent que 19 ou 20 postes, j'aurais été très inquiet parce que je sais que nous aurions dû tolérer une disparité plus grande que ce que j'aurais souhaité avoir comme consultant. Je me suis donc dit, comme je l'ai expliqué, que

pourvu que la disparité soit aléatoire et qu'elle s'annule donc au bout du compte, je pourrais vivre avec les résultats.

Qu'arrive-t-il lorsque vous avez de fait un questionnaire -- deux choses arrivent quand vous avez un questionnaire qui est un peu faible : (1) le processus s'en trouve ralenti, comme nous avons pu le constater; et (2) vous devez prévoir qu'il y aura une plus grande tolérance face à la disparité.

Q. D'accord. Je veux vous arrêter là parce que vous avez dit quelque chose encore que je mets en doute.

Vous dites que la disparité s'annule. C'est-à-dire si vous recherchez de la partialité fondée sur le sexe.

R. Si c'est aléatoire, par définition ça s'annule. S'il se dégage une tendance, alors ce n'est pas aléatoire.

Q. Je vous dirais que ce qui arrive, si c'est aléatoire, c'est que la partialité fondée sur le sexe s'annule.

R. Si c'est aléatoire, toute partialité est annulée.

Q. Toute partialité, d'accord. Mais ce qui n'est pas annulé, c'est la non-fiabilité. Si la disparité est grande, alors le niveau de fiabilité est plus faible. Je pensais qu'on s'était mis d'accord là-dessus hier.

R. Je pense qu'un statisticien vous dirait que si vous aviez affaire à un nombre de cas relativement restreint, cela serait très vrai. C'est moins vrai si le nombre d'évaluations augmente, et si la disparité continue d'être aléatoire -- autrement dit, les plus et les moins ont tendance à s'annuler -- vous pouvez quand même obtenir une fiabilité satisfaisante avec un grand nombre d'évaluations.

271. Au cours du contre-interrogatoire, M. Willis a par ailleurs réitéré à l'avocat de l'intimé les propos qu'il avait déjà tenus dans son témoignage (volume 78, à la p. 9566, lignes 19-22) :

[TRADUCTION]

Je pense que j'ai déjà dit que j'estimais, compte tenu de tout le travail que nous avons effectué concernant la collecte des données, que les données étaient assez bonnes pour les besoins de cette étude.

272. Plus loin, M. Willis précise ce qui suit (de la p. 9567, ligne 23, jusqu'à la p. 9568, ligne 14) :

[TRADUCTION]

Maintenant je vous demande, Monsieur Willis, que faudrait-il enlever pour que vous disiez finalement : «Je ne défendrai pas cette étude»?

R. Premièrement, j'ai plusieurs fois affirmé que la qualité des informations était assez bonne. J'aurais tiré la sonnette d'alarme si j'avais cru que la qualité avait été faible au point où nous n'aurions pu nous fier aux résultats.

Deuxièmement, j'ai exprimé la conviction que je ne pourrais valider les résultats de l'étude si nous ne pouvions effectuer un examen interne des évaluations existantes... ce qui a été fait.

Mais ce qu'il faudrait? Il est possible qu'en prenant connaissance de l'analyse finale je convienne que nous ne pouvons utiliser les résultats, mais ça je ne le sais pas.

### C. LE PROCESSUS D'ÉVALUATION

273. Dans une étude d'envergure telle celle menée par le Comité mixte, qui porte sur un nombre considérable de postes, M. Willis utilise des comités d'évaluation multiples. Il préfère pouvoir se limiter à un comité, mais face à une myriade d'emplois, on doit pouvoir compter sur plusieurs comités pour effectuer une évaluation efficace et appropriée. En tout, 16 comités ont été établis pour évaluer les questionnaires.

#### (i) Comité directeur

274. Le défi que doit relever M. Willis est de concevoir un processus qui permette aux divers comités d'être constants les uns par rapport aux autres sur une période relativement longue. Pour guider les travaux des comités multiples et à titre de mesure de protection, M. Willis crée un comité directeur de l'évaluation. Selon lui, il est nécessaire et essentiel dans une étude sur la parité salariale d'effectuer des comparaisons entre des emplois dissemblables. Le Comité directeur a comme

fonction première d'établir les relations entre les différents emplois et de dresser le cadre de référence pour les comités d'évaluation multiples,

bref de définir ce qu'il est convenu d'appeler la «discipline» du Comité directeur.

275. On désigne par l'expression postes-repères les postes qui ont été évalués par le Comité directeur. Au total, le Comité directeur a évalué 501 postes. L'évaluation de postes-repères revêt une importance critique lorsqu'on fait appel à des comités multiples.

276. Le Comité directeur était composé de 10 membres, la moitié étant des représentants patronaux, l'autre moitié des représentants syndicaux. Un représentant de l'employeur et un représentant des syndicats ont été désignés comme coprésidents. M. Willis n'a pas choisi les membres du Comité directeur; cette tâche a été laissée à la discrétion des parties. Il a toutefois recommandé que les membres puissent concevoir leur travail à l'échelle de l'administration fédérale, qu'ils possèdent des compétences analytiques et conceptuelles, qu'ils soient déterminés à mener à terme une tâche difficile et qu'ils soient capables de mettre de côté leur affiliation syndicale ou patronale afin de pouvoir adopter une approche équilibrée vis-à-vis des évaluations. Les parties ont tâché de former le Comité directeur pour qu'il reflète cet équilibre.

277. M. Willis a témoigné que la composition du Comité directeur était bien équilibrée quant au nombre d'hommes et de femmes qui en faisait partie et quant à la diversité de leurs antécédents. Le Comité directeur était aussi formé d'un nombre égal de représentants syndicaux et patronaux.

278. Selon M. Willis, la réussite de l'évaluation d'emploi tient à l'uniformité avec laquelle on interprète les facteurs d'évaluation d'un comité à l'autre. Il utilise trois méthodes pour vérifier cette uniformité, lesquelles ont toutes été employées dans l'étude sur la parité salariale. La première méthode est utilisée dans les situations où un consultant assiste un comité d'évaluation ou fait fonction de conseiller. Ici le consultant évalue le même emploi en utilisant les mêmes informations que les membres du comité, tout en recherchant des tendances qui diffèrent de l'évaluation indépendante qu'il a effectuée. La deuxième méthode consiste à comparer chaque évaluateur au comité dans son ensemble. Enfin, la troisième méthode consiste à comparer les comités entre eux. Les tests de fiabilité inter-évaluateurs (FIE) et les tests de fiabilité inter-comités (FIC) seront décrits et examinés plus en détail dans la section suivante.

279. Les évaluations des postes-repères fournissent un large cadre de référence pour les comités d'évaluation et servent à assurer l'uniformité; on les utilise en quelque sorte pour contrôler la qualité dans le processus d'évaluation. Plus particulièrement, c'est le terme «discipline» qu'on emploie lorsqu'on veut qualifier l'attitude plus ou moins libérale ou conservatrice avec laquelle le Comité directeur interprète les termes employés dans les évaluations.

280. Les consultants doivent veiller à ce que les comités d'évaluation appliquent une discipline uniforme. La discipline adoptée par le Comité directeur impose une lourde responsabilité aux comités d'évaluation multiples, en ce sens que ces derniers, dans leurs évaluations, doivent

64

respecter la ligne de conduite dictée par le Comité directeur dans ses propres évaluations. Autrement dit, si le Comité directeur évalue un facteur donné d'une certaine façon, c'est ainsi que les autres comités doivent l'évaluer.

281. M. Willis a expliqué que s'il était loisible aux comités d'évaluation multiples de créer leur propre discipline, il s'ensuivrait des évaluations qui manqueraient d'uniformité. Il pourrait y avoir uniformité au sein du comité même, c'est-à-dire que ce dernier pourrait traiter tous les postes d'une manière juste et équitable, mais le degré de libéralisme caractérisant leur interprétation des termes pourrait varier. Si le Comité directeur évalue un facteur d'une certaine façon, les comités d'évaluation multiples doivent s'en tenir à celle-ci, sinon il en découle une surestimation ou une sous-estimation des questionnaires. Voici ce que dit M. Willis à ce propos (volume 60, de la p. 7396, ligne 18, jusqu'à la p. 7397, ligne 4) :

[TRADUCTION]

Chaque comité d'évaluation adopte ce que j'ai appelé une discipline, qui se traduit par le conservatisme ou le libéralisme avec lequel on traite les facteurs d'évaluation. Une fois cette discipline établie, si une évaluation est supérieure ou si l'emploi est évalué de façon plus libérale que ce qu'indiquent les autres évaluations, je dirais qu'il y a eu surestimation. Si l'évaluation était plus conservatrice que ce à quoi je me serais attendu comparativement à l'uniformité générale du comité, alors je dirais qu'il y a eu sous-estimation.

282. M. Willis estimait qu'il était très important que le Comité directeur fournisse une base d'évaluation solide pouvant servir de cadre de référence aux autres comités. En ce qui concerne la qualité des questionnaires utilisés par le Comité directeur -- si on la compare à celle des questionnaires utilisés par les autres comités --, M. Willis a affirmé que la qualité des premiers était supérieure à celle des seconds.

283. M. Willis a demandé à Mme Drury de choisir pour le Comité directeur des postes-repères représentatifs de la profondeur et de l'ampleur de l'organisation. Le Comité mixte a formellement approuvé les critères de sélection à sa réunion du 10 juillet 1987. Ces critères précisent que les postes-repères doivent être représentatifs de tous les groupes professionnels, des différents niveaux organisationnels, des emplois à forts contingents, des emplois standard et des groupes professionnels à prédominance masculine et féminine qui existent dans l'échantillon qui sert en fait de population pour l'étude globale. De plus, on devait tâcher de bien échantillonner les postes spécialisés et d'éviter le plus possible de retenir les niveaux consécutifs d'une famille d'emplois donnée.

284. M. Willis a par ailleurs donné à Mme Drury un autre critère pour la sélection des postes-repères, à savoir choisir des questionnaires de la plus haute qualité. Dans ce contexte, un questionnaire de qualité, selon M. Willis, est un questionnaire complet qui contient des renseignements

définitifs et concrets. M. Willis estimait qu'il était très important que le Comité directeur dispose de questionnaires de la plus haute qualité.

285. Le Comité directeur n'a pu s'offrir le luxe de recevoir tous les questionnaires avant de choisir ceux qui allaient servir de postes-repères. Le Comité mixte a donné instruction à M. Willis d'amorcer le travail du Comité directeur dès que les 50 premiers questionnaires auraient été reçus. En fait, certains questionnaires rentraient toujours après que le Comité directeur eut terminé son travail. Si M. Willis était dans l'ensemble satisfait que le Comité directeur avait fourni un bon cadre de référence, il n'était cependant pas en mesure de dire si chacun des critères approuvés par le Comité mixte pour la sélection des questionnaires par le Comité directeur avait été respecté lors du choix des postes-repères.

286. Au début du travail du Comité directeur, M. Willis agissait à titre de président du comité. Après un certain temps, il a abandonné ce rôle aux coprésidents du Comité directeur, qui se sont relayés toutes les semaines. Le rôle du président était d'animer la réunion, de maintenir une

position neutre de façon à ne pas influencer le groupe, d'écrire les évaluations au tableau et de diriger le groupe dans l'établissement d'un consensus. M. Willis a passé un certain temps avec les coprésidents pour leur montrer ce qu'il faisait et pourquoi il faisait certaines choses. Il a fallu environ trois semaines aux coprésidents avant qu'ils assument cette tâche. A partir de ce moment-là, M. Willis s'est assis à l'arrière de la salle en tant qu'observateur et on faisait appel à lui de temps à autre lorsqu'il s'agissait d'interpréter certains points. Il a aussi agi à titre de facilitateur dans les séances de «récapitulation» ou d'«examen provisoire» (une autre partie du processus qui sera expliquée plus tard). Sa participation a été telle jusqu'à la fin des travaux du Comité directeur. Chaque fois qu'arrivait le moment d'une séance d'examen, il assumait l'animation du groupe.

287. Après que le Comité directeur eut terminé son travail, M. Willis a proposé, par souci d'efficacité, qu'une partie des postes-repères du Comité directeur soient désignés comme des «postes-repères primaires». A mesure que les comités d'évaluation supplémentaires entrèrent en scène, ils avaient besoin de consulter les postes-repères du Comité directeur. Au lieu de mettre l'ensemble des évaluations des postes-repères à la disposition de chaque évaluateur, il a décidé de désigner des postes-repères primaires dont les évaluations seraient fournies à tous les évaluateurs. Toutefois, chaque comité d'évaluation a pu recevoir la totalité des postes-repères.

288. La sélection des postes-repères primaires reposait principalement sur la fréquence d'utilisation prévue et sur d'autres facteurs tels que le niveau organisationnel, le groupe professionnel et les facteurs les plus représentatifs des emplois évalués. A la demande de M. Willis, chacun des membres du Comité directeur a produit une liste de postes-repères que M. Willis a raffiné; au bout du compte, une centaine de postes-repères primaires avaient été identifiés.

#### (ii) Comités d'évaluation multiples

66

289. Chacun des comités d'évaluation restants comptait sept membres représentant à parts égales les parties syndicale et patronale. Un des membres faisait fonction de président syndical ou patronal du comité. Ici encore, M. Willis a laissé la sélection des membres à la discrétion des parties. Le Tribunal a entendu des témoins de l'Alliance et de l'Institut qui lui ont signalé qu'on avait pris soin de choisir des individus s'exprimant bien, ayant un bon sens de l'analyse, capables de défendre les évaluations et désireux de travailler en équipe. Du point de vue de

l'équilibre entre les sexes, l'Alliance a tenté en vain de recruter des nombres égaux d'hommes et de femmes. Leurs évaluatrices, toutefois, faisaient souvent partie de groupes professionnels à prédominance masculine.

290. M. Willis estime que, du point de vue de la perception, il est important qu'un comité soit composé d'hommes et de femmes. Si un comité ne compte que des femmes, a-t-il fait remarquer, on pourrait avoir l'impression que l'étude est axée sur les femmes, l'inverse étant aussi vrai s'il était exclusivement formé d'hommes. D'après l'expérience de M. Willis, toutefois, si un comité regroupe des gens «capables», leur sexe est sans importance. Selon lui, les antécédents de l'évaluateur comptent plus que son sexe.

291. M. Willis avait recommandé qu'aucun spécialiste en classification de la fonction publique fédérale ne fasse partie des comités d'évaluation. Il n'en a toutefois pas été ainsi. Sept évaluateurs nommés par l'employeur avaient une connaissance approfondie du système de classification dans la fonction publique fédérale. Ils ont fait partie de quatre comités d'évaluation ainsi que du Comité directeur. Ce qui préoccupait M. Willis au sujet des individus ayant une expérience de la classification, c'est que ceux-ci amorcent leur travail d'évaluation avec une «idée préconçue» (il a employé le terme anglais «baggage»). Selon lui, une personne sans expérience aucune peut davantage être objective que quelqu'un qui arrive avec des années d'expérience en classification.

292. Cette «idée préconçue», d'après M. Willis, tient au fait que l'individu peut déjà avoir une certaine connaissance et compréhension de la valeur relative des emplois au sein d'une organisation. Ces idées préconçues prennent notamment la forme de suppositions concernant le travail et sont probablement inconscientes. À ses yeux, il s'agit d'un parti pris reposant sur une information incomplète et pouvant amener l'individu à avoir des intentions cachées.

293. Personne n'est exempt d'idées préconçues d'une forme ou d'une autre, selon M. Willis. On peut, cependant, les neutraliser au maximum en ayant l'esprit ouvert et en faisant preuve d'objectivité et d'équité dans tous les cas, de manière à ne pas influencer indûment une évaluation.

294. Chaque comité d'évaluation (les cinq premiers comme les neuf subséquents) a en tout temps consisté en sept membres; dans bien des cas, cependant, il a fallu remplacer des membres. Le Tribunal a entendu le témoignage direct de 17 évaluateurs.

295. Un des 17 évaluateurs, Mme Christine Netherton, qui faisait partie de la première version du comité no 1 (il a continué de fonctionner après que le Comité directeur eut terminé son travail), a témoigné au sujet des idées préconçues. Une membre de son comité avait de l'expérience en classification. Selon Mme Netherton, celle-ci avait de la difficulté à apprécier les autres points de vue à cause de ses antécédents en classification. Lorsque ce genre de problème surgissait, le comité tentait d'en discuter avec la personne. En cas d'impasse, il obtenait l'aide d'un consultant.

296. M. Willis a également constaté ce problème chez les évaluateurs du comité no 3. La première version de ce comité a éprouvé de nombreux problèmes. Certains de ces problèmes étaient attribuables au fait que quelques-uns des évaluateurs de la partie patronale avaient des antécédents en classification. Du côté syndical, il y avait des évaluateurs qui étaient résolus à augmenter les cotes des groupes à prédominance féminine plus qu'il n'était justifié. Selon M. Willis, ce comité se trouvait presque constamment «paralysé ou en situation d'impasse». Voici en quels termes M. Willis a décrit la situation (volume 57, de la p. 7090, ligne 12, jusqu'à la p. 7093, ligne 20) :

[TRADUCTION]

R. Le [comité] numéro 3 comptait des individus qui, du côté syndical, étaient déterminés à coter le plus haut possible les postes des titulaires appartenant aux groupes à majorité féminine, tandis que, du côté patronal, il y avait deux ou trois évaluateurs qui étaient d'anciens spécialistes de la classification qui semblaient résolus à les contrôler le plus possible. C'était presque une situation d'impasse.

La présidente du comité no 3 était une représentante syndicale et, en dépit du fait que nous avons soigneusement expliqué aux présidents l'importance de rester neutres -- autrement dit, le président, pour sa propre crédibilité et afin de ne pas exercer d'influence indue, devait surveiller attentivement sa façon de diriger ou d'animer le groupe --, cette présidente en particulier est presque devenue une quatrième évaluatrice syndicale. Pas qu'elle a produit comme tel des évaluations, mais elle a pris part aux discussions de manière à faire pencher la balance du côté syndical plutôt que de rester neutre.

Bien sûr, la présidente avait la possibilité d'amener le groupe à un consensus et de passer à autre chose. Or elle ne passait pas à autre chose avant d'être certaine que son côté était bien représenté. C'était un cas où j'estimais qu'il fallait à tout

prix retirer la présidente, et c'est ce que j'ai recommandé au Comité mixte.

Q. Et quelle suite a-t-on donnée à votre recommandation?

R. Aucune. Le côté patronal a appuyé ma recommandation, mais le côté syndical a refusé d'y souscrire.

68

Q. Et quel a été le résultat de cette situation? Comment, à votre avis, cela a-t-il affecté le processus d'évaluation au sein du comité no 3?

R. Fait intéressant, les deux tendances semblaient s'équilibrer. Un des évaluateurs patronaux était un ancien gestionnaire de la classification qui avait beaucoup d'influence. La plupart du temps, l'affrontement se soldait par une impasse.

Le problème, c'est qu'ils étaient lents et qu'ils évaluaient de plus en plus lentement. Je m'attendais à ce que les comités produisent de huit à neuf évaluations par jour, mais aucun d'entre eux n'a réussi à le faire. Ce comité particulier n'arrivait pas à évaluer plus de deux ou trois postes par jour, ce qui frustrait d'ailleurs beaucoup les membres. J'estimais donc que la situation était néfaste non seulement à la productivité du groupe, mais aussi à la santé et au bien-être des membres eux-mêmes.

Q. Vous avez mentionné deux conséquences de cette situation, à savoir l'effet sur la santé des membres et la faible productivité du comité. Que pensez-vous en fait des évaluations effectuées par ce comité?

R. Je ne peux pas dire que nous avons relevé de parti pris quelconque finalement. Je suis sûr que nous aurions constaté une tendance quelconque si les positions n'avaient pas été partagées à raison de trois d'un côté et trois de l'autre. Je pense que les évaluations, pour autant que nous ayons pu le déterminer, étaient correctes.

Q. Qu'est-il arrivé au comité no 3 finalement?

R. Lorsque nous avons passé de cinq à neuf comités, nous avons pu retirer la présidente et la placer comme membre votante dans un

autre comité, un des neuf comités. Fait plutôt étonnant, son attitude a semblé changer du tout au tout à ce moment-là.

Q. Vous voulez-dire que son attitude s'est améliorée?

R. De l'avis du consultant qui siégeait à ce comité et de l'avis du président du comité, elle se conduisait de façon plus consciencieuse qu'auparavant.

Q. En ce qui concerne les évaluations qu'a produites le comité dont elle a fait partie subséquentement, avez-vous une opinion à exprimer à ce sujet?

R. Je n'ai constaté aucun problème, ni concernant cette personne ni concernant le comité comme tel.

297. Dans le contexte d'une étude sur la parité salariale, on obtient les meilleures évaluations de comités véritablement hétérogènes. Le profil idéal d'un comité d'évaluation des emplois dans ce genre de contexte est

69

d'avoir des individus ayant des antécédents différents et des expériences différentes, de compter un nombre à peu près égal d'hommes et de femmes, d'avoir des individus provenant de syndicats différents et exerçant des fonctions différentes au sein de ministères différents et à des niveaux divers. Ce que M. Willis recherche, ce sont des personnes qui peuvent évaluer de façon consciencieuse et équitable, une tâche difficile.

298. Selon M. Willis, la partialité peut découler de l'utilisation de systèmes d'évaluation des emplois, pas nécessairement du plan d'évaluation mais de la part de l'évaluateur. C'est d'ailleurs pourquoi il estime que le processus proprement dit est plus important que l'instrument d'évaluation. Un comité hétérogène ne peut garantir qu'il ne se glissera pas de parti pris dans un processus d'évaluation des emplois; toutefois, dans ce genre de comité on a une meilleure chance d'obtenir un résultat objectif. Selon M. Willis,

[TRADUCTION]

[...] nous avons besoin de gens consciencieux, capables d'analyser une situation et sur qui on peut compter pour qu'ils tâchent de faire ce qui doit être fait, par opposition à protéger leur propre secteur ou domaine.

(volume 29, à la p. 3788, lignes 18-22)

299. Le Tribunal a entendu des témoignages sur les antécédents, l'âge, les compétences, le poste, les forces et les faiblesses des membres des comités d'évaluation multiples. A quelques exceptions près, la preuve confirme qu'en général les critères d'un comité équilibré définis par M. Willis ont été respectés. Par contre, la preuve montre clairement que certains évaluateurs sont arrivés avec une opinion préconçue ou un parti pris, et qu'on ne pouvait compter sur eux pour évaluer les emplois objectivement. Dans la mesure où ces évaluateurs ont pu influencer sur la fiabilité des résultats, nous examinerons d'autres mesures de protection utilisées par M. Willis dans le processus d'évaluation afin de déterminer comment bien le processus a fonctionné.

(iii) Processus d'évaluation des questionnaires

300. M. Willis a décrit la différence entre ce que l'on appelle couramment une évaluation d'emplois traditionnelle et une évaluation d'emplois à des fins de parité salariale. Il a expliqué qu'on utilisait l'approche traditionnelle depuis le début des années 1940 surtout pour évaluer les emplois de gestion. Celle-ci a pour objet d'établir une base pour appliquer des différences salariales entre les divers niveaux de gestionnaires. Par opposition, la parité salariale, d'expliquer M. Willis, nécessite que l'on compare des emplois dissemblables à tous les niveaux d'une organisation et sur le marché.

301. Les deux approches reposent sur des méthodes très différentes bien que, dans chaque cas, on fasse appel à des comités d'évaluation. En général, pour l'approche traditionnelle, les comités d'évaluation sont composés de gestionnaires, les informations sont recueillies à l'aide de

descriptions de travail et les entrevues sont menées par des consultants. Le rôle du consultant devient plus discret une fois que le comité a reçu la formation nécessaire. Pour reprendre les propres mots de M. Willis, «la courbe d'apprentissage s'accroît de façon assez dramatique».

302. Lorsque l'évaluation s'effectue dans un contexte de parité salariale, M. Willis préfère des comités : (1) qui sont «équilibrés», c'est-à-dire composés d'un nombre égal d'hommes et de femmes; (2) qui sont représentatifs des divers niveaux organisationnels; et (3) au sein desquels on trouve des gens ayant des antécédents divers.

303. M. Willis a par ailleurs expliqué qu'un comité chargé d'évaluer les emplois à des fins de parité salariale doit avoir reçu une formation pour savoir examiner un questionnaire et analyser l'importance d'un emploi. Durant ce processus, les membres doivent mettre de côté leurs idées personnelles sur la façon dont les emplois ont tendance à se relier les uns aux autres. Les problèmes auxquelles l'évaluateur fait face dans une situation de parité salariale sont différents de ceux qu'il doit régler dans une évaluation d'emplois traditionnelle. Les problèmes qui surgissent dans le premier cas découlent principalement des «idées que les gens se font des rapports qui existent entre les postes».

304. D'après M. Willis, si les évaluateurs sont à l'aise face à une évaluation d'emplois traditionnelle, c'est qu'ils comprennent en général bien les emplois en cause, par exemple, lorsqu'il s'agit d'un groupe de gestionnaires qui évaluent les postes au sein de leur propre organisation. Ainsi, les «idées que les gens se font des rapports qui existent entre les postes» deviennent moins importantes dans ce contexte que lorsqu'il s'agit d'évaluer les emplois en fonction de la parité salariale, les consultants devant alors former les évaluateurs à jeter un coup d'oeil tout à fait différent sur les choses.

305. Le processus Willis exige que les évaluateurs s'acquittent de leur tâche de façon indépendante. Il est d'ailleurs prévu une procédure particulière qui doit être suivie durant les évaluations. On peut la décrire comme suit : chaque membre du comité lit le questionnaire par lui-même; ensuite les évaluateurs discutent des informations et soulèvent des questions sur le contenu de l'emploi, ce qui constitue, selon M. Willis, la dernière étape de la collecte des données; au stade de la discussion, M. Willis permet aux membres du comité de partager toute connaissance concrète qu'ils peuvent avoir de la nature du travail accompli et du contexte dans lequel il est exécuté; si un évaluateur ou un comité a besoin d'informations supplémentaires, c'est le moment de rédiger des questions qu'on transmettra à un réviseur; lorsque les membres du comité ont une connaissance commune des faits, chaque évaluateur est tenu de coter -- de façon indépendante et en secret -- chaque facteur du poste en question; subséquemment, le consultant ou le président recueille les bordereaux d'évaluation des évaluateurs et les transfère sur un tableau noir, ce qui donne un aperçu visuel à partir duquel le comité peut faire des comparaisons.

306. Il s'ensuit une période de discussion au cours de laquelle les évaluateurs débattent de leurs divergences; si quelqu'un a attribué une

cote légèrement différente pour un facteur donné, on lui demande de justifier sa cote à partir de faits concrets. M. Willis s'attend à ce que les autres membres du comité d'évaluation écoutent les raisons des évaluateurs minoritaires; il dénomme cette partie du processus «processus consensuel». Chaque évaluateur peut à ce stade-ci rajuster sa cotation des facteurs, mais uniquement s'il peut étayer le rajustement sur des faits concrets. La cote à laquelle le comité est arrivé par consensus est ensuite enregistrée, puis l'on rédige -- à l'aide des critères figurant dans le plan d'évaluation et illustrés par les postes-repères -- une justification raisonnée qui explique essentiellement ce qui a motivé l'évaluation de chaque poste.

307. Bien que M. Willis ait dit au Comité mixte que seul le Comité directeur avait à produire une justification raisonnée, le Comité mixte a décidé que les comités d'évaluation multiples devaient en rédiger également. Des problèmes ont surgi à propos de ces justifications parce que certaines étaient mal rédigées et difficiles à comprendre. Par conséquent, il y a eu des retards dans leur transcription. Par le passé, M. Willis n'a pas utilisé de justifications raisonnées parce qu'il ne considère pas qu'elles sont essentielles au processus d'évaluation, bien que, a-t-il reconnu, elles puissent être utiles. Il a conseillé aux membres des comités d'évaluation multiples d'utiliser les justifications comme guide mais, dans tous les cas, il voulait que les évaluateurs retournent au questionnaire du Comité directeur et qu'ils le lisent, plutôt que se fier à une justification. A son avis, il était impossible pour un évaluateur de saisir tout ce qu'il avait à savoir afin d'évaluer un poste dans une justification raisonnée d'une ou deux pages.

308. Soit avant d'en arriver à un consensus ou après y être arrivé, selon la préférence de chaque comité, ce dernier examine les postes-repères du Comité directeur et choisit des emplois semblables ou dissemblables afin de s'assurer que ses cotes concordent avec les postes-repères du Comité directeur. S'il n'y a pas concordance, le comité rajuste son évaluation en conséquence.

309. M. Willis a expliqué qu'il devient assez évident, aux yeux des consultants surtout, lorsqu'un individu manifeste une préférence pour un sexe ou pour l'autre durant ce processus puisqu'il est difficile pour l'individu de fournir des informations concrètes à l'appui d'une préférence fondée sur des sentiments. M. Willis n'exige pas qu'il y ait unanimité pour un consensus, uniquement un accord des deux tiers des membres du comité d'évaluation. Tout évaluateur faisant partie de l'autre tiers a la possibilité de convaincre le groupe que sa cote est la bonne. Comme l'a affirmé M. Willis, «[...] dans le consensus final, nous avons au moins les deux tiers des gens qui estiment que l'évaluation est correcte» (volume 38, à la p. 4737).

310. Les comités d'évaluation ont en général suivi le processus conçu par M. Willis, c'est-à-dire que chacun lisait le questionnaire de façon indépendante, que les membres discutaient entre eux des faits pour mieux les comprendre, que chaque membre cotait individuellement chacun des sous-facteurs, que l'on affichait les cotes individuelles sur un tableau

72

noir, qu'on arrivait à un consensus et, enfin, qu'on choisissait les postes-repères appropriés du Comité directeur.

311. M. Willis a témoigné qu'un bon comité possédant de bonnes informations sur les emplois peut habituellement évaluer entre 8 et 10 questionnaires par jour. A partir de cette estimation, le Comité mixte a initialement établi cinq comités d'évaluation multiples, en tenant pour acquis que chaque comité pourrait évaluer environ 750 postes; toutefois, la productivité du Comité directeur et des comités d'évaluation multiples a été beaucoup plus faible que prévue au départ. Par conséquent, en vue de faire face à la question du retard et de régler d'autres problèmes, M. Willis a recommandé que les comités d'évaluation soient restructurés et que leur nombre soit porté de cinq à neuf, ce que le Comité mixte a approuvé le 3 mars 1989.

312. Un grand nombre des problèmes observés par M. Willis et ses consultants étaient liés aux cinq comités d'évaluation de départ. Nous expliquerons maintenant en détail les circonstances entourant ces problèmes.

73

#### (iv) Formation des comités d'évaluation multiples

313. Les évaluateurs avaient besoin de formation sur l'utilisation du plan Willis. M. Willis a lui-même donné cette formation au Comité directeur en octobre 1987. Il a témoigné qu'il était satisfait de la formation du Comité directeur (volume 62, à la p. 7698).

314. Ce sont M. Willis et ses consultants qui ont formé les cinq premiers comités d'évaluation. M. Willis a rencontré les cinq comités d'évaluation le premier jour, puis ensuite il a réparti les membres en comités d'évaluation et leur a assigné à chacun un consultant. Lorsque le nombre des comités est passé de cinq à neuf, tous les nouveaux membres ont reçu une formation individuelle ou, s'il s'agissait d'un comité entièrement nouveau, la formation a été donnée à tout le comité.

315. Un des buts que recherche M. Willis lorsqu'il forme un comité est de faire en sorte que chacun puisse appliquer le plan Willis en toute confiance. Sa formation consiste habituellement en explications sur le processus Willis et en formation sur le tas avec ses consultants jusqu'à ce que les évaluateurs se soient bien familiarisés avec l'utilisation du plan. L'approche de M. Willis est surtout basée sur les expériences pratiques. La formation s'étend habituellement sur une période de deux semaines; vers la fin de la première journée ou peut-être le deuxième jour, M. Willis distribue un questionnaire et demande au groupe d'effectuer une évaluation. Il met les instructeurs en garde contre les hypothèses ou suppositions qu'ils pourraient faire à propos du travail, les incitant plutôt à rechercher des faits en remplissant le questionnaire.

316. M. Willis forme ses propres consultants à l'utilisation du processus Willis. Une partie de leur formation est orientée vers la question des attitudes reliées au travail stéréotypé. Cette partie de la formation, tant avec les consultants que les comités, est informelle. La perspective qu'il veut faire adopter est d'ignorer si un emploi est à prédominance masculine ou féminine. Il discute d'attitudes avec ses consultants et leur enseigne des moyens d'y faire face, notamment en examinant les éléments et composantes de l'emploi, en subdivisant le poste en un certain nombre de composantes et en examinant chaque composante sans égard au sexe du titulaire. A leur tour, ses consultants enseignent la même méthode lorsqu'ils forment les comités.

317. On a demandé M. Willis de commenter une publication de la Commission de l'équité salariale de l'Ontario (pièce PSAC-71), qui a été établie pour aider à la mise en oeuvre et à l'administration de la Loi sur l'équité salariale (Ontario). La publication contenait des renseignements sur la formation des comités d'évaluation d'emplois. M. Willis approuvait en principe la liste des éléments qui, selon la Commission, devaient faire partie de la formation d'un comité d'évaluation, à savoir : l'historique de l'évaluation des emplois; la façon dont on fixait les traitements et salaires par le passé; les processus liés à l'équité salariale et à la détermination des salaires; la façon dont la partialité fondée sur le sexe peut entrer dans les systèmes d'évaluation; les tendances concernant la participation des femmes à la population active; la justification de

l'équité salariale; et les mécanismes spécifiques du système utilisé par l'organisation en question.

318. Cependant, M. Willis préfère son approche qui, depuis une vingtaine d'années, s'est révélée plus pragmatique que la liste détaillée

de critères préconisée par la Commission de l'équité salariale de l'Ontario. Il a fait les observations suivantes à ce propos (volume 209, de la p. 27088, ligne 23, jusqu'à la p. 27089, ligne 16) :

[TRADUCTION]

R. Nous avons appris par expérience que la meilleure façon de faire face aux différences dans divers genres d'emplois -- et, en passant, des emplois exclusivement féminins ou des emplois exclusivement masculins, ça n'existe plus aujourd'hui. On trouve des hommes et des femmes dans tous les emplois et il y a toutes sortes d'emplois.

Il y a certaines caractéristiques des emplois masculins et des emplois féminins qui sont en quelque sorte cachées et il existe une variété d'emplois telle, en particulier dans le secteur public, que notre expérience nous a appris que la meilleure façon de faire face à la situation est d'éviter entièrement de nous concentrer sur le travail des hommes par opposition au travail des femmes et, plutôt, de chercher à décomposer le travail en facteurs pour ensuite examiner ces facteurs sans nous demander s'il s'agit de l'emploi d'un homme ou d'une femme, tout en nous assurant de bien faire ressortir tous les éléments cachés, quels qu'ils soient.

319. M. Willis n'a pas acquis sa propre formation sur les questions liées à l'égalité des hommes et des femmes dans un programme officiel, mais plutôt «à l'école de l'adversité» (volume 209, à la p. 27168).

320. En répondant à d'autres questions de l'avocat de la plaignante, en contre-interrogatoire, M. Willis a reconnu que la sensibilisation aux questions touchant l'égalité des sexes n'était «pas une mauvaise chose». En réponse au témoignage sur ce sujet des experts Weiner et Armstrong, qui préconisent le genre de formation recommandée par la Commission de l'équité salariale de l'Ontario, M. Willis, s'appuyant sur sa propre expérience, a affirmé qu'une sensibilisation de ce genre pouvait être utile, mais que ce n'était pas absolument nécessaire. Il confirme sa pensée à ce propos notamment dans le passage suivant de son témoignage (volume 209, à la p. 27096, lignes 9-17) :

[TRADUCTION]

Q. Alors vous êtes convaincu que vous pouvez mener une étude sur la parité salariale et effectuer des évaluations d'emplois équitables sans le genre de formation proposé par les spécialistes Armstrong et Weiner?

R. Je dirais que nous avons acquis amplement d'expérience dans l'évaluation d'emplois masculins et féminins sans jamais avoir reçu comme tel de formation en la matière; l'orientation fournie par le consultant est suffisante.

321. Le mécanisme ou le moyen que M. Willis utilise pour assurer des résultats solides et fiables en l'absence d'une formation en règle sur les questions touchant l'égalité des sexes consiste dans la participation du consultant. Sauf pendant une période de trois semaines, M. Willis a personnellement observé le travail du Comité directeur. Pendant son absence de trois semaines, il a été remplacé par sa consultante, Mme Drury.

322. Cinq consultants, dont M. Willis, ont travaillé à l'étude sur la parité salariale. Lorsque les cinq premiers comités d'évaluation ont commencé leurs évaluations, chaque comité avait un consultant désigné pour la formation et la consultation.

323. M. Willis a témoigné que le rôle du consultant est d'évaluer les postes en privé pendant que le comité d'évaluation fait ses propres évaluations. Il a fait remarquer qu'initialement le Comité directeur aurait une courte période pour discuter de l'emploi particulier choisi pour fins d'évaluation. Comme il avait coutume de le faire, M. Willis avait sa propre liste de questions auxquelles il fallait répondre par rapport à un questionnaire donné. Si les membres du Comité directeur ne faisaient pas ressortir ces informations, M. Willis soulevait lui-même les questions pertinentes. Dans les comités d'évaluation multiples établis par la suite, ce sont ses consultants qui ont rempli ce rôle.

324. M. Willis a résumé le rôle du consultant comme étant, en général, de faire fonction d'animateur ou de facilitateur, de dispenser la formation nécessaire et de répondre aux questions du comité sur les techniques d'évaluation. Celui-ci ne cesse par ailleurs jamais d'observer le fonctionnement du comité et d'en prendre le pouls. Lorsque les consultants évaluent eux-mêmes l'emploi, ils ne communiquent pas au comité le résultat de leur évaluation. L'objet de ces évaluations est de permettre au consultant de suivre de près les évaluations des comités. De l'avis de M. Willis, les consultants ont deux avantages par rapport aux comités : (i) ils sont des évaluateurs professionnels; et (ii) ils n'ont pas d'idées préconçues ni de parti pris.

325. Selon M. Willis, un des inconvénients du rôle du consultant tient au fait que, n'étant pas du milieu, celui-ci ne connaît pas l'environnement de l'organisation aussi bien que les membres du comité d'évaluation et

ignore donc comment les différences sont perçues au sein de l'organisation. Par ailleurs, d'ajouter M. Willis, le consultant risque toujours d'être influencé par sa connaissance d'emplois semblables dans d'autres organisations, emplois qui peuvent en fait être différents de ceux visés par l'étude.

326. Le consultant examine non seulement les faits qu'invoque chaque évaluateur pour justifier sa propre évaluation, mais il examine aussi le travail du comité, vérifiant -- et c'est là le point le plus important -- comment ce dernier évalue chacun des sous-facteurs du plan

76

Willis. Durant ce processus, les consultants exercent ce que M. Willis décrit comme un «jugement empirique».

327. M. Willis a déclaré que le Comité directeur était un bon comité et un comité efficace. S'appuyant sur ses propres observations et sur l'information reçue de Mme Drury, il était satisfait du degré d'uniformité que le Comité directeur avait manifesté par rapport à sa propre discipline. S'il avait à se prononcer sur la qualité générale de son travail, il dirait que le Comité évaluait les postes en se basant sur les faits.

328. A partir de ses propres observations, M. Willis a identifié deux individus qui semblaient être des évaluateurs marginaux ou minoritaires. Le Comité directeur n'a pas en général été influencé par ces individus. M. Willis a tiré ses conclusions sur la façon dont les autres membres du Comité directeur recevaient les commentaires de ces évaluateurs et y réagissaient en s'appuyant sur les raisons invoquées par les autres évaluateurs dans leurs évaluations et sur le consensus général du groupe, l'influence des évaluateurs marginaux n'étant apparente dans ni l'un ni l'autre des cas.

329. En ce qui concerne la formation des comités d'évaluation multiples, M. Willis a constaté qu'après cinq jours de formation la majorité des membres étaient à peine à l'aise avec le système, mais qu'ils l'étaient devenus progressivement au bout de deux semaines. Les évaluateurs qui ont témoigné à l'audience ont eux aussi affirmé avoir été de plus en plus à l'aise avec le plan Willis à mesure que leur travail progressait.

330. M. Willis a reconnu qu'il fallait faire preuve d'une vigilance constante pour maintenir et comprendre le plan afin que les évaluateurs ne retournent pas à des jugements antérieurs dans leurs évaluations. C'est pourquoi M. Willis a régulièrement rencontré les cinq comités d'évaluation

initiaux en vue d'examiner les problèmes et de proposer des solutions. De plus, les consultants de M. Willis ont rédigé des notes consultatives techniques dans lesquelles ils répondaient aux questions posées par les comités concernant l'interprétation des aspects techniques du plan Willis.

(v) Évaluations du Comité directeur

331. M. Willis a présenté au Tribunal ses conclusions concernant le travail effectué par le Comité directeur durant l'étude sur la parité salariale. A maintes reprises il a affirmé que les évaluations du Comité directeur étaient impartiales, qu'il n'avait rien à redire sur le travail du Comité directeur, que les informations utilisées par celui-ci reposaient sur les faits et que, à son avis, le Comité directeur avait accompli un excellent travail.

332. A plusieurs reprises pendant l'étude sur la parité salariale, on a demandé à M. Willis de se prononcer sur la qualité des cotes attribuées par le Comité directeur. En réponse à l'avocat de l'intimé, il a fait les observations suivantes (volume 75, à la p. 9202, lignes 14-23):

[TRADUCTION]

77

J'ai probablement examiné «à mort» ces 503 évaluations et les différences. C'est sur une période de six mois qu'on m'a constamment posé des questions à leur sujet. Chaque fois que j'ai revu les évaluations moi-même ou avec d'autres consultants, nous sommes toujours arrivés à la même opinion, c'est-à-dire que s'il y avait certaines différences, les postes-repères du Comité directeur étaient satisfaisants, reconnaissant que nous n'avons pas affaire à une science exacte, mais plutôt à un art.

333. L'approche que M. Willis a adoptée pour valider les résultats des évaluations consistait à réévaluer lui-même des questionnaires choisis ou à le faire faire par un de ses consultants. La première vérification des évaluations du Comité directeur a été effectuée par un des consultants de M. Willis, en l'occurrence Mme Drury, au printemps de 1988. M. Willis a expliqué que l'examen de Mme Drury n'avait pas pour objet de valider les résultats du travail accompli par le Comité directeur, mais de vérifier si les évaluateurs du Comité savaient comment utiliser le plan Willis et s'ils le comprenaient et l'interprétaient comme il faut.

334. M. Willis voulait savoir si les membres du Comité directeur évaluaient les postes de façon uniforme. Les évaluateurs eux-mêmes

souhaitaient une telle vérification pendant qu'ils apprenaient encore le processus d'évaluation. On n'a pas dit au Tribunal exactement combien de questionnaires du Comité directeur Mme Drury avait examinés. Cette dernière a effectué son examen au printemps de 1988, et le Comité directeur avait amorcé son travail d'évaluation à l'automne de 1987.

335. Pour les besoins de son examen, Mme Drury a utilisé uniquement les questionnaires qui avaient été évalués pendant son absence des discussions du Comité directeur. Au bout du compte, elle a relevé 12 postes pour lesquels elle avait produit une évaluation légèrement différente de celle du Comité directeur. Les écarts mis en évidence par Mme Drury ne concernaient que des postes à prédominance féminine et les membres féminins du Comité directeur ont contesté ses conclusions et écrit une lettre en guise de protestation. Cette lettre était adressée à M. Willis et à la Commission. M. Willis croit que la controverse ne tenait pas tant au fait que Mme Drury avait critiqué les évaluations du Comité directeur qu'au fait qu'elle avait semblé s'en prendre aux postes féminins.

336. M. Willis a revu les 12 évaluations relevées par Mme Drury. Il a conclu que les évaluations de cette dernière étaient «bien fondées», avis qu'il a communiqué à la Commission. Au total, pour trois des douze questionnaires l'écart entre les cotes de Mme Drury et celles du Comité directeur était inférieur à 2,5 %. Mme Drury a ensuite rencontré de nouveau le Comité, qui a révisé ses cotes à la lumière de ses réévaluations; sept des douze évaluations en question sont toutefois demeurées inchangées. Sur les sept, Mme Drury estimait que le Comité directeur avait sous-estimé la valeur du poste dans deux cas et qu'il l'avait surestimée dans les cinq autres cas. M. Willis ne se préoccupait pas du petit nombre de différences en soi, il s'inquiétait plutôt du fait que sur les sept évaluations divergentes, cinq allaient dans un sens et deux dans l'autre. Ce fait l'a amené à envisager la possibilité d'une partialité systématique.

337. M. Willis a de nouveau rencontré Mme Drury et le Comité directeur. Quatre des cinq postes qui, selon Mme Drury, avaient été surévalués étaient des postes en sciences infirmières. Après avoir discuté du contenu de ces postes avec le Comité directeur, M. Willis a conclu que les évaluations du Comité étaient satisfaisantes et il a appuyé les évaluations initiales. En discutant plus tard de cette situation avec M. Willis, Mme Drury a admis que son expérience passée concernant des postes en sciences infirmières avait été acquise dans l'État du Connecticut et que cela avait pu teinter ses évaluations. M. Willis a soutenu qu'au Connecticut on n'exigeait pas des membres de la profession infirmière des

connaissances aussi approfondies que celles qu'on exige au Canada des infirmières et infirmiers. Donc, selon M. Willis, l'évaluation de Mme Drury était probablement «un peu à côté».

338. En dernière analyse, M. Willis ne croyait pas que le Comité directeur avait réagi de façon excessive et systématique aux réévaluations de Mme Drury. Il a exprimé ce point de vue dans une lettre datée du 5 décembre 1988 qu'il a adressée à la coprésidente patronale du Comité mixte, Mme Lise Ouimet, et dans laquelle il dit notamment ce qui suit :

[TRADUCTION]

Au printemps de 1988, nous avons répondu à une demande du Comité directeur d'examiner et de commenter les évaluations qu'il avait effectuées jusque-là. Mme Jan Drury a examiné le travail du Comité et a fait des recommandations concernant douze évaluations. Dans quatre de ces évaluations, la cote globale a été rajustée de 10,0 % à 10,9 %; dans quatre autres évaluations, le rajustement a été de 11,0 % à 15 %; enfin, une évaluation a été rajustée par plus de 15 %.

Le groupe a examiné ses évaluations et ses explications, tant écrites que verbales, et il a changé ses évaluations sur deux des neuf postes (y compris celui pour lequel il y avait un écart d'un peu plus de 15 %); il restait donc sept questionnaires dont l'évaluation différait de celle de Mme Jan Drury par une marge de 10 % à 15 %. Sur ces sept questionnaires, l'évaluation de Mme Drury était supérieure à celle du Comité pour deux postes et inférieure pour cinq postes. Cinq de ces sept postes sont des postes en sciences infirmières.

Les membres du Comité directeur, d'après les commentaires qu'ils ont faits, estiment qu'il y a une légère différence entre les fonctions qu'exerce le personnel infirmier de la fonction publique du Canada affecté à des services spécialisés et le personnel infirmier des États-Unis, qui est celui que connaît Mme Drury. Par exemple, le Comité directeur a accordé plus de poids que Mme Drury au counseling que les infirmières et infirmiers de l'unité des délinquants sexuels (poste 83) donnaient à ces derniers.

Je ne rejeterais pas entièrement l'avis du Comité directeur sur ce point sans de plus amples informations et j'estime que ces

faibles différences ne méritent pas qu'on s'y attarde. Par contre, si vous n'êtes pas d'accord, je suggère qu'on soumette ces postes infirmiers au Comité directeur qui, après avoir obtenu des informations additionnelles concernant la signification des affectations spéciales, pourra les réévaluer.

(pièce R-35)

339. Un autre procédé de protection ou d'appoint que M. Willis a conçu pour faire face aux divergences d'opinions entre les comités d'évaluation multiples et le Comité directeur consistait à permettre aux comités d'évaluation multiples -- voire à les y encourager -- de soumettre leurs divergences au Comité mixte avec des explications. M. Willis avait proposé dans son plan que le Comité directeur soit reconvoqué pour examiner ces écarts, et soit qu'il explique ses évaluations d'une façon plus détaillée ou qu'il rajuste ses évaluations afin qu'elles concordent avec les résultats des comités d'évaluation multiples. Il estime qu'il s'agit là d'une démarche très importante.

340. M. Willis a expliqué qu'un comité, après qu'il a évalué un certain nombre de postes, acquiert une certaine confiance dans sa capacité d'évaluer et qu'il y a inévitablement des différences mineures dans la façon dont un emploi est perçu. Pour remédier à la situation, a-t-il expliqué, on pourrait par exemple dire aux comités d'évaluation qu'ils doivent adopter la discipline du Comité directeur, quels que soient les désaccords éventuels. Mais M. Willis souhaitait une communication plus ouverte. Si les comités d'évaluation multiples avaient des réserves à propos d'une évaluation du Comité directeur, M. Willis était d'avis qu'ils avaient l'obligation et le droit de signaler ces différences, et que le Comité directeur devrait se pencher sur les cas «litigieux» que lui soumettraient les comités d'évaluation.

341. A la suite de ce qui précède, les comités d'évaluation multiples ont soumis 48 cas où ils contestaient les évaluations du Comité directeur. Les membres du Comité mixte n'arrivaient pas à s'entendre sur l'opportunité de reconvoquer le Comité directeur pour qu'il réexamine ces évaluations. A un moment donné, on a demandé aux consultants de réviser de façon indépendante environ 33 rajustements proposés par les comités d'évaluation. M. Willis a témoigné dans les deux tiers des cas. Les écarts étaient tellement minimes qu'il ne valait presque pas la peine de les prendre en considération, et ces cas ont été discutés individuellement avec les comités d'évaluation. M. Willis estimait qu'il restait environ 14 évaluations de questionnaires à faire réviser par le Comité directeur.

342. M. Willis ne voulait pas dire quel changement devait être apporté parce qu'il ne voulait pas se substituer au Comité directeur. Il a relevé

14 questionnaires qui lui paraissaient «problématiques» ou à tout le moins douteux et justifiant un réexamen.

343. M. Willis tenait fortement à ce que l'on reconvoque le Comité directeur pour concilier les divergences d'interprétation entre celui-ci et les comités d'évaluation multiples. Le Comité mixte a finalement décidé que le Comité directeur ne serait pas reconvoqué, préférant créer une

80

version réduite de ce dernier (le «mini-comité directeur»). Le mini-comité directeur était composé d'un nombre restreint d'anciens membres du Comité directeur. Ils étaient trois : M. Willis, Mme Joanne Labine, une représentante syndicale, et M. Michel Cloutier, un représentant patronal. Ces deux derniers membres avaient été identifiés comme des évaluateurs «marginaux» au sein du Comité directeur. En effet, ces deux membres, Mme Labine et M. Cloutier, avaient été identifiés par M. Willis dans ses observations directes du Comité directeur, conclusion qu'un statisticien indépendant avait par la suite confirmée au moyen d'une analyse statistique. Il s'agit là d'une des décisions incompréhensibles qu'a prises le Comité mixte.

344. Comme on pouvait s'y attendre, M. Willis a remis en question la décision du Comité mixte de choisir ces deux évaluateurs marginaux et a suggéré qu'on choisisse deux autres personnes. Selon M. Willis, on lui a «opposé une fin de non-recevoir» et il n'avait pas le pouvoir de renverser la décision du Comité mixte. M. Willis estimait que ces deux individus étaient mal préparés pour représenter le Comité mixte parce qu'ils avaient fait preuve de partialité fondée sur le sexe dans leurs évaluations initiales.

345. Le mini-comité directeur a examiné les modifications que le consultant avait proposé qu'on apporte aux postes-repères remis en question. La représentante syndicale a souscrit au point de vue du consultant et le représentant patronal a rejeté toutes les recommandations de M. Willis.

346. Le mini-comité directeur a ensuite fait les trois propositions suivantes au Comité mixte sans avoir consulté M. Willis :

[TRADUCTION]

OPTION 1

Il est proposé que le mini-comité directeur examine les changements que les consultants recommandent qu'on apporte aux postes-repères (33) du Comité directeur.

- Si les deux membres du mini-comité directeur en conviennent, les postes-repères et les justifications raisonnées seront modifiés.
- Si les membres du mini-comité directeur, après avoir consulté N.D. Willis ou Jane [sic] Drury, ne parviennent pas à s'entendre, alors le mini-comité directeur et le consultant détermineront s'il est dans le meilleur intérêt de l'étude de retirer le ou les poste(s)-repère(s) en question.

## OPTION 2

Il est proposé que les postes-repères remis en question ne soient pas modifiés. Dans les cas où le mini-comité

81

directeur convient que la cote est non convergente, on devrait retirer le ou les poste(s)-repère(s) en question.

Cette proposition repose en partie sur notre opinion selon laquelle il est trop tard pour tenter de modifier les postes-repères et demander aux comités de rajuster leurs habitudes de cotation.

Il y a lieu de signaler que deux options ci-dessus nécessiteraient qu'on tienne de nouvelles séances de récapitulation dans les cas où la modification ou l'enlèvement d'un poste-repère du Comité directeur s'ensuivraient.

## OPTION 3

Il est proposé que, si un comité ne peut en arriver à un consensus sur une cote, le questionnaire soit renvoyé au mini-comité directeur, qui tranchera.

Il est en outre recommandé qu'on n'accepte plus de remises en question des postes-repères du Comité directeur.

347. Le Comité mixte a retenu l'option 2. Comme on pouvait s'y attendre, le mini-comité directeur n'a pu s'entendre sur les postes-repères dont les cotes manquaient de cohérence et, par conséquent, aucun des postes-repères n'a été retiré. Bien que déçu, M. Willis ne pensait pas que cela portait atteinte à l'intégrité du processus. A ce stade-ci, il estimait que les évaluations étaient «intactes» et raisonnables.

348. M. Wisner a effectué les réévaluations des postes-repères contestés et proposé une augmentation moyenne générale de 4,2 % pour les 14 postes-repères visés. L'objet de la révision effectuée par le consultant était de déterminer si les écarts traduisaient une partialité ou un biais. L'analyse exécutée par M. Willis n'a révélé aucune partialité, et ce dernier estimait qu'il pouvait vivre avec la décision du Comité mixte de ne pas rappeler le Comité directeur. M. Willis ne croyait pas que les écarts globaux entre les consultants et le comité que cette analyse avait mis en lumière influaient de façon importante sur l'étude.

349. M. Willis a témoigné que l'analyse de M. Wisner illustre, en pourcentage, l'écart entre ce dernier et l'approche du Comité directeur. Tout au long de l'étude, a-t-il ajouté, les consultants ont cherché à éviter d'imposer leurs évaluations au Comité directeur. A la question de savoir quand il aurait imposé l'évaluation du consultant à un comité, M. Willis a répondu ce qui suit (volume 57, de la p. 7053, ligne 14, jusqu'à la p. 7055, ligne 10) :

[TRADUCTION]

LA PRÉSIDENTE : Peut-être pouvez-vous nous dire : A partir de quel moment recommandez-vous fortement quelque chose ou insistez-vous fortement sur une suggestion ou conseillez-vous que l'on prenne telle ou telle mesure? Qu'est-ce qui fait qu'à un moment donné

82

vous dites au comité «Il faut faire ceci ou cela», ou vous arrive-t-il en fait jamais de lui faire une telle recommandation?

LE TÉMOIN : Oui. Premièrement, le consultant qui siège avec le comité est au courant des informations contenues dans le questionnaire et de la discussion qui a porté sur le poste. Si nous devons constater à ce moment-là que les membres ne discutaient pas des faits et qu'ils ne les utilisaient manifestement pas de façon juste et équitable, nous soulèverions la question avec le comité d'évaluation lui-même durant ses travaux.

Par contre, après coup, en examinant une série d'évaluations, nous pourrions diverger légèrement d'opinion avec le comité, mais nous chercherions alors à savoir s'il ne s'agissait pas d'une honnête divergence d'interprétation de la part du comité. L'écart pourrait s'expliquer par ce que le consultant sait à propos de ce genre de poste -- et nous pouvons nous aussi, en tant que consultants, faire légèrement fausse route -- par opposition aux connaissances qu'en possède le comité, qui peut être mieux placé pour comprendre le contenu des postes faisant partie de son organisation.

Si nous devons déceler l'émergence d'une tendance, là nous prendrions des mesures très fortes. Puisqu'il s'agit là de jugements de valeur, il nous faut toutefois faire preuve d'une certaine tolérance. Un écart de 5 ou 6 % au niveau de la cote globale ne signifie pas forcément que le comité est dans l'erreur de 5 ou 6 %. Ce que nous voudrions savoir ici, c'est plutôt : Y a-t-il une tendance ici ou la différence est-elle aléatoire? Si l'écart est aléatoire, alors il n'y a aucune raison de s'inquiéter, sauf s'il est possible que les membres ne sachent pas comment utiliser l'instrument comme il faut.

LA PRÉSIDENTE : Pourquoi recommandiez-vous fortement que le Comité directeur se réunisse de nouveau?

LE TÉMOIN : Juste à cause de la psychologie des comités... ils insistaient beaucoup sur cela. Même s'il pouvait y avoir une très légère différence pour un poste donné, ils n'ont pas l'esprit tranquille s'ils n'ont pas pu au moins faire valoir leur point de vue.

350. Selon M. Willis, à la suite de la décision du Comité mixte de ne pas reconvoquer le Comité directeur, les évaluateurs des comités d'évaluation multiples ont éprouvé beaucoup de frustration. Les consultants ont été obligés de leur dire que le Comité mixte avait pris une décision de principe et qu'aucune modification ne serait apportée aux postes-repères à la suite des évaluations qu'ils avaient contestées. M. Willis a suggéré aux comités de chercher à contourner le problème. Il pense que de nombreux comités d'évaluation étaient enclins à choisir d'autres évaluations du Comité directeur pour les comparer à leurs propres

évaluations, ignorant les évaluations du Comité directeur qui aient été remises en question.

351. M. Willis a fait remarquer que la frustration était à son comble lorsque les comités d'évaluation s'attendaient à ce que le Comité directeur se réunisse de nouveau. Après qu'on leur a dit que le Comité ne serait pas reconvoqué, ils sont devenus plus résignés.

(vi) Évaluations des comités d'évaluation multiples

352. Certains des cinq premiers comités d'évaluation avaient tendance à négocier plutôt qu'à coopérer pour en arriver à un consensus. M. Willis a constaté que les comités d'évaluation s'équilibraient en général assez bien les uns les autres, mais cela se traduisait manifestement par une baisse de la productivité. Parmi les cinq comités qui ont été formés initialement, le comité no 3 était davantage enclin à se disputer et moins productif que les autres. M. Willis a formé les membres de ce comité et les a dirigés durant les trois premières semaines de leurs travaux. Il a rencontré sa présidente toutes les semaines en vue d'essayer de travailler avec elle, car il considérait qu'elle faisait partie du problème. Il a passé énormément de temps avec ce comité, travaillant avec lui et surveillant ses évaluations. Il connaissait mieux ce comité et ses problèmes que tout autre comité. M. Willis a remarqué que le comité no 3 avait des individus du côté syndical qui semblaient résolus à coter les emplois des groupes à prédominance féminine le plus élevé possible et deux ou trois membres du côté patronal, dont certains avaient des antécédents en classification, qui étaient déterminés à garder le plus possible les évaluations à l'intérieur de certaines limites. Il a qualifié cette situation d'«impasse».

353. M. Willis a témoigné que la présidente du comité no 3, une représentante syndicale, assumait parfois le rôle d'évaluatrice et participait aux discussions d'une façon qu'il jugeait inappropriée pour la présidente. Le rôle de la présidente est de rester neutre et d'animer les discussions du comité. M. Willis a subséquemment recommandé au Comité mixte de retirer la présidente du comité no 3 et il a fini par recommander la dissolution du comité. Le Comité mixte a toutefois rejeté ses recommandations et rien ne s'est produit qui était de nature à améliorer la situation du comité no 3 jusqu'à ce qu'on réforme les comités et qu'on en porte le nombre à neuf.

354. M. Willis a décrit un comité d'évaluation qui fonctionnait bien comme une équipe travaillant ensemble et dont chacun des membres s'efforçait d'évaluer de manière juste et équitable. Les réserves qu'il avait à l'endroit du comité no 3 ne concernaient pas les cotes attribuées aux postes, mais la façon dont le comité évaluait les postes. Ce comité débattait de ses divergences jusqu'à ce qu'il se mette d'accord sur les cotes par suite d'épuisement.

355. M. Willis ne croyait pas que la situation d'impasse dans laquelle se plaçaient les évaluateurs patronaux et syndicaux du comité no 3 influait négativement sur le processus d'évaluation dans ce comité. Ni lui ni ses consultants n'ont pu détecter de partialité systématique dans les

84

évaluations du comité no 3. Cependant, à cause de cette situation certains membres ont éprouvé des troubles de santé et la productivité du groupe a beaucoup souffert.

356. Le comité no 4, dans sa version première, a aussi donné du trouble à M. Willis. Ce comité a d'abord été excellent, jusqu'à environ mars 1989 en fait. Toutefois, dans les derniers stades, à cause des remplacements et de la réorganisation du comité, des problèmes ont surgi. En avril 1989, M. Willis a demandé au comité no 4 d'effectuer une dernière séance de récapitulation. Pendant cette séance, la présidente du comité est venue le trouver, presque en larmes. Selon M. Willis, elle lui a dit : [TRADUCTION] «Je ne peux plus le prendre. Tout s'est effondré, tout le monde est sur les nerfs, tout le monde se crie après. Nous avons un travail à faire et je démissionne.» Dans le procès-verbal de la réunion du Comité mixte tenue le 31 octobre 1989 (pièce R-44), M. Willis fait remarquer ce qui suit à propos du rapport du consultant sur le comité no 4 : [TRADUCTION] «[Le] problème principal concernant le comité no 4, c'est qu'il a manqué d'objectivité, ce qui a entraîné comme conséquence désastreuse la création de deux camps ainsi que d'ordres du jour séparés et de points de vue arbitraires et divergents.»

357. A ce moment-là, le comité avait évalué 52 postes. M. Willis a ensuite demandé aux membres restants du comité d'exprimer par écrit leurs préoccupations individuelles concernant les évaluations et de proposer les changements qu'ils estimaient nécessaires. Puis il a dissolu le comité. Par la suite, un consultant de M. Willis, M. Robert Barbeau, a examiné les préoccupations exprimées et a fait des recommandations, après quoi on lui a demandé de prendre des mesures appropriées. Les membres du comité ont fait des suggestions sur un total de 25 emplois et il n'y en avait qu'un seul à propos duquel le consultant différait sensiblement d'opinion par rapport aux membres du comité. Selon M. Willis, il s'est agi d'un cas où le consultant avait influencé les évaluations, encore que cette influence fût minime.

358. M. Willis n'a constaté aucun problème au sein des comités nos 1 et 2 durant la première phase des cinq comités d'évaluation.

359. En ce qui concerne le comité no 5, M. Willis a remarqué que les évaluateurs avaient tendance à adopter des positions extrêmes, d'un côté ou de l'autre, mais pas aussi extrêmes que le comité no 3, et que la productivité du comité «suivait en général un rythme soutenu». Il a relevé une représentante syndicale qui manifestait une préférence à l'égard des emplois féminins et deux représentants patronaux masculins qui en faisaient autant à l'égard des emplois masculins. M. Willis a par ailleurs témoigné qu'une représentante syndicale manifestait aussi une préférence pour les postes masculins. D'après son analyse, l'influence de deux de ces évaluateurs, une représentante syndicale et un des représentants patronaux masculins, avait tendance à s'annuler. Les autres membres du comité, a-t-il signalé, n'étaient pas influencés par ces deux membres et avaient tendance à ne pas tenir compte de leurs positions.

360. M. Willis a par ailleurs trouvé que les évaluations produites par le comité no 5 étaient en général «pas mal bonnes». Il a identifié deux

85

membres du comité comme étant des évaluateurs «marginiaux», mais il a plus tard recommandé qu'ils deviennent présidents de comité -- lorsque leur nombre a été porté à neuf -- parce qu'à ses yeux ils étaient de bons évaluateurs.

361. Le Tribunal a entendu le témoignage de trois évaluateurs ayant fait partie du comité no 5. Chacun a confirmé la rigueur et la diligence avec lesquelles le comité s'acquittait de sa tâche. Leur témoignage a corroboré une fois de plus le point de vue de M. Willis selon lequel les évaluateurs marginaux n'avaient pas influencé le consensus du comité.

362. Deux évaluateurs qui avaient fait partie de la première version du comité no 5 ont témoigné que les questionnaires discutés dans ce comité étaient difficiles. Une d'entre elles, Mme Mary Crich, a expliqué que les longues discussions du comité avaient porté sur des emplois masculins très difficiles.

363. Une autre évaluatrice du comité no 5, Mme Pauline Latour, a affirmé ce qui suit (volume 171, à la p. 21604, lignes 20-25) :

[TRADUCTION]

R. Nous avons eu de la difficulté... les questionnaires que nous avons dans le comité 5, j'ai l'impression qu'ils étaient plus difficiles à évaluer. Il y en avait plusieurs pour lesquels

nous semblions avoir des questions sans réponse. Alors nous avons définitivement retourné plus de questionnaires dans le comité 5.

364. Plus loin dans son témoignage, Mme Latour a apporté les précisions suivantes (volume 171, de la p. 21605, ligne 12, jusqu'à la p. 21606, ligne 9) :

[TRADUCTION]

Q. Vous avez dit tantôt qu'il y avait des emplois que vous trouviez plus difficiles à évaluer que d'autres. Pourriez-vous décrire lesquels -- donnez-nous peut-être des exemples d'emplois qu'en tant que comité vous avez trouvé plus difficiles que d'autres.

R. Vous allez peut-être trouver que je vous répons de façon détournée, mais, par exemple, les emplois que nous avons le plus de facilité à évaluer étaient ceux qui ressemblaient à plusieurs autres que nous avons déjà cotés. Par exemple, nous avons évalué de nombreux emplois de secrétaire représentant tout un éventail, allant des dactylos aux cadres supérieurs. Nous comprenions bien la nature du travail.

Parfois nous avons affaire à seulement quelques postes qui étaient reliés et nous ne savions pas vraiment comment le travail s'insérait dans la section où travaillait le titulaire. Alors, parce qu'il y avait si peu de liens avec les autres postes, et il y avait pas mal de postes qui n'étaient pas reliés à d'autres

86

postes, nous avons vraiment beaucoup de difficulté à saisir le niveau de complexité du poste.

365. Le Tribunal a entendu le témoignage direct de 15 témoins, qui étaient des évaluateurs dans les comités nos 1, 2, 3, 4, 5, 8 et 9, au sujet de leurs expériences et perceptions pendant qu'ils faisaient partie de leurs comités respectifs. La preuve concernant les comités nos 6 et 7 a été présentée par M. Willis et un de ses consultants, M. Owen. Ni l'un ni l'autre n'ont exprimé des réserves sérieuses au sujet de ce qu'ils avaient observé dans ces comités.

366. En ce qui concerne les témoignages directs de ces 15 témoins, le Tribunal a été impressionné par le degré d'engagement de ces derniers face à l'étude. S'il est vrai que l'évaluation d'emploi est un processus

systematique qui pose un véritable défi intellectuel, il n'en demeure pas moins que ces individus ont tâché de s'entendre sur une évaluation pour chaque poste, et ce, huit heures par jour et cinq jours par semaine sur de longues périodes. M. Willis a observé des variations dans la productivité des comités, dont voici le bilan (basé sur un total de 3 185 questionnaires) : (i) comité 1 - 466; (ii) comité 2 - 431; (iii) comité 3 (avant l'expansion à 9 comités) - 165; comité 4 - première version - 200 évaluations; deuxième version - 52 évaluations et après l'expansion à 9 comités - 160 évaluations; comité 5 - 430 évaluations. Après l'expansion à 9 comités, comité 6 - 197 évaluations, comité 7 (francophone) - 149 évaluations; comité 8 (francophone) - 150 évaluations et comité 9 - 145 évaluations.

367. A la lumière de son expérience dans d'autres études, M. Willis s'attend à trouver certains conflits au sein d'un comité d'évaluation à cause de la diversité des antécédents et des perspectives des évaluateurs. Néanmoins, M. Willis a témoigné que le degré et la nature des conflits qu'il a observés durant l'étude au sein des comités d'évaluation l'ont mis mal à l'aise.

368. Le Comité mixte avait prévu certains des problèmes qui ont surgi dans les comités d'évaluation multiples. Le sous-comité des tests du plan d'évaluation Willis, dans son rapport du 20 juillet 1987 (pièce HR-11A, onglet 19), a fait des recommandations en réponse aux problèmes qu'il avait éprouvés durant une période d'essai de deux semaines. Entre autres, il y avait des incompatibilités de caractères, la fatigue causée par la concentration constante et le stress associé au fait d'être détaché de son emploi normal pour de longues périodes. A la suite de cette expérience, le comité a recommandé la rotation des membres entre les comités d'évaluation, le raccourcissement de la journée ou de la semaine de travail et le recours à des suppléants qui remplaceraient les membres désignés pendant des périodes données. Le Comité mixte n'a jamais donné suite à ces recommandations et on n'a pas communiqué au Tribunal les motifs pour lesquels celles-ci avaient été rejetées.

369. Les évaluateurs ont témoigné qu'ils avaient dû subir de la tension à titre de membres de comité, du stress dans la réalisation d'un consensus, des incompatibilités de caractères, un manque de souplesse de la part de certains évaluateurs, des difficultés vis-à-vis de certains

présidents et des cris de la part de certains évaluateurs. Il arrivait que des évaluateurs sortent des réunions d'évaluation tellement ils étaient frustrés. Ces difficultés étaient aggravées par le remplacement fréquent

de certains membres du comité. Il s'ensuivait un changement dans la dynamique du groupe qui nécessitait des rajustements de la part tant des nouveaux que des anciens membres.

370. A ces problèmes se greffait un milieu de travail rigide orchestré et contrôlé par le chef du Secrétariat de l'étude sur la parité salariale (SEPS), qui était apparemment plus souple avec les évaluateurs patronaux qu'avec les évaluateurs syndicaux. Le chef, M. Pierre Collard, surveillait de près l'heure d'arrivée et de départ des évaluateurs, la période prévue pour le déjeuner et les pauses café. Il insistait pour que les portes soient constamment closes durant les délibérations (entraînant des problèmes de ventilation), limitait l'utilisation du téléphone et gardait toutes les fournitures sous clé (occasionnant des pertes de temps lorsqu'on avait besoin d'articles). Ces contraintes très rigoureuses ont intensifié les frustrations déjà expérimentées par les membres du comité. De plus, certains évaluateurs provenant «de l'extérieur de la province» devaient attendre pendant de longues périodes avant de se faire rembourser leurs frais de déplacement. Cette question, notamment, n'a pas été résolue dans un délai raisonnable.

371. Plusieurs des évaluateurs qui ont témoigné à l'audience ont souligné leur volonté et la nécessité d'adhérer aux postes-repères du Comité directeur. De plus, ont-ils dit, les évaluations devaient être fondées sur les faits contenus dans les questionnaires, aucune autre considération extérieure ne devant entrer en ligne de compte.

372. La seule critique que le Tribunal a entendue concernant la volonté des comités de suivre la discipline du Comité directeur était que, pendant une brève période au début du processus d'évaluation, le comité no 1 avait tendance à suivre sa propre discipline plutôt que celle du Comité directeur. Ce problème a été corrigé dès que les consultants de M. Willis l'ont repéré.

373. Au début de 1989, M. Willis a commencé à exprimer au Comité mixte certaines préoccupations non reliées aux évaluations proprement dites, mais plutôt à des «aspects circonstanciels» qui en avaient découlé. Il a qualifié ces incidents de «fumée» parce qu'ils tenaient en grande partie de la rumeur, incidents qui se seraient produits tant à l'intérieur qu'à l'extérieur des salles de réunion des comités. La façon dont les comités d'évaluation travaillaient et ce qu'il a décrit comme des affrontements entre les parties syndicale et patronale le rendaient de plus en plus mal à l'aise. M. Willis ne pouvait relever quoi que ce soit qui aurait pu indiquer que l'on commençait à manifester de la partialité fondée sur le sexe, mais d'après ses propres observations et celles de ses consultants il savait qu'il «se passait des choses» et que certaines attitudes non

souhaitables se développaient, situation qui n'était pas sans le préoccuper énormément.

374. Il est arrivé à plusieurs occasions, pendant qu'il assistait aux séances de travail des comités, que M. Willis remarque que la position

88

prise par un évaluateur en particulier était très empreinte de partialité. Habituellement, l'évaluateur en question refusait de changer sa cote même s'il ne pouvait la justifier par des faits concrets. La fréquence de ces occasions a commencé à préoccuper M. Willis. Il a constaté ces cas des deux côtés, syndical et patronal, et surtout parmi les cinq comités de la première formation.

375. M. Willis a dit être au courant que des représentants syndicaux tentaient de recruter d'autres évaluateurs dans leur «bloc». Il n'avait jamais vu ce phénomène dans aucune des études d'évaluation auxquelles il avait participé. M. Willis n'a pas directement été témoin d'incidents concernant ce recrutement. M. Owen, toutefois, l'a mis au courant d'un incident où un évaluateur en avait approché un autre à propos des évaluations. M. Owen a témoigné sur les circonstances entourant cet incident, qui s'est produit en février 1989. Il a raconté avoir entendu une conversation entre deux évaluatrices qui sont entrées dans une pièce où il travaillait. Il a entendu l'une d'elles dire à l'autre «je ne pense pas que nous en faisons assez pour les emplois des femmes». Selon M. Owen, l'autre évaluatrice est devenue agitée, elle a haussé la voix et elle a répondu : «Je ne suis pas venue ici pour mousser certains genres d'emplois. Je suis venue ici pour évaluer le travail le plus honnêtement possible.»

376. M. Owen a par ailleurs témoigné qu'il a observé dans les comités certains comportements qu'on pouvait associer à des factions. Il y avait des évaluateurs syndicaux qui semblaient traiter certains emplois de la même façon que des évaluateurs syndicaux dans d'autres comités. Il les a identifiés comme des membres de l'Alliance. Ce qui troublait M. Owen, c'était que dans son expérience professionnelle antérieure -- il avait formé et animé plus de 50 comités d'évaluation -- il n'avait jamais observé ce genre de comportement. Il a également été témoin de pratiques de cotation inusitées, de longues discussions visant à faire adopter un choix particulier et de la sélection de postes-repères ne convenant pas à l'évaluation en cause. A une autre occasion, à l'époque des cinq comités d'évaluation initiaux, on a demandé à M. Owen de présider le comité no 3 parce que la présidente désignée travaillait à titre d'évaluatrice ailleurs. Lorsque la présidente est revenue dans la pièce, une discussion très controversée concernant une évaluation était en cours. La présidente

a demandé à M. Owen de trancher la procédure à suivre et a demandé des rappels au règlement semblables à ce qui est prévu dans Robert's Rules of Order. Or M. Owen ne connaissait pas du tout ce code de procédure et n'a donc pu fournir une réponse appropriée. La présidente a réagi en ordonnant aux évaluateurs de l'Alliance de sortir, ce qu'ils ont fait en claquant la porte. M. Owen a interprété ce malheureux incident comme une tentative de la part d'un côté de contrôler le comité.

377. Comme M. Willis, M. Owen était frustré de ne pouvoir intervenir ou agir.

378. Un autre incident relaté par M. Owen s'est passé à l'automne de 1988. Les membres de l'Alliance n'ont pas assisté aux travaux de leurs comités un jour en particulier qu'ils avaient décidé de «chômer» pour montrer leur appui aux questions de parité salariale à la table de négociation. La négociation collective, semble-t-il, était en cours et les

syndiqués discutaient de la question de savoir si les propositions sur la parité salariale allaient être retirées du processus de négociation. Deux des membres de l'Alliance ayant refusé de prendre part à ce congé de maladie concerté ont dit à M. Owen qu'ils craignaient des représailles de la part de leur syndicat.

379. M. Willis estimait que les conflits au sein des comités tenaient à l'attitude trop répandue qui amenait les gens à se définir comme appartenant à des «camps opposés». Il a confirmé n'avoir jamais vu tant de participants à une étude sur la parité salariale ayant des antécédents en classification, un des facteurs «importants qui ont contribué au conflit en l'occurrence».

380. M. Willis a témoigné que s'il avait fait partie de la fonction publique fédérale et s'il avait eu le contrôle du processus d'évaluation et le pouvoir décisionnel voulu, il aurait apporté certains changements et poursuivi l'étude. Sa préférence aurait été de retirer le personnel qui créait des problèmes et d'engager plus de consultants pour travailler étroitement avec chaque comité.

381. Selon l'opinion experte de M. Willis, la partialité fondée sur le sexe peut se manifester très subtilement dans une étude sur la parité salariale, et il estimait que pour défendre les résultats il devait se rassurer lui-même que les évaluations ne présentaient aucun problème. M. Willis n'était pas sûr si les problèmes comme tels qui existaient

avaient donné lieu à des résultats biaisés. Il a déclaré ce qui suit à ce propos (volume 69, à la p. 8654, lignes 8-14) :

[TRADUCTION]

J'ai affirmé qu'il y avait une contradiction intéressante. J'avais de très sérieuses réserves au sujet d'attitudes, de choses que nous avons observées. Toutefois, lorsque nous avons essayé d'examiner les résultats des comités et de comparer des emplois semblables, nous n'avons pas réussi à détecter une tendance claire qui aurait fait problème.

382. M. Willis a témoigné que, durant les travaux des cinq premiers comités d'évaluation, il a relevé dix évaluateurs qui, pensait-il, manifestaient des préférences pour les emplois masculins ou féminins. À son avis, la majorité manifestait une préférence à l'égard des emplois féminins. Dans ces cas, son approche consistait à conseiller les personnes mises en cause. À ce stade, M. Willis n'était pas en mesure de dire si les évaluateurs en question influençaient les évaluations du groupe. Il s'inquiétait de ne pouvoir justifier ses conclusions autrement que par ses observations personnelles. Il a alerté ses consultants, qui connaissaient déjà ces individus. Ses consultants et lui ont continué de suivre les individus visés et d'examiner les résultats des évaluations dans leur ensemble.

383. Comme autre approche, les consultants ventilaient les évaluations par groupes professionnels et déterminaient si ces individus influençaient le groupe, puis ils cherchaient à établir si, globalement, la partialité

90

semblait constituer un problème. Ce genre de suivi n'a pas permis de déceler de partialité importante.

384. Lorsque M. Willis conseillait les évaluateurs, il s'installait avec eux en privé dans une pièce et discutait de leurs évaluations et des changements qu'il s'attendait d'eux. Dans son témoignage, M. Willis a affirmé n'avoir constaté aucun changement dans les évaluations des individus qu'il avait conseillés. Durant ces séances de consultation, il a appris de certains évaluateurs patronaux qu'ils évaluaient comme ils le faisaient afin de contrebalancer les évaluations effectuées par les évaluateurs syndicaux. La plupart des évaluateurs que M. Willis a conseillés n'ont pas nié leur comportement.

385. Tout au long de l'étude, M. Willis a aussi conseillé les comités. Il observait les comités d'évaluation pendant leurs travaux d'évaluation. Dans ses interventions, il tentait de ramener les évaluateurs aux faits, les incitait à examiner le questionnaire et à discuter du poste réel plutôt que de faire des suppositions ou de s'en tenir à des stéréotypes. A propos de l'efficacité des conseils qu'il a donnés aux comités, M. Willis a déclaré ce qui suit (volume 57, de la p. 7087, ligne 9, jusqu'à la p. 7088, ligne 5) :

[TRADUCTION]

Q. En ce qui concerne le dernier genre de counseling que vous venez de donner aux comités d'évaluation en tant que groupe... Je vous ai déjà demandé votre opinion sur son efficacité pour les individus. Maintenant j'aimerais connaître votre opinion sur le succès que vous avez obtenu auprès des comités d'évaluation comme groupes.

R. C'est un peu difficile à dire. Ces comités étaient plutôt inhabituels si je les compare à la plupart de ceux avec lesquels je travaille, en ce sens que je n'ai pu, dans mes observations, déceler de partialité dans les évaluations ni de tendance quelconque. Par contre, je n'avais pas affaire à des comités qui travaillaient ensemble en vue d'atteindre un résultat juste, équitable et consciencieux.

Ce que j'avais dans de nombreux comités, c'était le syndicat d'un côté et l'employeur de l'autre, et ils étaient à couteaux tirés. Ce n'était pas comme cela dans tous les comités, mais c'était le cas dans plusieurs d'entre eux. La mesure dans laquelle notre counseling a pu les aider, dans certains cas, était négligeable. [c'est nous qui soulignons]

386. Plus tard dans son témoignage, on a demandé à M. Willis ce qu'il avait voulu dire dans l'extrait ci-dessus lorsqu'il avait affirmé : «[...] je n'avais pas affaire à des comités qui travaillaient ensemble en vue d'atteindre un résultat juste, équitable et consciencieux.» M. Willis a expliqué que le mot important était «consciencieux». A ses yeux, cela signifie que l'employé travaille fort et atteint ses propres normes personnelles. En l'occurrence, M. Willis a témoigné que tous les

évaluateurs qu'il a observés dans chaque comité s'acquittaient de leur tâche de façon consciencieuse. D'autre part, les consultants cherchaient à

faire appliquer une norme pour que chaque emploi soit traité d'une manière juste, objective et impartiale. Dans ce contexte, M. Willis a déclaré avoir été à même de constater que cette norme n'était pas appliquée avec la constance voulue, encore que ce phénomène ne fût pas répandu parmi tous les évaluateurs et tous les comités.

387. Le Tribunal conclut du témoignage des évaluateurs participants à qui on a demandé comment ils envisageaient les évaluations que ceux-ci étaient honnêtes, dévoués et consciencieux. Ces témoins ont dit avoir observé le même engagement de la part de la plupart des membres des comités.

388. Des questions précises ont été posées aux évaluateurs qui ont témoigné à propos des préoccupations de M. Willis, dont il a parlé en utilisant le mot «fumée». Les questions concernaient les rumeurs selon lesquelles certains comités votaient «en bloc»; autrement dit les évaluateurs syndicaux votaient ensemble afin d'obtenir la même cote pour les éléments ou sous-facteurs et tous les évaluateurs patronaux en faisaient autant pour obtenir eux aussi la même cote. De plus, les membres des comités auraient utilisé d'autres moyens de communication, le langage gestuel notamment, pour indiquer comment certains évaluateurs cotaient les postes de façon à influencer les décisions.

389. Aucun des évaluateurs qui ont témoigné n'ont observé ce genre de comportement ni quelque autre forme de communication organisée que ce soit qui aurait eu pour objet de surestimer les emplois féminins et de sous-estimer les emplois masculins. On aurait apparemment discuté de langage gestuel dans un contexte social, ce qu'un témoin a attribué aux frustrations que chacun éprouvait face à un processus d'évaluation difficile. Le tout s'était passé dans un contexte de plaisanterie.

390. Le Tribunal a entendu des témoignages directs concernant trois incidents de conduite inappropriée. Dans le premier incident, les deux évaluatrices en cause étaient les représentantes de l'Alliance ayant pris part à la conversation entendue par M. Owen et dont il a été fait état plus tôt. Une évaluatrice du comité no 4 a témoigné qu'elle avait été approchée par une autre évaluatrice de son comité qui avait voulu savoir si elle évaluait d'une manière juste les postes à prédominance féminine. La témoin a eu l'impression que cette personne voulait qu'elle augmente ses cotes. Elle a répondu en disant qu'elle était là pour évaluer ces postes d'une façon juste et du meilleur de sa connaissance par comparaison avec l'ensemble des postes. Pour autant que la témoin fût concernée, l'incident s'était clos ainsi.

391. Le deuxième incident mettait aussi en cause deux évaluatrices de l'Alliance. La témoin a affirmé qu'elle avait été approchée par une autre

évaluatrice qui voulait la rencontrer à l'extérieur de la pièce pour discuter comment évaluer les emplois. Essentiellement, cette dernière voulait qu'elle favorise les emplois à prédominance féminine situés dans une fourchette supérieure de la même façon qu'elle le faisait, elle. La

92

témoin estimait qu'il ne s'agissait pas d'une approche objective et elle lui a répondu qu'elle continuerait de coter les postes objectivement.

392. En ce qui concerne les premier et deuxième incidents, les deux témoins ont déclaré que l'affaire n'avait eu aucun effet sur leur façon d'évaluer. La preuve est claire que, dans le premier incident, la personne ayant approché sa collègue et lui ayant fait la demande que l'on sait était reconnue par le comité pour ses cotations biaisées, que celui-ci avait tenté en vain de modifier. Comme celle-ci a refusé de changer, le reste du comité a choisi de l'ignorer.

393. Le troisième incident concernait une évaluatrice de l'Institut. Cette dernière a témoigné qu'il y avait eu dans sa chambre d'hôtel une soirée sociale à laquelle avaient pris part 10 ou 15 évaluateurs. Une conversation s'est engagée plus tard dans la soirée entre la témoin et quatre autres évaluatrices de l'Alliance. L'évaluatrice de l'Institut a témoigné avoir soutenu qu'on devait évaluer les postes de manière objective, ce qui avait rendu deux des membres de l'Alliance très agressives à son endroit. Celles-ci avaient répondu que l'étude était une occasion pour les femmes d'obtenir qu'on fasse quelque chose pour elles, et que rien n'allait être accompli si les emplois des femmes n'obtenaient pas des cotes supérieures, l'étude étant leur dernière chance. Selon l'évaluatrice de l'Institut, «la situation est ensuite devenue un peu trop personnelle». Un autre témoin de l'Alliance a déclaré à l'audience que cet incident avait été une attaque verbale contre l'évaluatrice de l'Institut.

394. Pour ce qui est du troisième incident, l'évaluatrice de l'Institut croyait que les personnes qui l'avaient affrontée dans sa chambre d'hôtel occupaient une fonction d'autorité au sein de l'Alliance et qu'elles pouvaient convoquer des réunions et influencer d'autres évaluateurs de l'Alliance. Lors de son témoignage, elle a admis qu'elle n'avait plus de raison de croire que tel était le cas et qu'elle n'avait plus l'impression que les évaluateurs de l'Alliance s'étaient concertés en vue d'agir malhonnêtement.

395. M. Willis se rappelle avoir discuté des problèmes propres aux comités d'évaluation avec le mini-comité mixte, un sous-comité du Comité mixte. Ce sous-comité a été formé pour régler les problèmes des comités

d'évaluation liés à la procédure. M. Willis a témoigné avoir discuté avec deux de ses membres, M. Gaston Poiré et Mme Elizabeth Millar, de certains des évaluateurs qui, à son avis, créaient des problèmes. Il a proposé que certains individus soient retirés des comités d'évaluation, mais il n'a pas reçu le soutien actif qu'il s'attendait de recevoir. Ainsi, le Comité mixte a réaffecté les individus problématiques lorsque le nombre des comités est passé de cinq à neuf. Selon M. Willis, après l'expansion des comités certains d'entre eux ont bien fonctionné, tandis que d'autres ont continué d'avoir des problèmes, mais pas autant que les cinq comités initiaux. «Il n'y avait rien de pire que le comité no 3 initial», a-t-il affirmé. Selon lui, c'était le fond du baril, mais après, la situation «s'est améliorée» (volume 69, à la p. 8653).

396. M. Willis considérait comme inacceptable ce qui se passait dans les comités d'évaluation. Il a conclu qu'il lui fallait effectuer une

93

analyse plus approfondie des résultats avant d'être en mesure d'appuyer le résultat de l'étude. Bien que, en janvier et février 1989, il n'eût pas décelé de partialité fondée sur le sexe dans les évaluations, il a déclaré ce qui suit dans son témoignage (volume 58, de la p. 7229, ligne 13, jusqu'à la p. 7230, ligne 3) :

[TRADUCTION]

R. Je pense qu'une des choses qui caractérisaient toute l'étude, la plainte concernant l'égalité de rémunération pour fonctions équivalentes, était d'évaluer un large éventail de postes sans partialité fondée sur le sexe. Je pense que tout le processus que nous avons établi et le système d'évaluation qui a été utilisé, la façon dont nous avons essayé de travailler avec les groupes, tout cela visait principalement à éviter que des évaluations reflètent des relations traditionnelles, ou à faire en sorte qu'on ne puisse rien relever qui pourrait être considéré comme de la partialité fondée sur le sexe.

J'estime qu'à tous les stades de l'étude il était primordial que nous exercions une vigilance de tous les instants et que nous renforçons la nécessité de produire des évaluations objectives, justes et équitables de tous les genres de postes.

397. Dans une lettre datée du 4 mai 1989 qu'il a adressée aux présidents du Comité mixte, M. Scott Gruber, un consultant de M. Willis,

a recommandé qu'on effectue une analyse spéciale des résultats des comités d'évaluation. La lettre dit notamment ce qui suit :

[TRADUCTION]

La présente lettre décrit notre proposition d'une analyse spéciale des résultats des comités d'évaluation, que nous estimons opportune et appropriée. La question à étudier est la suivante :

Les évaluations produites par les cinq comités d'évaluation (1 à 5) sont-elles conformes à celles effectuées par le Comité directeur?

[...]

Pour cette analyse, la méthodologie suivante sera appliquée :

1. Un échantillon sera prélevé au hasard parmi les résultats d'évaluation de chaque comité. La taille de l'échantillon sera 10 % des postes évalués, avec un minimum de 25 par comité. Cette dernière disposition permet un examen raisonnable des efforts des comités dont la productivité est faible. Suivant ces lignes directrices, l'échantillon total comprendra environ 140 postes.
2. Un consultant de M. Willis connaissant bien les évaluations du Comité directeur examinera chacun des 140 questionnaires

94

et effectuera des comparaisons avec des questionnaires appropriés ou correspondants du Comité directeur.

3. S'appuyant sur cet examen, le consultant étudiera ensuite le bien-fondé des évaluations finales, établies par consensus lors des séances de récapitulation, émanant des cinq comités ainsi que les questionnaires des postes-repères du Comité directeur qu'ils ont choisis. Les problèmes et tendances seront relevés par comité et pour l'ensemble du groupe.
4. Les informations relatives à la prédominance d'un sexe sur l'autre seront recueillies pour les postes de l'échantillon à ce stade. Des analyses additionnelles permettront de savoir si un ou plusieurs des comités ont manifesté dans leurs résultats une tendance à privilégier les groupes à

prédominance masculine ou féminine. D'autres variables que le sexe pourraient aussi être incluses dans l'analyse à ce stade-ci.

5. Un rapport décrivant le processus de la recherche, l'analyse et les résultats sera préparé et vous sera présenté.

[...]

Nous considérons qu'il s'agit là d'une étude d'assurance de la qualité visant à examiner les évaluations de cinq comités formés de personnes aux antécédents -- études, expérience et profession -- variés et ne pouvant refléter parfaitement les caractéristiques de la composition du Comité directeur [...] Une question fondamentale à explorer est de savoir si les comités ont appliqué comme il faut et de manière uniforme les évaluations des postes-repères du Comité directeur durant le processus de comparaison.

[...] Si les résultats indiquent que les cinq comités se sont acquittés de leurs tâches respectives en harmonie avec le Comité directeur, de nombreuses préoccupations concernant l'étude auront été réglées. En revanche, si des problèmes sont relevés il sera possible de prendre des mesures correctives et les neuf [nouveaux] comités pourront poursuivre leur travail en profitant des connaissances acquises.

(pièce HR-11B, onglet 32)

398. M. Willis a demandé qu'une enquête éclair visant à faire le point sur la validité des évaluations soit effectuée sur les 2 000 postes qui avaient été évalués à ce moment-là. Il a proposé que l'un de ses consultants examine 10 % des évaluations et compare les évaluations des comités à celles du consultant. De cette façon, il se convaincrait au moins qu'il n'y avait aucune indication de problème ou l'analyse révélerait la possibilité qu'un problème puisse exister. Son intention à ce moment-là était de commencer par une étude restreinte qui pourrait révéler des indications de discrimination. Si un problème était décelé, il prévoyait

faire une seconde étude, plus en profondeur celle-là, qui exposerait la portée du problème mis à jour par la première étude. Il n'a pas fait savoir directement au Comité mixte qu'il entrevoyait une approche à deux volets.

399. Le Comité mixte a accepté son idée de mener une étude restreinte et cette analyse a débuté au printemps de 1989. L'analyse -- intitulée Special Analysis of Evaluation Committee Results (la «Wisner 222») -- a été exécutée par le consultant de M. Willis, M. Jay Wisner (pièce PSAC-4). M. Wisner a réévalué 222 des évaluations qu'avaient effectuées les comités (les cinq initiaux et les neuf qui leur ont succédé). Lorsque l'échantillon de 222 postes a été constitué, les comités d'évaluation multiples évaluaient toujours des questionnaires et les neuf comités étaient à la tâche depuis environ trois mois.

(vii) Recyclage des comités d'évaluation multiples

400. Cette étape du processus Willis consiste à recycler un comité d'évaluation ou un évaluateur en particulier. Si le consultant avait remarqué un problème, la séance de recyclage avait pour objet de ramener le comité ou l'individu à la discipline du Comité directeur. Le recyclage pouvait être aussi informel que la formation dispensée durant le mandat du Comité directeur, lorsque M. Willis aidait le comité à interpréter le plan, ou il pouvait comporter des séances plus structurées comme certaines qui ont été données au cours des travaux des deux générations de comités (les cinq premiers et les neuf subséquents). La séance initiale de formation a été donnée aux cinq comités d'évaluation durant la semaine du 19 au 23 septembre 1988 (pièce HR-11B, onglet 27), et la séance officielle suivante de recyclage a eu lieu en mars-avril 1989, après l'expansion des comités. Entre ces deux sessions, les consultants ont donné de la formation moins structurée suivant les besoins.

(viii) Récapitulation

401. Un autre procédé d'appoint ou de protection qui fait partie du processus Willis est la récapitulation, qui est synonyme d'examen provisoire. Selon M. Willis, le premier examen provisoire a habituellement lieu après que 25 à 30 postes ont été évalués. Ces postes sont ensuite listés par ordre décroissant des cotes et les comparaisons sont effectuées entre les postes, facteur par facteur. L'objet est de détecter les évaluations qui ne présentent pas le même degré d'uniformité que les autres évaluations. Une séance finale de récapitulation a lieu lorsque tous les emplois ont été évalués. Cette méthode vise à assurer l'uniformité au sein d'un comité et révèle si un comité s'écarte de sa discipline. Un évaluateur qui n'a pas participé à l'évaluation initiale ne peut prendre part à la séance de récapitulation.

402. Le Comité directeur a tenu cinq séances de récapitulation qui ont donné lieu à des changements insignifiants. Dans l'ensemble, M. Willis était satisfait des résultats de la récapitulation accomplie par le Comité directeur. Chacun des autres comités d'évaluation a aussi tenu quatre ou

cinq séances du genre. Dans leur cas, l'accent était différent étant donné qu'il s'agissait davantage pour eux de vérifier s'ils respectaient la

96

discipline du Comité directeur. Concrètement, ils ont examiné leurs propres évaluations et les ont comparées à la discipline du Comité directeur afin de s'assurer qu'il y avait harmonie entre les deux.

403. Si les comités d'évaluation n'étaient pas conformes à la discipline du Comité directeur facteur par facteur, il s'ensuivrait un manque de compatibilité ou d'uniformité au niveau des évaluations globales. Le degré de libéralisme ou de conservatisme n'est pas toujours le même d'un facteur à l'autre. L'important, c'est que tous les postes soient traités de la même façon; autrement dit, si les comités cotent l'habileté en communications interpersonnelles de manière libérale, alors ils doivent faire preuve du même libéralisme à l'égard de tous les postes, et s'ils sont conservateurs quand ils cotent le facteur connaissances et compétences, alors ils doivent l'être pour tous les postes. M. Willis n'a toutefois pas exprimé d'opinion directe sur l'efficacité de la récapitulation effectuée par les comités d'évaluation multiples.

#### D. TESTS DE FIABILITÉ

404. Dans le cadre du processus Willis, M. Willis recommande en général que les évaluations fassent l'objet de tests de fiabilité.

##### (i) Tests de fiabilité inter-évaluateurs (FIE)

405. Le premier genre de fiabilité qu'on veut vérifier est la fiabilité inter-évaluateurs (FIE), qui vise à identifier les évaluateurs qui ont peut-être commencé à développer dans leurs cotations des tendances qui les éloignent de la ligne de conduite suivie par les autres membres de leur comité. M. Willis a introduit ce concept durant la phase de planification de l'étude sur la parité salariale.

406. M. Willis a expliqué que les tests FIE étaient recommandés pour deux raisons. Pour des raisons personnelles il estime que, lorsqu'il doit conseiller les évaluateurs qui manifestent de la partialité dans leurs cotes, il lui est utile d'avoir en main des données statistiques sur lesquelles étayer ses observations et opinions. Autrement, ce serait la parole du consultant contre celle de l'évaluateur. M. Willis trouve utile d'utiliser les résultats de ces tests avec l'évaluateur et de lui demander d'examiner la tendance qui se dégage de ses évaluations. Il est ainsi plus facile de discuter du problème avec l'évaluateur et de le convaincre de

changer. M. Willis a témoigné que, dans certains cas, les évaluateurs refusent de reconnaître que leurs évaluations sont biaisées s'ils ne sont pas confrontés à des preuves statistiques.

407. La deuxième raison pour laquelle M. Willis a introduit les tests FIE tient à l'envergure et à l'importance de l'étude sur la parité salariale. Comme les résultats de l'étude seront scrutés par le public, a-t-il expliqué, le fait de ne pas avoir effectué de telles vérifications pourrait leur être reproché.

408. M. Willis a clairement dit qu'il n'était pas nécessaire pour lui et ses consultants d'effectuer des tests FIE pour observer et identifier les évaluateurs marginaux ou minoritaires. Un consultant expérimenté saura

97

toujours les reconnaître, mais ces tests lui fournissent certaines preuves statistiques.

409. La recommandation de M. Willis concernant les tests FIE n'a pas été acceptée par le Comité mixte au stade de la planification initiale. M. Willis a présenté de nouveau ce concept lorsque le Comité directeur a amorcé ses travaux. La question de savoir s'il fallait ou non effectuer des tests FIE a fait l'objet d'une certaine controverse au sein du Comité mixte. A la réunion du Comité mixte du 13 janvier 1988 (pièce R-9), le côté patronal a reconnu en principe la nécessité d'effectuer des tests FIE en plus des tests de fiabilité inter-comités, mais il a mis en doute la proposition courante de M. Willis.

410. Le Comité mixte a formé un sous-comité de fiabilité inter-évaluateurs et des méthodes (le «sous-comité FIE») qu'il a chargé d'explorer cette question. On lui a donné le mandat suivant :

[TRADUCTION]

- a) déterminer les méthodes et recherches qui sont nécessaires pour tester la fiabilité des évaluateurs des comités d'évaluation et présenter des recommandations à ce sujet;
- b) évaluer les méthodes de recherche en fonction de l'étude sur la parité salariale dans son ensemble et présenter des recommandations à ce sujet.

(pièce HR-11A, onglet 26)

411. M. Willis n'était pas sûr au juste pourquoi le Comité mixte était réticent à donner le feu vert aux tests FIE, mais le sous-comité a finalement décidé d'engager le cabinet de consultants Tristat Resources pour effectuer les tests. Pour sa part, M. Willis a souscrit à cet arrangement. Les tests ont été réalisés par M. Richard Shillington, un statisticien qui a témoigné devant le Tribunal à titre d'expert.

412. M. Willis était déçu que les tests FIE ne débutent pas avant que le Comité directeur ait terminé ses travaux. A cause de cela, il n'a pu utiliser les résultats pour conseiller les évaluateurs du Comité directeur qui cotaient différemment des autres (les «marginiaux»). «En revanche, j'étais satisfait des tests comme tels», d'ajouter M. Willis.

413. Au départ, M. Willis avait proposé qu'on effectue les tests FIE au moins trois ou quatre fois durant les travaux du Comité directeur, ce qui aurait produit des données statistiques qu'il aurait pu utiliser dans ses discussions avec les évaluateurs qui manifestaient de la partialité fondée sur le sexe. Pendant les travaux du Comité directeur, M. Willis a identifié deux évaluateurs marginaux, observation que les tests FIE ont confirmée. Il a rencontré ces évaluateurs, mais comme aucun test FIE n'avait alors été fait, il n'avait aucune documentation sur laquelle étayer ses conseils.

98

414. M. Willis n'avait pas le pouvoir de retirer les évaluateurs qu'il avait identifiés comme marginaux. A l'époque, il estimait que leur parti pris était subtil, inefficace et sans effet néfaste sur le travail du Comité directeur. Dans les deux cas en question, les conseils de M. Willis n'ont eu pratiquement aucun effet. Ces deux évaluateurs avaient tendance à se neutraliser mutuellement, de signaler M. Willis. Les tests FIE ont confirmé l'identité de ces deux personnes.

415. Bien que les tests FIE n'aient pu aider M. Willis à conseiller les évaluateurs marginaux, à son avis les tests pouvaient servir «après coup pour établir l'uniformité du processus d'évaluation». Le rapport de M. Shillington sur les tests FIE qu'il avait appliqués aux évaluations du Comité directeur a été présenté le 31 juillet 1988. Ce rapport est connu sous le nom de «rapport Tristat».

416. La première intervention de M. Shillington dans l'étude sur la parité salariale remonte au printemps de 1988. Un membre du sous-comité FIE représentant le Conseil du Trésor l'a approché pour lui demander s'il était intéressé à prendre part au travail du sous-comité. Bien qu'il fût engagé par l'employeur, il considérait le sous-comité FIE comme son client.

Dans le contexte des tests FIE, M. Shillington a fait des tests statistiques pour analyser et interpréter la fiabilité inter-évaluateurs. Ces tests visaient à déterminer si les évaluateurs travaillaient de façon uniforme, notamment dans leur façon de traiter les questionnaires des groupes professionnels à prédominance masculine et féminine.

417. M. Shillington a interprété son rôle comme étant d'aider le sous-comité FIE à élaborer une méthodologie qui, en combinaison avec les données, pouvait servir à répondre à leurs questions et les aider à prendre certaines décisions. Le sous-comité FIE voulait surtout identifier les évaluateurs qui semblaient manifester une préférence pour un sexe ou un parti pris contre l'autre sexe dans leurs questionnaires, mais il voulait vérifier d'autres aspects également. Il souhaitait notamment déterminer si ces évaluateurs exerçaient une influence au sein de leur comité. (Dans ce contexte, on entendait par influence le fait pour ces évaluateurs de pouvoir ramener la cote de consensus du comité vers leur propre cote initiale.)

418. M. Shillington a utilisé une combinaison de tests statistiques -- test t, chi carré et scores z (qui sont semblables aux tests t) -- pour comparer les écarts entre les cotes individuelles des évaluateurs et les moyennes des comités afin de déterminer si les questionnaires masculins et féminins étaient traités différemment.

419. M. Shillington a identifié deux évaluateurs du Comité directeur qui manifestaient un parti pris systématique contre l'un ou l'autre sexe dans leurs cotes. Le premier était un représentant patronal masculin qui attribuait aux postes à prédominance masculine une cote supérieure à celle du comité, et l'autre était une représentante syndicale qui accordait aux postes à prédominance féminine une cote supérieure à celle du comité. Ces évaluateurs étaient les mêmes que ceux que M. Willis avaient identifiés et qui ont fini par devenir membres du mini-comité directeur. D'après les résultats des tests de fiabilité, il n'y avait pas d'écart important entre

leurs cotes et celles du comité sur aucun questionnaire; les tests ont plutôt révélé dans leur cas une tendance légère et subtile, mais tout de même assez fréquente, à coter différemment de leurs collègues.

420. Le sous-comité avait également demandé à M. Shillington de repérer les évaluateurs «influent», lui posant la question suivante : «Y avait-il des évaluateurs en particulier qui semblaient capables de faire cela plus souvent que les autres évaluateurs?» Pour répondre à la question, M. Shillington a examiné les questionnaires où la cote de

consensus n'était pas située près du point milieu des cotations afin de déterminer comment souvent des évaluateurs en particulier se trouvaient dans la situation d'avoir apparemment ramené la cote du groupe vers la leur. A l'aide de cette méthode, certains évaluateurs ont été identifiés comme étant influents.

421. On a ensuite demandé à M. Shillington de préciser la mesure dans laquelle les évaluateurs chez qui on avait noté une partialité fondée sur le sexe exerçaient une influence. Les résultats des tests ont montré que les individus qui exerçaient une influence importante sur le comité n'étaient pas ces deux évaluateurs en question et que les évaluateurs qui influençaient le plus le groupe ne manifestaient aucun parti pris basé sur le sexe.

422. Le troisième et dernier volet des tests FIE consistait à déterminer quels questionnaires feraient l'objet d'un nouvel examen. Cette étape a découlé de l'identification des évaluateurs influents. Le sous-comité a utilisé les résultats des tests pour mettre en évidence les questionnaires pour lesquels la cote de consensus semblait plus élevée ou plus faible par comparaison avec les cotes initiales. Il a conclu qu'environ 103 questionnaires, qualifiés d'«exceptionnels» (unusual), devraient être réexaminés. Un seul facteur, les conditions de travail, avait fait passer 43 de ces questionnaires dans la catégorie des «exceptionnels».

423. Dans son témoignage, M. Shillington a expliqué les limitations des méthodes de test dont il était fait état dans son rapport du 31 juillet 1988. Lorsqu'il s'est agi de comparer les cotes initiales des évaluateurs aux cotes moyennes du comité, de faire remarquer M. Shillington, il a fallu faire l'hypothèse que les comités étaient moins biaisés que les évaluateurs eux-mêmes. Dans ce contexte, la moyenne générale du comité est considérée comme plus fiable.

424. Une autre des limitations relevées par M. Shillington est expliquée dans le rapport Tristat :

[TRADUCTION]

De plus, le fait qu'un évaluateur ait systématiquement favorisé certaines professions dominées par un ou l'autre sexe n'implique pas une préférence fondée sur le sexe. Comme les sexes n'étaient pas distribués également dans la population, il a pu s'agir simplement d'un parti pris pour ou contre un autre facteur quelconque commun aux professions dominées par les hommes ou les

femmes. Par exemple, si un évaluateur avait eu un parti pris en faveur des études supérieures, il aurait été identifié comme ayant une préférence pour les hommes étant donné que cette caractéristique est plus fréquente parmi les postes de cadres. De même, tout parti pris d'un évaluateur pour les compétences techniques ou le travail physique aurait paru comme une préférence fondée sur le sexe.

(pièce HR-39, à la p. 5)

425. A propos de la limitation expliquée plus haut dans l'extrait du rapport Tristat, M. Shillington a fait les remarques suivantes dans son témoignage (volume 86, de la p. 10653, ligne 10, jusqu'à la p. 10656, ligne 11) :

[TRADUCTION]

LE TÉMOIN : Les statistiques mathématiques ne sont pas très utiles ici. C'est essentiellement une question d'interprétation.

Merci d'avoir signalé cette limitation à mon attention. Lorsque j'ai essayé de résumer ce rapport, je ne l'ai pas mentionnée et c'était une limitation importante.

Les statistiques mathématiques peuvent être utiles pour établir qu'un individu traitait les questionnaires provenant des groupes à prédominance masculine différemment de ceux provenant des groupes à majorité féminine. Mais elles ne peuvent pas faire grand-chose pour vous aider à comprendre pourquoi.

La limitation qui est exprimée dans la section que vous avez signalée, à savoir qu'il pouvait s'agir d'un rapport indirect avec une préférence pour les études ou le travail manuel, ou quelque chose du genre, est certainement une considération valable. Quelqu'un qui aurait eu une forte préférence et qui aurait pensé que l'on sous-estimait les études supérieures ou qu'on sous-estimait ou surestimait le travail à l'extérieur aurait pu donner l'impression d'avoir manifesté une préférence ou une partialité -- j'emploierai les deux mots de façon interchangeable pour un instant -- fondée sur le sexe, et vous n'auriez aucune façon de savoir si c'était directement relié au sexe ou avec quelque chose qui est en corrélation avec le sexe.

Si on avait, comme je l'ai expliqué, inséré dans le processus des questionnaires hypothétiques, c'est-à-dire des questionnaires construits pour donner l'impression d'une différence fondée sur le sexe bien qu'ils eussent été identiques à tous les autres points de vue, alors on aurait pu répondre à certaines questions qui se posent ici.

Voici en fait ce que vous auriez pu faire. Si nous avions affaire à quelqu'un qui avait une préférence pour les études supérieures et si cette préférence se traduisait dans une partialité fondée

101

sur le sexe, pourrions-nous construire trois ou quatre questionnaires qui sont tous semblables en ce sens qu'ils requièrent tous des études avancées mais diffèrent quant à leur composition hommes-femmes? Vous pourriez alors examiner les questionnaires pour ces individus et chercher à savoir si, en comparant deux emplois qui exigeaient des études supérieures mais qui étaient légèrement différents du point de vue du sexe, si ces individus traitaient ces emplois différemment ou non. Puis vous pourriez essayer de préciser si c'était véritablement une préférence fondée sur le sexe qui était en jeu ou s'il s'agissait plutôt d'un parti pris pour les études supérieures.

La préférence pour un ou l'autre sexe décelée par l'analyse statistique pourrait en fait être une relation indirecte avec une autre préférence quelconque. Essentiellement, c'est une question d'interprétation. Les mathématiques ne peuvent pas vraiment vous aider, sauf, j'imagine, pour former votre jugement. Plus la relation est forte, plus l'écart dans le traitement des questionnaires masculins et féminins est grand, plus on serait enclin à conclure -- à juger -- que c'est véritablement une préférence fondée sur le sexe qui est en jeu et non quelque chose qui serait en corrélation avec cela.

Vous parlez d'une préférence fondée sur le sexe par opposition à une partialité fondée sur le sexe. Nous utilisons ces deux expressions de façon assez interchangeable dans notre travail parce qu'il est possible qu'on étiquette quelqu'un de partial alors qu'en réalité l'écart en question pouvait avoir trait à une préférence pour les études ou le travail manuel, ce qui est une relation secondaire. Nous mettrions quiconque en garde contre l'idée de parler alors de préférence.

Pour ce qui est des deux évaluateurs en question, la relation dans les données était tellement forte que j'aurais de la difficulté à croire que ce n'est pas le sexe des titulaires qui explique pourquoi ils ont traité les questionnaires comme ils l'ont fait.

426. Le sous-comité de fiabilité inter-évaluateurs a produit son propre rapport sur les tests de fiabilité effectués par M. Shillington. Il a présenté son rapport le 15 juillet 1988, environ deux semaines avant la publication officielle du rapport Tristat. Le rapport du sous-comité du 15 juillet diffère du rapport Tristat du 31 juillet 1988 en ceci qu'il parle des 103 questionnaires «problématiques» qu'il était «nécessaire» de réévaluer. En revanche, le rapport Tristat qualifie ces questionnaires d'«exceptionnels» et signale qu'il «y aurait lieu de les réexaminer» et non de les réévaluer. Selon M. Shillington, le sous-comité, dans son rapport, a utilisé des termes plus forts que lui; pour sa part, il estimait que les questionnaires désignés devaient être «examinés, rien de plus».

427. M. Shillington était présent à la réunion du Comité mixte du 15 juillet 1988 lorsque le rapport du sous-comité FIE a été déposé. Ce rapport, rédigé par un représentant patronal, indiquait que des «influences» s'étaient exercées dans 103 des quelque 500 postes-repères, ce

qui rendait nécessaire un examen plus approfondi et peut-être une réévaluation et une nouvelle séance de récapitulation.

428. M. Shillington a témoigné que l'emploi du terme «influences» dans ce contexte ne reflétait pas ce qui avait été convenu dans le sous-comité FIE. Le rapport qu'il avait présenté, lui, le 31 juillet 1988 représentait son meilleur souvenir des opinions exprimées par le sous-comité. A son avis, ses propres recherches ne permettaient pas d'affirmer que des influences s'étaient exercées à l'égard des 103 questionnaires en question.

429. M. Willis a lui aussi, dans son témoignage, parlé de cet aspect du rapport Tristat (pièce HR-39) et du rapport du sous-comité FIE (pièce HR-11B, onglet 26B). Il ne souscrivait pas à la section du rapport du sous-comité FIE portant sur les évaluateurs ayant exercé une «influence». Il a fait remarquer que le sous-comité avait semblé ne pas tenir compte du fait que M. Willis considérait qu'il était nécessaire et souhaitable que les évaluateurs aient l'occasion de rajuster leurs cotes lors du consensus à partir d'informations concrètes qu'on leur aurait présentées. Le fait que la cote ait dans certains cas été influencée par un évaluateur minoritaire plutôt que par les évaluateurs majoritaires n'est pas en soi

indicatif d'un problème. Voici les propos que M. Willis a tenus à cet égard (volume 38, à la p. 4803, lignes 7-20) :

[TRADUCTION]

Il y avait, si je me souviens bien, une couple d'évaluateurs dans le comité d'évaluation qui ont effectivement exercé une influence, non pas parce qu'ils avaient un parti pris mais parce qu'ils étaient brillants, doués en analyse et respectés de tous. Habituellement lorsqu'ils affirmaient quelque chose à propos d'une évaluation ou lorsqu'on leur demandait de fournir des informations concernant les faits à l'appui de leur cote, ils avaient en général des raisons très solides, et chacun respectait ces raisons. Donc il est arrivé que d'autres membres du Comité directeur réagissent à eux.

Je ne considère pas qu'il s'agit là d'une limitation. Je ne pense pas que c'était une des étapes qui avaient été incorporées au processus.

430. En fin de compte, M. Willis a de fait procédé à une réévaluation des 103 questionnaires relevés par le sous-comité. Ce dernier avait demandé à M. Willis et à ses consultants de vérifier l'hypothèse selon laquelle les 103 questionnaires présentaient un problème. Un des consultants de M. Willis, M. Jay Wisner, a effectué les réévaluations et préparé un rapport pour le Comité mixte. Son analyse est contenue dans le rapport intitulé *Analysis and Conclusions Concerning the Master Evaluation Committee's Work* et est daté du mois de juillet 1988 (pièce R-22). M. Willis a témoigné avoir examiné les évaluations de M. Wisner et avoir apporté des modifications mineures au rapport. C'était «notre» conclusion, a-t-il signalé, qu'un examen systématique d'autres évaluations n'était pas justifié, non plus qu'il n'y avait lieu de reconvoquer le Comité directeur. M. Willis trouvait les évaluations convenables et il n'avait rien à redire

103

sur les résultats globaux, compte tenu des écarts aléatoires raisonnables constatés entre les évaluations. De plus, il estimait que, à ce stade-ci, on devait aller de l'avant comme prévu avec l'étude sur la parité salariale.

431. Les conclusions de M. Willis sont contenues dans un rapport au Comité mixte daté de juillet 1988. Nous reproduisons ci-dessous les conclusions concernant le réexamen de l'analyse de fiabilité inter-évaluateurs (FIE) :

[TRADUCTION]

Après avoir soigneusement scruté les questions soulevées par l'analyse que le comité FIE a faite des évaluations du Comité directeur, nous concluons que la recommandation principale de ce rapport, à savoir que le Comité directeur soit reconvoqué pour réexaminer un grand nombre de ces évaluations, n'est pas fondée.

Nous avons réexaminé les évaluations que l'analyse de fiabilité inter-évaluateurs avait relevées comme étant «exceptionnelles». Nous n'avons trouvé aucune indication comme quoi certains évaluateurs auraient exercé une «influence indue» sur l'évaluation du groupe arrêtée par consensus. A notre avis, la très grande majorité des évaluations listées par le comité FIE résultent de l'application exacte et uniforme du plan d'évaluation par le Comité directeur et ne devraient pas être changées.

En ce qui concerne les quelques postes pour lesquels nous recommandons une réévaluation, nous n'avons trouvé aucune influence systématique, de la part d'une minorité, qui aurait donné lieu à des évaluations que nous réprouverions; dans certains cas, nous recommandons que la cote s'éloigne davantage du point milieu des évaluations initiales. Nous croyons que le réexamen éventuel par le Comité directeur des dix évaluations qui, pensons-nous, devraient être révisées ne devrait pas forcément retarder la convocation des cinq sous-comités. Nous recommandons que ces examens soient combinés à l'examen des postes-repères demandé par l'un des cinq sous-comités.

Nous n'avons aucune réserve sérieuse quant à la façon dont le Comité directeur comprend et applique le plan d'évaluation. Sa façon d'appliquer le plan d'évaluation aux postes (sa «discipline») diffère à certains égards de la façon dont les consultants procéderaient. Cependant, compte tenu de la manière dont on a déterminé la composition du Comité directeur, la discipline de ce dernier constitue un reflet plus fidèle de la valeur des postes telle qu'on la conçoit couramment dans l'administration publique fédérale que ce qu'un consultant pourrait déterminer d'un point de vue extérieur. Ce genre d'adaptation du plan au climat et aux conditions d'une organisation par un comité d'évaluation est normal et approprié. Nous serions inquiets s'il était prouvé qu'on avait appliqué de façon non uniforme les facteurs d'évaluation à l'intérieur d'une

famille d'emplois donnée ou d'une famille à l'autre. Or nous n'avons trouvé aucune indication d'un tel manque d'uniformité. Nous croyons que le cadre des évaluations des postes-repères et la sélection des principaux postes-repères par le Comité directeur constituent une base solide pour l'évaluation des postes restants par les cinq sous-comités. Nous n'avons pas trouvé matière à inquiétude et nous appuyons la poursuite de l'étude tel qu'il a été prévu.

(pièce R-22, à la p. 8)

432. Sur les 103 questionnaires du Comité directeur réévalués par M. Wisner et examinés par M. Willis, dix ont été évalués différemment par les consultants et trois seulement de ces derniers étaient sensiblement différents. Sur les trois, un avait un écart plus marqué. M. Willis estimait que, s'il y avait lieu de reconvoquer le Comité directeur, ce ne devrait être que pour réexaminer ce questionnaire dont l'écart était plus marqué.

433. En septembre 1988, le côté patronal du Comité mixte était toujours insatisfait de la façon dont le plan Willis avait été appliqué par le Comité directeur et continuait d'exprimer des réserves au sujet du travail de ce dernier. A la réunion du 15 septembre 1988 du Comité mixte, la partie patronale a soutenu qu'il fallait analyser plus en profondeur des postes-repères problématiques mentionnés par M. Willis dans son rapport de juillet 1988 sur les évaluations du Comité directeur. Le côté patronal a relevé 100 postes-repères avec des problèmes et a soumis 46 d'entre eux à M. Willis avec une liste de questions, d'observations et d'anomalies. En réponse à la demande de la direction, Willis & Associates a mené un examen indépendant de ces questionnaires et a tenté de faire une évaluation «fraîche», sans tenir compte des évaluations antérieures effectuées par le Comité directeur, mais en appliquant la discipline générale d'évaluation établie par celui-ci.

434. Un rapport de ce travail a été soumis au Comité mixte en septembre 1988 (pièce R-28). Cette analyse a été exécutée par le consultant de M. Willis, M. Wisner. Willis & Associates a donné raison à l'employeur sur un certain nombre de questionnaires contestés, mais, au bout du compte, il n'a décelé aucune partialité fondée sur le sexe. A propos de la discipline adoptée par M. Wisner dans son évaluation indépendante des 46 questionnaires, M. Willis a fait les observations suivantes dans son témoignage (volume 56, à la p. 6936, lignes 3 à 11) :

[TRADUCTION]

Je pense qu'il connaissait assez bien les évaluations du Comité directeur à ce stade. Nous avons pu discuter ensemble des endroits où le Comité avait été conservateur et des points sur lesquels il avait été un peu libéral de sorte que [M.. Wisner] a pu faire le suivi nécessaire, mais de façon assez indépendante. Je dirais toutefois qu'en général il semble avoir été un peu plus libéral que le Comité directeur, mais à peine et sans que cela ne porte à conséquence.

105

435. En dernière analyse, MM. Willis et Wisner ont retenu sur les 46 une évaluation qui, à leur avis, avait été mal comprise par le Comité directeur. Dans son rapport, M. Willis a répondu explicitement à chacune des questions posées par la partie patronale. Selon les consultants, l'analyse supplémentaire avait confirmé le fait que le Comité directeur s'était très bien acquitté de sa tâche en appliquant le système d'évaluation à un large éventail de postes. Voici ce que dit le rapport :

[TRADUCTION]

Nous estimons qu'une base solide a été fournie pour l'évaluation des 3 900 postes restants et que, à ce stade-ci, il n'y a aucune raison logique de s'attendre à rien de moins qu'un résultat de haute qualité et défendable de l'étude.

(pièce R-28, à la p. 4 de l'addenda)

436. Dans ce rapport, M. Willis fait en outre des observations générales sur les sources éventuelles de divergences entre le Comité directeur et les consultants. Il explique que ces écarts peuvent naître de trois façons, chacune étant liée à un parti pris systématique de la part des évaluateurs. Il décrit ainsi ces trois sources de divergences :

[TRADUCTION]

3. Il y a trois façons qui peuvent expliquer comment le Comité directeur et les consultants peuvent en venir à évaluer différemment les mêmes postes :

Mauvaise interprétation du questionnaire. Cela peut se produire si on n'a pas tenu compte de certaines parties du questionnaire ou si on ne leur a pas accordé l'importance voulue.

Interprétations divergentes des faits donnés. Les consultants peuvent fonder leurs interprétations sur une expérience plus vaste de l'évaluation d'autres emplois ayant des responsabilités fonctionnelles semblables. Par contre, les membres des comités d'évaluation comprennent parfois mieux que les consultants la culture organisationnelle qui existe au sein de l'administration fédérale, ce qui donne lieu à un point de vue légèrement différent sur les emplois.

Mauvaise application ou interprétation du système d'évaluation. Normalement ce facteur n'entre en jeu qu'à l'étape où on apprend à évaluer.

Dans chacun de ces cas, la source de l'écart peut être un parti pris systématique de la part des évaluateurs.

(pièce R-28, aux pp. 1 et 2 de l'addenda)

106

437. A la fin de novembre 1988, les représentants patronaux faisant partie du Comité mixte étaient toujours insatisfaits de l'analyse que MM. Wisner et Willis avaient faite des évaluations du Comité directeur. Mme Ouimet a fait part de ses préoccupations à M. Willis dans une lettre de quatre pages (pièce HR-19). Ce dernier a tenté de répondre à ses interrogations dans une lettre de six pages datée du 5 décembre 1988. Nous reproduisons ci-dessous la partie de cette lettre dans laquelle il essaie de convaincre Mme Ouimet qu'une certaine variation entre les évaluateurs est inévitable et lui expose les raisons de cette variation. M. Willis dit ce qui suit :

[TRADUCTION]

Tolérance dans la variation des évaluations

Comme je l'ai affirmé dans l'addenda aux réponses que j'ai données à la partie patronale du Comité mixte syndical-patronal sur la mise en oeuvre de la parité salariale, il est normal qu'il y ait une certaine variation dans l'interprétation des informations sur les postes fournies aux évaluateurs. Une tolérance de plus ou moins 10 pour cent au niveau de la variation aléatoire des évaluations est acceptable entre deux équipes évaluant les mêmes postes et ayant en main des faits complets et exacts.

Sur le plan pratique, l'analyse et l'estimation de la fiabilité en matière d'évaluation nécessitent que l'on porte des jugements en tenant compte d'un certain nombre de facteurs variables, notamment :

- Les informations utilisées sont-elles complètes, concrètes et définitives? Plus l'information est déficiente, plus grand sera normalement le biais aléatoire.
- La nature de l'emploi. Est-ce un poste exceptionnel ou complexe, ou un poste que les évaluateurs ne devraient normalement pas avoir de difficulté à comprendre (p. ex., un chercheur scientifique ou un préposé au nettoyage)? Pour évaluer un poste comme il faut, l'évaluateur doit pouvoir en comprendre le contenu.
- Comment éloigné, sur le plan hiérarchique, le poste se trouve-t-il de l'expérience ou des connaissances des évaluateurs? Ce facteur est semblable au précédent en ce sens qu'un évaluateur peut avoir de la difficulté à conceptualiser un emploi qui est plusieurs niveaux organisationnels au-dessus de sa propre expérience.
- Peut-on déceler une tendance dans les variations des évaluations? Semble-t-il y avoir un biais systématique, ou s'agit-il d'une variation aléatoire? Un parti pris ou biais systématique est beaucoup plus significatif qu'une variation qui est simplement due à une différence dans

107

l'interprétation ou la compréhension des exigences d'un emploi.

- Si les évaluations des comparaisons ont été faites par un consultant, l'écart peut-il s'expliquer par le fait qu'il aurait perçu différemment la culture ou le système de valeurs de l'organisation, d'où son point de vue légèrement différent sur les emplois en question?

Essentiellement, il faut porter un jugement de valeur lorsqu'on doit décider quelle marge de variation on est prêt à accepter dans les cotations et s'il existe un problème ou non. Une analyse de cette nature ne peut s'effectuer en des «termes précis et quantitatifs».

Sur les quatorze évaluations du Comité directeur qui s'écartaient de plus de 10 pour cent des évaluations des consultants, j'estime qu'une seule, celle du poste no 428 (Chef, Section de la préparation des expositions), n'a pas été comprise du Comité directeur et devrait être soumise de nouveau à ce dernier avec des questions pour qu'il réévalue le poste.

(pièce R-35, aux pp. 3-4)

438. En dernière analyse, les représentants patronaux du Comité directeur n'ont pas appuyé sans réserve les évaluations des postes-repères réalisées par ce dernier. S'ils étaient prêts à poursuivre l'étude, leur intention était toutefois de procéder à d'autres examens des évaluations des postes-repères. Or l'employeur n'a pas soulevé cet autre examen dans la présentation de sa preuve.

439. Le Tribunal a entendu des témoignages limités de la part des évaluateurs à propos des évaluations contestées du Comité directeur. Aux yeux de Mme Pauline Latour, une des évaluatrices qui a témoigné devant nous sur la contestation des postes-repères par les comités, cette contestation revêtait peu d'importance. Un seul poste-repère a causé de la difficulté à son comité (le comité no 5). Ce dernier était d'avis que le Comité directeur avait coté le poste en question plus haut qu'il aurait dû le faire (volume 171, aux pp. 21641-43).

(ii) Application des tests de fiabilité inter-évaluateurs (FIE)  
aux comités d'évaluation multiples

440. M. Shillington a aussi appliqué des tests FIE aux comités restants (les cinq initiaux et les neuf subséquents) en utilisant les mêmes méthodes que celles qu'il avait employées pour identifier les évaluateurs marginaux au sein du Comité directeur. M. Willis a reçu deux rapports sur les tests FIE qui ont été appliqués aux comités susmentionnés. C'est en mai 1988 qu'il a reçu les premiers résultats, qui reposaient principalement sur les cinq premiers comités d'évaluation. Les deuxièmes résultats lui sont parvenus en juillet 1989 et étaient fondés sur les neuf comités élargis.

441. A sa réunion du 25 août 1989, le sous-comité FIE a fait savoir au Comité mixte que l'analyse des cotes individuelles pour la période clôturant à la fin de juillet 1989 avait permis de découvrir 11 évaluateurs marginaux ou produisant des cotes aberrantes : six femmes du côté syndical, trois femmes du côté patronal et deux hommes du côté patronal. Sept

d'entre eux ont exprimé une préférence apparente envers les postes à prédominance masculine et quatre ont manifesté une préférence apparente pour les postes à majorité féminine.

442. Le sous-comité a par ailleurs signalé que sept des évaluateurs marginaux avaient déjà été repérés dans les premiers résultats. Toutefois, M. Willis, dans son témoignage, a pu se rappeler huit de ces évaluateurs qui avaient déjà été identifiés. Le sous comité FIE avait signalé dans les deuxièmes résultats l'identification des évaluateurs marginaux précédemment repérés afin de confirmer l'opinion de M. Willis quant à l'inefficacité de ses interventions et conseils après la communication des premiers résultats.

443. Le Comité mixte a décidé que les noms des évaluateurs marginaux ne seraient révélés qu'à M. Willis et au chef du SEPS. M. Willis croyait savoir que le Comité mixte avait décidé de garder le nom de ces personnes secret afin de les protéger. Le Comité mixte avait précédemment décidé qu'il n'allait retirer aucun évaluateur des comités et qu'il ne serait pas productif de divulguer leur nom à ce moment-là.

444. M. Shillington a préparé certaines pièces décrivant ce qu'il a appelé la «dimension attitudinale de ces évaluateurs marginaux». Il était incapable d'expliquer pourquoi ces différences survenaient ou ce qu'elles étaient. Les pièces HR-117 et HR-133 montrent que les préférences à l'égard des postes à prédominance masculine ou féminine ne sont pas liées au fait d'être associé au côté syndical ou patronal ni au fait d'être un homme ou une femme. En ce qui concerne ce dernier cas, certaines évaluatrices ont manifesté une préférence envers les postes masculins, tandis qu'aucun évaluateur de sexe masculin n'a exprimé de préférence pour les postes à majorité féminine.

(iii) Tests de fiabilité inter-comités (FIC)

445. M. Willis a témoigné que les tests de fiabilité inter-comités (tests FIC) visaient à déterminer si les évaluations d'une série de comités étaient reliées. Tel qu'il l'a expliqué, ces tests permettent de faire le point sur l'uniformité entre les comités et d'identifier les comités qui ont besoin de recyclage. Les tests FIC, a-t-il fait remarquer, ne sont pas conçus pour détecter la partialité. Dans le contexte de l'étude sur la parité salariale, ils visaient à déterminer si les comités d'évaluation s'adaptaient avec succès à la discipline du Comité directeur.

(iv) Application des tests FIC aux comités d'évaluation multiples

446. Le processus a consisté en général à prendre une série de questionnaires et à soumettre chaque questionnaire à chacun des comités d'évaluation. Chaque comité a ensuite évalué le même questionnaire et le consultant a tenté de déterminer dans quelle mesure les différents comités avaient coté le même emploi de façon semblable ou différente. Selon M. Willis, les premiers tests FIC ont eu lieu au début de 1989 et comprenaient en tout 26 tests. Ces tests se sont poursuivis jusqu'en juillet 1989.

447. Le Comité mixte a créé un sous-comité des tests de fiabilité inter-comités (le «sous-comité FIC») qu'il a chargé d'établir la politique et de superviser les procédures pour les tests. Le sous-comité FIC consistait en trois représentants patronaux, deux représentants syndicaux, M. Willis, un de ses consultants et deux représentants de la Commission. Le mandat du sous-comité a été formulé comme suit dans le rapport que le sous-comité FIE a présenté le 3 mars 1989 :

[TRADUCTION]

- examiner les résultats des tests administrés aux comités d'évaluation par rapport à la référence fournie par les consultants;
- examiner la cote de référence fournie par les consultants;
- déterminer les différences significatives dans les cotes de consensus des comités par rapport aux postes-repères et à la référence;
- recommander, si nécessaire, la formation et le recyclage que le consultant devrait donner ainsi que toute autre ligne de conduite que le Comité mixte devrait envisager;
- cerner les problèmes liés à la procédure ou au processus ainsi que les possibilités d'amélioration, notamment les révisions qu'il y aurait lieu d'apporter à la formulation des justifications raisonnées.

448. Le sous-comité des tests de fiabilité inter-évaluateurs (FIE) a demandé à la Commission d'effectuer les tests proprement dits. C'est donc la Commission qui a déterminé le moment où seraient exécutés les tests, qui a distribué les questionnaires et qui a expliqué le processus aux comités. Le Comité mixte a demandé à M. Willis d'évaluer les questionnaires prévus pour les tests et de fournir une «cote de référence ou de base» pour chacun des emplois sur lesquels porteraient les tests.

449. La cote de référence était l'évaluation indépendante, faite par les consultants, des questionnaires compris dans le test. Dans chaque cas, M. Willis faisait revoir le questionnaire par deux consultants qui devaient évaluer le poste de façon indépendante, leurs évaluations étant ensuite comparées avec celles effectuées par les cinq ou les neuf comités. L'objet

110

de la comparaison entre la cote de référence et la cote du comité était de repérer tout écart entre, premièrement, les comités individuels et, deuxièmement, entre le consensus des comités d'évaluation multiples et les évaluations des consultants. On se trouvait ainsi à cerner les domaines où les comités d'évaluation multiples avaient besoin de recyclage à cause de leur difficulté à interpréter les facteurs d'évaluation.

450. M. Willis a utilisé des justifications raisonnées dans les tests FIC afin d'analyser les écarts entre les cotes de référence et les cotes des comités. L'application qu'il a faite des justifications raisonnées des comités dans ces tests visait un objectif différent de l'utilisation qu'il en a faite en général dans les comités d'évaluation. M. Willis a expliqué pourquoi, dans les tests FIC, il fallait en l'occurrence utiliser les justifications raisonnées, ces raisons étant différentes des motifs pour lesquels il ne voulait pas qu'on les applique dans les évaluations effectuées par les comités; dans ce dernier cas, il voulait que les membres se concentrent sur le questionnaire même.

451. La cote de référence du consultant était comparée à la cote de consensus de chaque comité ainsi qu'à la cote de consensus globale des cinq -- et plus tard des neuf -- comités d'évaluation.

452. M. Willis a eu peu à dire sur la procédure adoptée par le sous-comité FIC et il était opposé à son approche. Dans d'autres études, il fournissait toujours une liste des questionnaires à ses clients, puis il insérait le questionnaire dans le dossier de questionnaires de telle façon que les comités ignoraient quels questionnaires faisaient partie du test. En ce qui concerne l'étude sur la parité salariale, on réservait un moment spécial où l'on distribuait les questionnaires aux évaluateurs, qui devenaient alors conscients du test. La Commission a choisi les questionnaires au hasard et a approché les consultants de M. Willis environ une heure avant le test pour les laisser sélectionner les questionnaires qui devraient faire partie du test. Les consultants n'ont pas eu la possibilité de sélectionner des questionnaires qui étaient plus complets. Selon M. Willis, cela a frustré les évaluateurs, qui n'ont d'ailleurs pas tous rempli les tests avec le même degré de sérieux.

453. La procédure suivie pour les tests FIC par la Commission était très rigoureuse au départ. L'administration d'un test était d'abord annoncée. La Commission était sur place pour superviser le test et avait un observateur dans chaque pièce. On distribuait les questionnaires et on informait les évaluateurs qu'ils ne pouvaient pas quitter la pièce durant le test proprement dit.

454. Dans d'autres études, si les comités avaient besoin de plus de renseignements sur les questionnaires du test, ils pouvaient prendre le téléphone et poser leurs questions à des personnes que M. Willis avait spécialement affectées à cette tâche. Pour les tests FIC, rien de tel n'était prévu. Par conséquent, chaque comité pouvait faire ses propres hypothèses et compléter lui-même les informations manquantes. Les comités devaient mettre leurs hypothèses par écrit, mais le problème était que chaque comité énonçait des hypothèses différentes, ce qui, d'expliquer

111

M. Willis, a entraîné des variations dans les évaluations. Pour ces raisons, M. Willis avait des réserves à propos des résultats des tests FIC.

455. Aux yeux de M. Willis, les comités ne prenaient pas ces tests aussi sérieusement que leurs évaluations effectives. Il a remarqué beaucoup de rancœur chez les évaluateurs, sentiment qui s'est accru avec le temps. De plus, les comités étaient obligés d'interrompre leurs évaluations normales pour évaluer les questionnaires du test. M. Willis exerçait des pressions sur les comités pour qu'ils maintiennent leur rythme; il y a eu des moments, toutefois, où les comités subissaient deux tests par semaine. Dans son témoignage, M. Willis a fait le commentaire suivant à ce sujet (volume 58, à la p. 7166, lignes 13-16) :

[TRADUCTION]

Ils détestaient cela à chaque étape du processus et certains d'entre eux se prêtaient souvent aux tests avec une attitude qui n'était pas des plus sérieuses.

456. Le 6 février 1989, M. Willis a produit un rapport sur les neuf premiers tests exécutés entre le 7 novembre 1988 et le 5 janvier 1989. Il y examine la variation parmi les cinq comités d'évaluation initiaux et conclut essentiellement que ces derniers ont appris le système Willis et évaluent les postes conformément à la discipline du Comité directeur lorsqu'ils «se sentent à l'aise avec cette discipline».

457. Le rapport que le sous-comité FIC a présenté le 3 mars 1989 était fondé sur l'analyse des 11 premiers tests effectués. Il y était noté ce qui suit : (i) les consultants ont dû réviser le programme de formation initiale avec les comités et examiner les problèmes mis en évidence; (ii) la comparaison de familles d'emplois recoupées a suscité certaines préoccupations et le processus d'évaluation des emplois devrait être modifié en conséquence; et (iii) il devrait être accordé plus de soin à la justification raisonnée et l'on devrait mettre au point un processus d'évaluation des emplois révisé.

458. On a demandé à M. Willis de décrire le degré de variation entre les cotes des consultants et le consensus des cinq comités pour les 11 premiers tests. Il a répondu ce qui suit (volume 58, à la p. 7227, lignes 1-9) :

[TRADUCTION]

R. Compte tenu des divers handicaps ainsi que des frustrations et préoccupations qu'on nous a exprimées, je pense qu'ils ont très bien fait. J'étais très positif, particulièrement après avoir discuté avec chaque comité de leurs différences, des raisons pour lesquelles ils avaient choisi les hypothèses qu'ils avaient choisies. Même si je reconnaissais que le recyclage additionnel était souhaitable, j'étais très satisfait de la façon dont ils s'acquittaient de leur tâche.

112

459. Le sous-comité FIC a joint à son rapport une description d'un Processus d'évaluation améliorée dont il a recommandé l'adoption par les comités. Le processus révisé prévoyait la comparaison de postes-repères en dehors de la famille d'emplois. On y demandait aux évaluateurs de citer d'abord les postes-repères par rapport aux cotes attribuées aux facteurs avant de transmettre les cotes indépendantes au président du comité pour qu'il les affiche.

460. M. Willis a témoigné que les comités avaient essayé le processus amélioré d'évaluation des emplois et qu'ils ne l'avaient pas trouvé vraiment pratique; en fait, il nécessitait plus de temps que le processus initial. Les comités d'évaluation ayant résisté au changement, on l'a finalement abandonné.

461. M. Willis a aussi produit un rapport sur les tests FIC. Le rapport du sous-comité était en général semblable au sien, a-t-il dit, mais M. Willis ne souscrivait pas à toutes les conclusions du sous-comité.

462. Essentiellement, les tests FIC visaient à déterminer si les comités interprétaient et appliquaient le plan Willis conformément à la discipline du Comité directeur. Ces tests fournissaient également aux consultants l'occasion de vérifier si les motifs invoqués par les comités pour justifier leurs cotes reflétaient une partialité fondée sur le sexe.

463. M. Willis a examiné soigneusement, facteur par facteur, ce que chaque comité avait fait, pourquoi il l'avait fait, et comment il avait atteint le consensus pour chaque facteur et pour le total. Lorsque la cote d'un comité différait de celles des autres comités, il explorait les raisons de l'écart. Si ces raisons donnaient à croire que, d'une façon ou d'une autre, elles étaient influencées par un sexe ou par l'autre ou encore par un genre d'emploi en particulier, le consultant pouvait utiliser cette information pour y faire le suivi nécessaire.

464. Dans le contexte de ces tests, on a demandé à M. Willis s'il avait trouvé une indication quelconque de préférence pour un sexe ou pour l'autre dans le travail des comités durant la première série des tests FIC. Voici ce qu'il a répondu (volume 58, de la p. 7227, ligne 10, jusqu'à la p. 7228, ligne 19) :

113

[TRADUCTION]

Q. A ce moment-là -- et nous sommes, semble-t-il, à la fin de l'hiver ou au début du printemps de 1989, lorsque seulement la moitié des tests avaient été administrés -- aviez-vous des indications d'après ces tests, ou autrement, que les évaluateurs aient pu manifester une préférence pour un sexe ou pour l'autre?

R. Il n'y a pas de rapport avec les tests FIC, mais au début de l'année j'ai commencé à exprimer des réserves, qui n'étaient pas reliées directement aux évaluations proprement dites, mais je m'inquiétais de faits circonstanciels qui s'étaient produits. J'étais devenu de plus en plus mal à l'aise concernant la façon dont les comités travaillaient, à propos notamment des affrontements entre les côtés syndical et patronal ainsi que de certaines choses que j'avais observées.

Après avoir discuté avec mes consultants qui travaillaient avec les groupes, et après avoir analysé nous-mêmes la façon dont les comités évaluaient dans les faits ainsi que les résultats que les groupes professionnels obtenaient au sein des comités et d'un comité à l'autre, je n'ai pu déceler rien de précis qui aurait

donné à croire qu'il se développait une partialité fondée sur le sexe.

Néanmoins, j'avais de forts sentiments contradictoires parce que je savais qu'il se passait des choses, que certaines attitudes étaient manifestes -- ce qui n'était pas sans me préoccuper beaucoup.

Ainsi, à ce stade de l'étude j'avais certains problèmes quant à ma propre tranquillité d'esprit. J'ai discuté de ces problèmes individuellement et avec les membres du mini-comité directeur, et collectivement avec eux comme groupe. J'estimais qu'il me faudrait faire une analyse quelconque, une analyse plus approfondie des résultats, avant de pouvoir appuyer le résultat de l'étude.

465. M. Willis ne croyait pas qu'il était possible de détecter la partialité fondée sur le sexe simplement en examinant les résultats des tests FIC. Lorsqu'une telle partialité existe, selon M. Willis, les évaluateurs parlent en général de leurs conclusions ou opinions au lieu des faits contenus dans le questionnaire. Il a donné instruction à ses consultants de surveiller attentivement ce genre de comportement, mais il ne croit pas qu'un consultant puisse décider s'il y a partialité ou non en examinant simplement une cote, emploi par emploi. Dans son témoignage, il a expliqué que, pour une évaluation donnée, le consultant doit examiner les raisons pour lesquelles les comités ont fait les choix qu'ils ont faits et ce qui est dit dans la justification raisonnée, puis il doit interroger personnellement les évaluateurs en vue de cerner les facteurs qui expliquent les écarts. De l'avis de M. Willis, ces tests n'ont fourni aucune preuve concluante de partialité fondée sur le sexe et on ne devrait

pas tenir compte des informations obtenues des tests parce que les comités n'ont pas pris les tests au sérieux.

466. A la fin de mai ou au début de juin 1989, M. Willis a recommandé au sous-comité FIC de mettre fin aux tests parce qu'il était devenu très clair à ses yeux que les comités d'évaluation étaient de plus en plus frustrés par cette procédure. Il a aussi conclu que les tests ne pouvaient plus être d'aucune utilité. M. Willis croyait savoir que c'était sur l'insistance du représentant du Conseil du Trésor siégeant au sous-comité que les tests se poursuivaient. Le sous-comité n'a pas accepté sa recommandation et a poursuivi les tests jusqu'en juillet 1989. M. Willis estimait que la réaction à ces tests qu'il avait observée dans les comités

d'évaluation pouvait influencer sur la fiabilité des résultats (volume 59, à la p. 7291).

467. Bien que les tests se soient poursuivis, M. Willis n'a plus comme tel analysé les résultats. Il a examiné les tests restants que le sous-comité lui a remis et il a continué de rencontrer individuellement les comités.

468. Une version préliminaire du rapport final des 26 tests FIC a été rédigée par un des représentants patronaux faisant partie du sous-comité chargé de ces tests, M. Michel Papineau. Ce rapport est daté du 26 octobre 1989. M. Willis n'a pas participé à la rédaction du rapport, dont voici la conclusion :

[TRADUCTION]

Les résultats des tests tendent à étayer les préférences fondées sur le sexe décelées dans le rapport FIE et dans l'étude du consultant portant sur l'échantillon de 200 questionnaires déjà évalués par les comités. Les écarts sont tels qu'il fait peu de doute quant à savoir s'ils sont attribuables à des biais systématiques ou aléatoires. La proportion de ces écarts est assez significative pour dépasser le seuil de tolérance établi par le consultant. Il est donc fortement recommandé de mener d'autres études avant de tirer des conclusions à partir des résultats des évaluations.

(pièce HR-90, à la p. 4)

469. M. Papineau a conclu qu'il y avait preuve de partialité fondée sur le sexe dans les évaluations, mais M. Willis estimait pour sa part qu'on ne devait pas tenir compte de l'analyse des tests FIC pour deux raisons. Premièrement, les 26 évaluations constituaient un nombre trop restreint pour qu'on puisse en tirer des conclusions fermes; deuxièmement, les comités n'avaient pas pris les tests aussi sérieusement que les évaluations réelles et ils les avaient réalisés aussi rapidement qu'ils le pouvaient sans trop de discussion. M. Willis croit que les tests «n'avaient plus aucune utilité particulière après les 10 ou 12 premiers tests» (volume 59, à la p. 7297).

470. Pour ce qui est de l'affirmation du rapport selon laquelle d'autres études devaient être effectuées, M. Willis a témoigné qu'il avait décidé, sur la foi de la Wisner 222, qu'une étude plus approfondie

s'imposait et que cette première version du rapport FIC n'avait rien ajouté à sa conviction.

471. L'étude Wisner 222, d'expliquer M. Willis, n'avait aucun rapport avec les tests FIC. Il aurait réclamé la Wisner 222 que le Comité mixte ait donné ou non son aval aux tests de fiabilité, a-t-il ajouté. A ses yeux, il s'agissait de deux questions entièrement séparées.

472. Mme Elizabeth Millar, membre syndicale du sous-comité FIC et chef de la section Classification et égalité de salaire à l'Alliance, a témoigné avoir eu l'impression que les comités prenaient très au sérieux les tests FIC. Une des difficultés du sous-comité FIC, a-t-elle signalé, était de communiquer son feedback aux comités en temps utile. A son avis, le sous-comité FIC n'a pas fonctionné de manière efficace après mai 1989. Les représentants patronaux faisant partie du sous-comité ont semblé poursuivre des buts différents de ceux du reste du comité, selon elle. Ces représentants souhaitaient qu'on intensifie le rythme des tests, au lieu de le réduire, jusqu'à la fin du processus d'évaluation.

473. Par une note de service datée du 10 novembre 1989, Mme Millar a répondu au rapport préliminaire rédigé par M. Papineau. Essentiellement, elle trouvait le rapport inacceptable à l'Alliance parce qu'il ne reflétait pas les discussions et délibérations qui avaient eu lieu au sein du sous-comité FIC. L'analyse contenue dans le rapport ne concordait pas avec les constatations du comité et la conclusion de M. Papineau n'avait jamais été discutée par le sous-comité FIC. Dans son témoignage, M. Willis a convenu avec Mme Millar que la conclusion contenue dans l'ébauche du rapport allait peut-être trop loin. M. Papineau y laissait entendre que, à la lumière de la comparaison, l'existence d'une partialité fondée sur le sexe faisait peu de doute (volume 59, à la p. 7304).

474. M. Papineau, dans la note de service jointe au procès-verbal, fait part de son intention de déposer le rapport à la réunion suivante du Comité mixte, qui a eu lieu le 31 octobre 1989 (pièce R-44). Le rapport final n'a toutefois pas été déposé à ce moment-là, car le rapport [provisoire] n'avait été distribué qu'une semaine auparavant.

475. Mme Millar a témoigné que, au printemps de 1989, elle a remarqué un changement d'attitude de la part des représentants patronaux membres du sous-comité FIC à l'égard des consultants. Avant mai 1989, a-t-elle expliqué, l'attitude de l'employeur était plus accommodante et en harmonie avec les vues du consultant, de sorte que le sous-comité réussissait à s'entendre sur des points litigieux. Chacun reconnaissait que les comités d'évaluation avaient de la difficulté à comprendre le plan Willis et qu'ils avaient besoin d'aide supplémentaire sous forme de formation. Après mai 1989, les représentants de l'employeur sont devenus très critiques à

l'endroit des consultants et les réunions du sous-comité FIC sont devenues extrêmement houleuses. Mme Millar a relaté ce qui s'était passé à une réunion en particulier à laquelle M. Scott Gruber, un consultant de M. Willis, faisait état de l'un des tests qui avaient été faits. M. Gruber

116

avait rencontré tous les comités pour discuter des résultats et avait trouvé que les travaux progressaient bien. Les représentants du Conseil du Trésor ont contesté le rapport de M. Gruber. Selon Mme Millar, un représentant de l'employeur a fait un commentaire comme quoi le comité aurait dû s'attendre à quelque chose de mieux de la part des consultants.

476. Dans son témoignage, Mme Millar a relaté cet autre incident (volume 185, de la p. 23775, ligne 17, jusqu'à la p. 23776, ligne 11) :

[TRADUCTION]

A une réunion où M. Owen était le consultant, deux représentants du Conseil du Trésor ont produit des rapports dont nous ignorions qu'ils étaient en préparation. Ils avaient calculé l'écart entre la cote de chaque comité et la cote de référence, puis ils avaient utilisé ces calculs pour indiquer s'il existait un problème ou non.

M. Owen -- que j'ai décrit comme quelqu'un d'une grande politesse et gentillesse, sans compter sa très grande compétence -- est devenu extrêmement agité. Il a lancé son crayon sur le bureau et accusé les représentants du Conseil du Trésor de ne rien comprendre à l'évaluation des emplois ou au plan Willis. M. Willis m'a plus tard dit qu'il avait longtemps travaillé avec Fred Owen et que jamais il ne l'avait vu aussi fâché. Il va sans dire que ces rapports, les rapports qui n'avaient pas été commandés, n'ont jamais été acceptés par le sous-comité et n'ont jamais plus été ramenés sur le tapis.

M. Owen n'a pas été interrogé à propos de cet incident.

477. M. Willis a témoigné que les tests FIC n'avaient pas répondu à ses attentes. A l'avenir, il ferait faire ces tests de façon cachée ou secrète afin que les évaluateurs ne sachent pas qu'ils font l'objet d'épreuves. Par contre, il a fait les observations suivantes sur les résultats de ces tests (volume 59, à la p. 7352), lignes 2-7) :

[TRADUCTION]

En fin de compte, on peut affirmer, semble-t-il, qu'en dépit de l'absence d'appui de la part de l'employeur, en dépit de certaines variations dans la qualité des informations et en dépit de certains problèmes d'attitude, le résultat était dans des limites satisfaisantes.

478. On a demandé par ailleurs à M. Willis si, dans les 11 premiers tests FIC, il était possible de dire si les comités attribuaient des cotes supérieures en fonction du sexe du répondant. D'après lui, il ne semblait pas y avoir de parti pris fondé sur le sexe. Tout écart d'interprétation entre les consultants et les comités relativement au facteur «résolution de problèmes» dans le plan Willis tenait davantage au fait que des évaluateurs ne comprenaient pas clairement comment appliquer le système d'évaluation, qu'à n'importe quel autre facteur.

117

(v) Réévaluations Wisner

479. M. Willis a témoigné qu'il était clair à ses yeux que, tant du côté syndical que du côté patronal, il y avait un parti pris qui influait sur la façon dont les évaluateurs travaillaient ensemble. Il a observé un problème d'attitude chez certains des évaluateurs. La situation l'inquiétait de plus en plus à mesure que progressait l'étude, au point où il n'était plus sûr s'il pouvait défendre les résultats sans effectuer lui-même certaines analyses additionnelles. Ce n'était pas qu'il avait observé comme tel de la partialité fondée sur le sexe dans les évaluations. Ses préoccupations découlaient avant tout de ce qu'il considérait comme un problème d'attitude chez les évaluateurs. Comme il s'agissait d'une étude d'envergure et importante, il voulait s'assurer qu'aucune partialité subtile ne se glissait dans le processus.

480. M. Willis a recommandé au Comité mixte de mener une évaluation ponctuelle de la validité des évaluations, dans l'intention de procéder à une analyse plus en profondeur si jamais l'analyse préliminaire soulevait la possibilité de problèmes. Lorsqu'il a fait sa proposition au Comité mixte au printemps de 1989, il ne lui a pas dit qu'il avait l'intention d'adopter une approche à deux volets si la première étude révélait un problème. Il a fait sa recommandation au Comité mixte à peu près au moment où l'on venait de terminer les 11 premiers tests FIC. A ce moment-là, les comités avaient évalué environ 2 000 questionnaires, et M. Willis souhaitait faire examiner un échantillon de 10 % des évaluations achevées de façon indépendante par un de ses consultants.

481. M. Willis a témoigné que le seul moyen qu'il connaissait de déterminer si l'évaluation d'un évaluateur était entachée de partialité fondée sur le sexe était de chercher des tendances traduisant une préférence pour un sexe ou pour l'autre. A son avis, la seule façon de détecter la partialité fondée sur le sexe aurait été de faire appel à un tiers impartial, tel un de ses consultants, de lui faire réévaluer des questionnaires choisis, puis de comparer les résultats entre les comités et les consultants. M. Willis demande habituellement l'assistance d'un statisticien pour effectuer cette comparaison. Il appelle disparités les écarts entre les comités et les consultants.

482. Pendant l'audience, on a soulevé la question de savoir si les consultants devaient servir de point de référence pour la comparaison. M. Willis a fait remarquer que le Comité mixte avait convenu d'utiliser les consultants comme point de référence dans les études FIC, entente que le Comité avait formulée par écrit.

483. M. Willis estime que ses consultants qui ont participé à l'étude sur la parité salariale étaient impartiaux. Il a fait les commentaires suivants sur ce point (volume 208, à la p. 26934, lignes 10-16) :

[TRADUCTION]

Nous comprenons le système. Je ne pense pas que nous puissions affirmer que tous les consultants sont nécessairement impartiaux. Cependant, -- et notre expérience, nos antécédents, notre volonté

118

et notre philosophie en font foi -- nous avons toujours eu comme ligne de conduite de ne pas favoriser un côté ou l'autre, mais de rester neutres en somme.

484. Ce sont les disparités, d'expliquer M. Willis, qui permettent de conclure s'il existe une partialité fondée sur le sexe. Dans ce contexte, a-t-il ajouté, il y a «partialité» si l'on peut déceler une tendance à traiter différemment les emplois à prédominance masculine par opposition aux emplois à prédominance féminine, tendance qui se traduirait jusqu'à un certain point en parti pris fondé sur le sexe.

485. Les postes inclus dans l'étude Wisner 222 ont été choisis au hasard à partir d'une liste de toutes les évaluations fournies par le SEPS. L'échantillon prélevé comprenait au moins 10 % du nombre total de postes évalués par les neuf comités d'évaluation à l'époque de la Wisner 222. L'échantillon était représentatif de l'éventail complet des niveaux

d'évaluation et de la diversité des types de travail vus par les neuf comités.

486. M. Wisner n'a pas témoigné à l'audience. C'est M. Willis qui a décrit la méthode que M. Wisner avait utilisée dans son analyse. Premièrement, M. Wisner a lu le questionnaire sur le poste et les notes du réviseur, le cas échéant. Il a ensuite vérifié s'il existait un poste semblable parmi les postes-repères évalués par le Comité directeur. Lorsqu'un tel poste existait, M. Wisner examinait le questionnaire du poste-repère pour confirmer son impression et adoptait l'évaluation du poste-repère effectuée par le Comité directeur comme évaluation du consultant. En l'absence de fonctions semblables parmi les postes-repères, il effectuait une évaluation indépendante du poste, en s'appuyant au besoin sur les postes-repères appropriés. Un grand nombre des postes inclus dans l'échantillon ont nécessité cette étape. Après avoir lui-même évalué un poste, M. Wisner examinait l'évaluation que le comité en avait faite. Il surveillait plus particulièrement l'utilisation que le comité avait faite des postes-repères et examinait les faits qu'il avait invoqués à l'appui de son évaluation. M. Wisner rajustait ensuite son évaluation pour qu'elle soit appropriée à la lumière de la justification raisonnée du comité et des renvois aux postes-repères.

487. Lorsque M. Wisner constatait des différences entre son évaluation finale et l'évaluation du comité, il rédigeait une courte justification à l'appui de sa position.

488. M. Wisner a ensuite effectué une analyse spéciale des résultats. Celle-ci avait pour objet d'apprécier la qualité des évaluations des postes réalisées par les neuf comités d'évaluation. Tel qu'il l'a expliqué dans son rapport de juillet 1989, il a tenu compte des points suivants pour déterminer la «qualité» des évaluations :

[TRADUCTION]

1. L'utilisation appropriée du système d'évaluation Willis conformément au guide de mesure des postes ainsi qu'aux notes de formation et techniques publiées par le consultant.

2. La conformité des évaluations effectuées par les neuf comités avec les évaluations des postes-repères et la discipline d'évaluation établie par le Comité directeur.

3. L'absence de tout biais systématique dans les évaluations de la part des neuf comités.

(pièce PSAC-3, à la p. 1)

489. L'analyse de M. Wisner comportait aussi des tests statistiques. Selon M. Willis, M. Wisner est un statisticien. En ce qui concerne l'application du système d'évaluation Willis, M. Wisner n'a trouvé aucune indication, à deux exceptions possibles près, comme quoi on aurait de façon constante mal interprété ou appliqué les facteurs et dimensions d'évaluation. Pour ce qui est des deux exceptions, M. Willis a signalé que le nombre de postes échantillonnés était tellement faible qu'il était impossible d'affirmer quoi que ce soit de concluant.

490. En ce qui concerne la conformité globale des évaluations effectuées par les neuf comités avec la discipline d'évaluation du Comité directeur, M. Wisner a découvert que le comité et le consultant avaient attribué la même cote pour 70 des 220 postes, et que l'écart pour 34 des autres postes était de l'ordre de +/-2,5 %, de sorte que près de 47 % des postes de l'échantillon avaient à peu près la même évaluation globale. Il a conclu que les différences indiquaient une uniformité d'évaluation appréciable entre les neuf comités et les postes-repères du Comité directeur. Comme W. Wisner a constaté que plus de la moitié des postes différaient par plus de 2,5 %, il a recommandé une analyse plus approfondie des écarts.

491. Pour ce qui est du troisième point, en analysant les écarts entre les évaluations du consultant et celles des comités, M. Wisner a constaté que pour les postes à prédominance féminine, 35 avaient été sous-estimés par rapport à l'évaluation du consultant, 40 avaient été surestimés et 43 n'accusaient aucun écart; quant aux postes à prédominance masculine, 55 avaient été sous-estimés, 22 avaient été surestimés et 27 ne marquaient aucune différence. Dans son rapport, M. Wisner précise à la p. 5 :

[TRADUCTION]

Cela indique que les postes à prédominance féminine ont été surestimés un peu plus souvent que dans l'échantillon total, et que les postes à prédominance masculine ont été sous-estimés un peu plus souvent que dans l'échantillon total.

492. Et il conclut ce qui suit à la p. 8 :

[TRADUCTION]

Les conclusions de l'analyse décrite plus haut montrent que la conformité des évaluations effectuées par les neuf comités avec les postes-repères du Comité directeur laisse à désirer, et que les résultats des évaluations pourraient être entachés d'une

120

certaine partialité fondée sur le sexe. Le consultant est d'avis que, à la lumière de ces conclusions, une étude plus vaste des évaluations des neuf comités serait indiquée. Une telle étude permettrait de confirmer ou de réfuter les problèmes apparents décelés dans l'échantillon des postes examinés dans la présente étude. [c'est nous qui soulignons]

493. Selon M. Wisner, cependant, il faut interpréter ce rapport avec prudence. L'analyse statistique de la dominance d'un sexe sur l'autre et des écarts d'évaluation entre le comité et lui-même est fondée sur un nombre relativement restreint de postes, et ses conclusions «sur l'écart entre deux variables ne signifie pas qu'il y a eu une partialité délibérée ou inconsciente dans les évaluations». Les écarts peuvent s'expliquer par de nombreux autres facteurs, a-t-il poursuivi. Par exemple, les postes faisant partie des classifications à prédominance masculine comportent en général des fonctions et responsabilités plus complexes que la majorité des postes compris dans les classifications essentiellement féminines. Les écarts observés dans les évaluations pourraient s'expliquer par le fait que les comités aient eu tendance à sous-estimer les postes plus complexes par rapport à la discipline du Comité directeur telle qu'interprétée par le consultant.

494. Au troisième paragraphe de la lettre datée du 17 juillet 1989 qui accompagnait le rapport de M. Wisner et qui était adressée aux coprésidents du Comité mixte, M. Willis affirme ce qui suit :

[TRADUCTION]

Nos analyses révèlent l'existence d'une certaine divergence systématique par rapport aux évaluations du Comité directeur. Statistiquement, toutefois, la taille de l'échantillon examiné, 222 évaluations, était insuffisante pour que nous puissions tirer des conclusions précises sur l'ampleur du problème.

(pièce PSAC-4)

495. A l'audience, on a demandé à M. Willis de préciser ce qu'il voulait dire au juste dans sa lettre. Il a répondu ce qui suit (volume 58, à la p. 7249, lignes 1-5) :

[TRADUCTION]

R. Les résultats de notre analyse semblent révéler une certaine déviation systématique par rapport aux évaluations du Comité directeur. On pourrait l'interpréter comme une partialité fondée sur le sexe.

496. A la fin de l'étude Wisner 222, il restait encore mille évaluations environ à faire. Comme les neuf comités d'évaluation venaient à peine de commencer leur travail, M. Willis estimait qu'il était très important de faire une étude plus approfondie le plus tôt possible afin de corriger les problèmes éventuels. Il a recommandé au Comité mixte qu'une analyse supplémentaire soit effectuée sans délai.

121

497. Dans son témoignage, M. Willis s'est reporté au tableau suivant, qui figure à la p. 4 du rapport Wisner, pour justifier l'étude supplémentaire qu'il réclamait et pour expliquer ses craintes quant à l'existence possible de partialité fondée sur le sexe :

Tableau 1  
Écarts en pourcentages

Groupe	<-15 % à - 10,0	- 14,99 à - 5,00	- 9,99 à - 2,50	- 4,99 à 0,01	2,49 à 2,49	0 à 4,99	0,01 à 9,99	2,50 à 14,99	5,00 à 14,99	10,00 à 14,99	> 15
Féminin	6	8	7	5	9	43	9	10	4	9	8
Masculin	8	15	13	9	10	27	6	4	4	4	4
Total	14	23	20	14	19	70	15	14	8	13	12

498. M. Willis a expliqué que le tableau ci-dessus était une ventilation du groupe total des 222 évaluations. Sur la première ligne intitulée «Féminin», dans la colonne 0, le 43 de la cellule ombragée indique que M. Wisner et le Comité directeur étaient du même avis pour 43 évaluations. A droite de cette cellule (43) est indiqué le nombre d'évaluations du Comité directeur qui étaient supérieures à celles de

M. Wisner (elles totalisent 40), et à gauche de cette cellule (43) est indiqué le nombre d'évaluations du Comité directeur inférieures à celles de M. Wisner (elles totalisent 35). Sur la ligne «Masculin», à la colonne 0, le 27 de la cellule ombragée est le nombre d'évaluations sur lesquelles M. Wisner et le Comité directeur se sont entendus. Les colonnes de droite indiquent que le Comité directeur a attribué une cote supérieure à celle de M. Wisner dans 22 cas, et les colonnes de gauche indiquent que le Comité directeur a donné une cote inférieure à celle de M. Wisner dans 55 cas. Aux yeux de M. Willis, il y a là l'indication d'un début de tendance, car il y a environ deux fois plus d'évaluations de postes à prédominance masculine (55) qui ont reçu une cote inférieure que d'évaluations sur lesquelles le consultant et le Comité étaient d'accord (27) et d'évaluations (22) surestimées par comparaison au nombre élevé d'emplois à prédominance masculine, à l'intérieur de l'ensemble des groupes professionnels à majorité masculine, qui ont reçu des comités une cote plus basse que la cote du consultant.

499. Cet aspect de la Wisner 222 inquiétait M. Willis. Un autre aspect inquiétant du rapport était le fait que, pour un groupe professionnel à prédominance féminine (ST), les données indiquent une proportion relativement élevée de surestimations. Il s'agissait là, pour M. Willis, d'une indication, aussi faible fût-elle, de partialité fondée sur le sexe.

122

500. M. Willis a affirmé que la Wisner 222 était très limitée. Elle ne devait pas servir à déterminer de façon concluante s'il existait ou non de la partialité fondée sur le sexe et à en préciser l'ampleur. Cette étude contenait toutefois assez de preuves pour justifier un examen plus approfondi, qu'il jugeait nécessaire avant qu'il puisse défendre les résultats sans réserve.

501. A la suite de la publication de la Wisner 222, les syndicats ont envoyé une lettre à M. Durber pour lui faire part de leurs préoccupations. Dans cette lettre, qui est datée du 27 septembre 1989, la coprésidente syndicale, Mme Christine Manseau, affirme que les syndicats ne sont pas d'avis que le rapport de cette étude montre que les évaluations étaient entachées de partialité fondée sur le sexe. Au paragraphe 2 de la lettre elle dit :

[TRADUCTION]

Notre analyse montre que, en moyenne, il y a remarquablement peu de différence entre les cotes du consultant et celles des comités.

Sur les 118 postes féminins de l'échantillon, la cote moyenne du consultant est de 182 et celle des comités, de 181. Dans le cas des 104 postes masculins, la cote moyenne du consultant est de 273 et celle des comités est de 263, un écart de 3,7 %. Selon nous, ces écarts ne sont pas significatifs et nous remarquons qu'ils sont bien en-deça du niveau d'exactitude de +/-5 %, pour les cotes moyennes, que les parties avaient convenu d'utiliser lorsqu'il avait été question de la réduction de la taille de l'échantillon et de la détermination de la taille de l'échantillon global pour l'étude sur la parité salariale.

(pièce PSAC-5)

502. Selon Mme Kathryn Brookfield de l'Institut, après avoir reçu le rapport de l'étude Wisner 222, les syndicats se sont posés des questions sur la concordance entre les données et les conclusions qui en avaient été tirées. Mme Brookfield a témoigné que le syndicat, après avoir examiné la distribution des évaluations portant sur les emplois à prédominance féminine, n'avait trouvé aucune preuve de déséquilibre dans les évaluations; or le consultant, dans son rapport, était arrivé à cette conclusion. Les syndicats, de poursuivre Mme Brookfield, voulaient savoir pourquoi, dans la Wisner 222, les données et les conclusions ne concordaient pas. Jusqu'à ce que cette question fût résolue, les syndicats n'étaient pas assez confiants pour demander à M. Willis d'aller de l'avant et de reprendre l'analyse. Mme Brookfield a par ailleurs témoigné que les syndicats voulaient rencontrer les représentants du Conseil du Trésor, parcourir le rapport et discuter des différences pour voir si, ensemble, ils ne pouvaient pas y trouver une explication.

503. Il y avait énormément de controverse au sein du Comité mixte quant à savoir si M. Willis devait effectuer d'autres réévaluations. M. Willis a rencontré en privé les membres du mini-comité mixte ainsi que le plein Comité mixte afin de demander une analyse plus en profondeur. Il n'a jamais bronché de sa position selon laquelle une telle analyse

s'imposait, bien que, au cours de l'été de 1989, les positions extrêmes adoptées par certains évaluateurs aient semblé s'atténuer lorsque les membres des comités ont commencé à travailler avec des comités nouveaux, frais et, dans certains cas, réorganisés. M. Willis a témoigné ce qui suit à ce propos (volume 58, à la p. 7285, lignes 4-8) :

[TRADUCTION]

R. C'était une conviction profonde que j'avais. Je croyais simplement que vu l'importance, la taille de l'étude, j'avais besoin d'être plus sûr que je pourrais, en fait, défendre les résultats.

504. A ce moment-là, M. Walt Saveland, un employé de la Commission, a effectué un «examen technique» de l'analyse Wisner 222. M. Saveland était rattaché à la Direction des politiques et de la recherche de la Commission. M. Durber lui avait demandé d'aider à l'interprétation de la Wisner 222 et de cerner le problème de la partialité. Le rapport Saveland -- intitulé «Technical Observations and Suggestions on Willis & Associates "Special Analysis of Working Committee Results"» (pièce PSAC-6) -- a présenté une liste des emplois masculins auxquels la Commission devait accorder une attention prioritaire parce que les comités différaient de 10 % ou plus dans leurs évaluations par rapport à celles du consultant. Au bout du compte, cette liste ne contenait plus que 25 postes, même si 27 postes avaient été relevés. M. Wisner et les comités se sont mis d'accord sur deux autres postes qui, pour une raison ou pour une autre, avaient été oubliés.

505. D'après le rapport Saveland, la source de la partialité apparente fondée sur le sexe serait attribuable aux questionnaires relatifs aux postes à prédominance masculine. Le reste du rapport, à partir de la page 6, utilise un certain nombre de mesures statistiques qui, selon l'expert en statistique Sunter, sont «tout à fait absurdes» (volume 105, aux pp. 12696-97).

506. Le paragraphe 2 de ce rapport dit ce qui suit :

[TRADUCTION]

La preuve la plus importante d'une partialité apparente fondée sur le sexe se trouve parmi les emplois à prédominance masculine. Un rôle critique semble être joué par 27 emplois où les évaluations des comités sont de 5 à 15 % inférieures aux évaluations du consultant. (Des indications de partialité apparente fondée sur le sexe ont également été trouvées parmi les emplois à prédominance féminine qui étaient des emplois de bureau.)

(pièce PSAC-6)

507. M. Saveland, dans son rapport, affirme que [TRADUCTION] «c'est ce genre d'asymétrie sur la ligne des emplois masculins qui indiquent une apparente partialité fondée sur le sexe». Il a exploré les effets de l'asymétrie en étendant la norme d'accord relatif de +/-2,5 % à +/-5 %. Si

la norme élargie est appliquée à la catégorie «accord relatif» pour la ligne des emplois à prédominance féminine, on obtient une distribution parfaitement symétrique et une majorité importante de postes (76) sur lesquels il y a un accord relatif. Pour ce qui est des emplois à prédominance masculine, 56 font maintenant partie de cette dernière catégorie (les «accords relatifs»), mais les sous-évaluations l'emportent sur les surévaluations dans un rapport exact de trois pour un (36/12).

508. Le rapport présente, à la fin, des suggestions d'ordre technique. Il est proposé, par exemple, que l'on confie le réexamen des emplois litigieux à un comité existant ou nouvellement formé dont les membres sont expérimentés dans l'évaluation d'emplois. Ce comité de révision devrait se pencher sur tous les postes faisant partie d'une catégorie «suspecte», ce qui [TRADUCTION] «signifie tous les emplois à prédominance masculine actuels et additionnels (et peut-être aussi tous les emplois de bureau)». Selon le rapport, l'examen des seuls postes jouant un rôle critique par rapport à la partialité fondée sur le sexe risque de perdre de son objectivité. Le rapport propose des façons de traiter les cas où un comité de révision accepterait ou rejetterait une évaluation particulière d'un comité. Toujours selon le rapport, pendant que le comité de révision accomplit son travail, le consultant pourrait réévaluer les mêmes postes. M. Wisner serait le consultant à retenir de préférence pour la réévaluation parce qu'il offrirait la meilleure garantie de continuité. A la page 24 on peut lire ceci :

[TRADUCTION]

Si d'autres [consultants] font le travail pour Willis and Associates, il faudra mettre en place des procédures de contrôle de la qualité afin de nous assurer que les nouveaux consultants auront fait le travail précédent exactement de la même façon.

(pièce PSAC-6, à la p. 24)

509. M. Saveland était présent à la réunion du Comité mixte tenue le 31 octobre 1989. Il y a présenté son analyse de la Wisner 222. (Son rapport a été publié après cette réunion et porte la date du 10 novembre 1989.) Mme Brookfield a témoigné que M. Saveland, dans son exposé, avait souscrit à la position des syndicats, à savoir qu'il n'y avait dans le rapport aucune preuve de surestimation systématique des postes féminins. M. Saveland a en outre dit au comité que la plupart des différences entre M. Wisner et les comités d'évaluation concernaient 27 postes masculins.

510. M. Durber a lui aussi assisté à la réunion du Comité mixte le 31 octobre 1989. On lit au procès-verbal (pièce R-44), à la p. 9, que M. Durber a demandé au Comité mixte ce qu'il comptait faire au sujet de l'apparente partialité fondée sur le sexe mentionnée dans la Wisner 222. M. Durber a offert au Comité mixte l'aide de la Commission. A ce moment-là, le côté patronal du Comité était disposé à ce qu'on effectue d'autres examens des résultats de M. Willis. La position syndicale, communiquée par Mme Manseau, la coprésidente syndicale, était qu'avant cette réunion les syndicats n'étaient pas en mesure d'aller de l'avant avec l'étude Willis. Mme Manseau a promis de répondre à la partie patronale au

125

plus tard le 10 novembre 1989 quant à savoir si la partie syndicale irait de l'avant et qui la représenterait dans le processus conjoint.

511. A la suite de la réunion du Comité mixte du 31 octobre 1989, M. Durber a envoyé à Mme Manseau une lettre en date du 10 novembre 1989. Dans sa lettre, M. Durber fait remarquer que ce qui préoccupe la Commission, c'est l'apparence de partialité fondée sur le sexe; il précise que la Commission n'a encore tiré aucune autre conclusion, mais qu'elle s'attend à ce que les parties règlent la question de la partialité d'une façon qui satisfasse aux exigences de la Loi. M. Durber signale par ailleurs que M. Saveland, dans son rapport, parle d'examiner les 27 emplois masculins, et il formule le souhait qu'on effectue l'analyse séparée avec soin pour en assurer l'objectivité.

512. Dans l'espoir de comprendre le rapport Wisner 222, les syndicats ont approché leurs membres qui avaient fait partie du Comité directeur pour obtenir des renseignements qui auraient pu les aider à expliquer les écarts entre les évaluations faites par M. Wisner et celles effectuées par les comités. Mme Brookfield a témoigné avoir reçu des informations de l'ACCTA. Un membre de l'ACCTA, M. Rick Smith, s'est vu remettre 27 questionnaires de titulaires masculins et désigner par le côté syndical pour les analyser. Les informations qu'il a fournies ont été déclarées et admises sous la cote PIPSC-129. L'auteur du rapport n'a pas témoigné à l'audience. Ses conclusions sont présentées aux pages 2 et 3 du rapport dont voici la teneur :

[TRADUCTION]

En résumé, après avoir étudié attentivement les résultats des comités et ceux du consultant, je conclus que ce dernier a produit de façon constante des cotes supérieures pour plusieurs raisons. Certaines sont expliquées plus haut et d'autres sont signalées

dans ses justifications raisonnées. Les écarts en pourcentages que j'ai indiqués entre les comités et le consultant vont d'un niveau insignifiant (à mon avis), 5,4 %, à 17,4 %, qui est juste à la limite d'une tolérance d'erreur acceptable. Je ne trouve aucune preuve de partialité, encore que je ne puisse éliminer cette possibilité non plus. Les comités et le consultant ont produit des cotes complètes et solides, avec des justifications raisonnées logiques à l'appui. Il y a de légères différences dans tous les cas, mais cela est normal. Ma propre analyse des postes était souvent légèrement différente des deux, parfois plus proche de l'un ou de l'autre.

Le processus n'est pas une science exacte et dans le plan Willis il n'y a pas d'évaluation correcte ou incorrecte d'un emploi. Un consensus est le mieux qu'on puisse espérer obtenir et je n'ai aucune raison de ne pas accepter les cotes des comités telles quelles.

513. Selon Mme Brookfield, les syndicats avaient hâte de rencontrer les représentants du Conseil du Trésor avec toutes les informations que les syndicats avaient recueillies, notamment le rapport de M. Smith, supra,

126

afin de déterminer si les écarts entre les consultants et les comités pouvaient être expliqués.

514. Par une lettre datée du 27 novembre 1989 qu'elle a fait parvenir à Mme Manseau, Mme Ouimet a fait savoir que la partie patronale avait besoin d'une réponse à sa demande pour qu'on donne instruction à Willis & Associates d'effectuer d'autres analyses. Dans sa lettre, Mme Ouimet affirme que si, au plus tard le 1er décembre 1989, l'employeur n'a reçu aucune réponse, il «procédera unilatéralement» (pièce HR-17, document 22).

515. La réunion suivante du Comité mixte était prévue pour le 13 décembre 1989. Mme Brookfield a témoigné que le côté syndical n'avait pas eu l'occasion de discuter avec le côté patronal du rapport reçu de M. Rick Smith, de l'ACCTA. D'après la lettre du 27 novembre 1989 de Mme Ouimet, le côté patronal avait entrepris l'examen des 27 questionnaires relevés par M. Saveland de la Commission. Le deuxième paragraphe de la lettre dit :

[TRADUCTION]

Tel que demandé, nous sommes disposés à échanger des observations sur les 27 questionnaires relevés dans l'analyse de M. Willis le 8 décembre; les modalités d'un sous-comité seront discutées à la réunion du 13 décembre. Son travail, toutefois, est indépendant de la recherche exigée par Willis and Associates; ce travail doit commencer immédiatement et sera mené en parallèle avec celui du comité s'il est établi. Même si le comité trouve une explication pour les 27 questionnaires en question, nous avons quand même besoin de plus d'évaluations en vue de faire des estimations du biais pour les divers groupes d'emplois compris dans l'étude. Si tard dans le processus, nous ne pouvons nous payer le luxe d'attendre encore. Si, au plus tard le 1er décembre, nous n'avons pas reçu de vos nouvelles concernant le travail supplémentaire de Willis and Associates, nous procéderons unilatéralement.

(pièce HR-17, no 22)

516. La partie syndicale a interprété cette lettre comme voulant dire que, si on mettait en branle un processus mixte en vue de trouver une explication concernant les 27 questionnaires masculins, le Conseil du Trésor entendait donner son aval à la recommandation de M. Willis d'effectuer une autre analyse, avec ou sans le consentement des syndicats. C'est effectivement ce qui s'est produit lorsque la coprésidente syndicale a reçu de Mme Ouimet la lettre en date du 11 décembre 1995, qui disait notamment ce qui suit :

[TRADUCTION]

Nous demeurons fermement convaincus que l'incertitude entourant ces évaluations nécessite une étude plus approfondie. Nous acceptons la recommandation de Willis and Associates de réaliser une autre étude (ce qu'appuie, semble-t-il la CCDP). Nous avons donné notre aval à votre proposition d'examiner les 27 évaluations

127

qui, d'après la CCDP, seraient entachées de «partialité apparente», mais vous n'avez pas répondu à notre proposition pour que l'on effectue d'autres évaluations en même temps. Pour citer M. Durber, «[...] nous souhaitons que la question de la partialité fondée sur le sexe soit réglée le plus rapidement possible». Vos réponses à nos lettres nous forcent à conclure que vous ne voulez pas résoudre cette question dans l'avenir immédiat. Nous avons donc décidé, conformément aux recommandations formulées tant par le consultant que par la CCDP, d'aller de l'avant dès le

11 décembre, auquel moment le processus par lequel Willis and Associates pourra effectuer une autre analyse débutera. Nous vous tiendrons au courant de l'avancement de l'étude. Soyez assurés, par ailleurs, que la même méthodologie dont le Comité mixte avait unanimement convenu pour la première phase sera soigneusement appliquée. [c'est nous qui soulignons]

(pièce HR-17, no 7)

517. M. Willis a témoigné que l'employeur ne l'avait pas consulté avant d'annoncer à la partie syndicale qu'il avait décidé de procéder unilatéralement et de l'autoriser à faire d'autres réévaluations. A l'ouverture de la réunion du Comité mixte, le 13 décembre 1989, Mme Manseau a fait lecture d'une déclaration. A la demande de celle-ci, la déclaration a été annexée au procès-verbal, après quoi les syndicats se sont retirés et la séance a été levée. Nous reproduisons intégralement le texte de la déclaration.

[TRADUCTION]

DÉCLARATION DE CHRISTINE MANSEAU  
COPRÉSIDENTE DU COMITÉ MIXTE SYNDICAL-PATRONAL  
SUR LA MISE EN OEUVRE DE LA PARITÉ SALARIALE  
AU NOM DES SYNDICATS DE LA FONCTION PUBLIQUE

Depuis un certain temps, les syndicats représentés au Comité mixte n'ont pas l'impression d'être des partenaires égaux dans cette entreprise conjointe. Nous avons voulu discuter conjointement des conclusions de la CCDP concernant les résultats de Willis and Associates dans un contexte informel pour que le Comité mixte puisse peut-être en venir à une entente conjointe sur la façon de réagir aux recommandations. Nous avons proposé la formation d'un sous-comité pour examiner conjointement nos conclusions relatives aux évaluations du consultant dont il est fait état dans l'analyse spéciale Willis avant d'aller de l'avant avec une autre analyse -- on nous a refusé cela. Nous avons demandé que l'analyse supplémentaire ne soit pas amorcée unilatéralement, car nous estimions qu'une telle façon de faire compromettrait le caractère conjoint de l'étude et minerait la crédibilité du Comité mixte -- on nous a refusé cela.

A la lumière de la lettre de Mme Ouimet en date du 11 décembre annonçant que le Conseil du Trésor avait décidé d'aller de l'avant unilatéralement avec l'analyse supplémentaire que mènera Willis & Associates, nous jugeons que cette étude n'est plus conjointe.

Par conséquent, nous ne voulons plus participer à aucune discussion sur les questions encore en suspens pour le moment.

Nous demandons que cette déclaration soit enregistrée textuellement dans le procès-verbal et que la correspondance échangée depuis la dernière réunion du Comité mixte soit jointe au procès-verbal.

(pièce HR-11B, onglet 34)

518. Entre la réunion du Comité mixte du 25 août 1989 à laquelle M. Willis a pour la première fois recommandé une analyse supplémentaire et celle du 13 décembre 1989 où la partie syndicale s'est temporairement retirée de l'étude, il y a eu énormément de tension entre les parties. Cette tension s'était manifestée plus tôt durant les travaux des sous-comités des tests de fiabilité (inter-évaluateurs et inter-comités), mais c'est après la publication de l'étude Wisner 222 que les relations entre les parties patronale et syndicale ont commencé à se détériorer rapidement.

519. A partir du 25 août 1989, la partie syndicale voulait qu'on poursuive l'étude sur la parité salariale en vue de conclure la phase d'évaluation, de déterminer les méthodes d'indemnisation et de comparaison des salaires, et -- si une disparité salariale devait être relevée -- de poursuivre au besoin les réunions bilatérales et multilatérales. En revanche, après la réunion d'août la partie patronale était persuadée qu'il fallait à tout prix mener une autre étude et qu'il était impossible de régler la question de l'apparence de partialité fondée sur le sexe sans cette étude.

520. Les parties se sont repliées de plus en plus sur leurs positions au cours de l'automne de 1989, ce qui a fait grimper la tension. Entre le 7 novembre 1989 et le 11 décembre 1989, pas moins de 21 lettres, admises en preuve, ont été écrites par les coprésidents du Comité mixte, la même partie en signant jusqu'à trois dans une même journée. Dans son témoignage, Mme Brookfield a fait les commentaires suivants à ce sujet (volume 169, de la p. 21296, ligne 24, jusqu'à la p. 21297, ligne 9) :

[TRADUCTION]

Q. Avez-vous déjà, depuis les années que vous traitez l'un avec l'autre, eu à faire face à une telle avalanche de paperasse?

R. Non. Je pense que le document HR-17... je dirais que sur une période de six semaines, chaque question imaginable -- quatre ou cinq -- étaient débattues par correspondance, certaines simultanément, et je pense que cela montre assez bien jusqu'à quel point les gens avaient de la difficulté à communiquer les uns avec les autres, ce qui explique cette avalanche de lettres.

521. Puisque les syndicats refusaient de souscrire à une autre analyse, Mme Ouimet a informé M. Willis que l'employeur avait l'intention de demander à Willis & Associates d'accomplir le travail pour le compte du

129

Conseil du Trésor. Le 19 décembre 1989, M. Willis a écrit à Mme Ouimet pour l'aviser qu'il refusait de mener une autre analyse «unilatéralement» pour le compte du Conseil du Trésor. M. Willis a témoigné qu'il avait compris dès le départ qu'il n'avait de comptes à rendre qu'au Comité mixte. Selon lui, cette démarche était inappropriée. M. Willis souhaitait que l'on reconvoque le Comité mixte. Un représentant du Conseil du Trésor, M. Gaston Poirier, lui a demandé dans quelles circonstances il accepterait de mener l'analyse. M. Willis lui a répondu qu'il ferait une étude sur un plus gros échantillon si la Commission le lui demandait, car «la Commission des droits de la personne était un tiers objectif et cela faisait partie de son mandat» (volume 59, à la p. 7311).

522. Dans la lettre du 18 décembre 1989 qu'il a envoyée à Mme Ouimet, M. Willis mentionne pour la première fois ce que les informations découlant d'une seconde étude apporteraient. La partie pertinente de sa lettre dit ce qui suit :

[TRADUCTION]

Je crois qu'il est nécessaire d'élargir cette étude pour déterminer l'ampleur de toute partialité qui pourrait exister dans les évaluations. Ces données devraient servir de base au rajustement des résultats d'évaluation qui pourra se révéler nécessaire pour assurer une étude équitable et objective. [c'est nous qui soulignons]

(pièce HR-92)

523. Le 23 janvier 1990, l'Alliance a annoncé qu'elle se retirait de façon permanente du Comité mixte et trois jours plus tard, le 26 janvier 1990, le président du Conseil du Trésor a annoncé la mise en oeuvre de rajustements paritaires et fourni l'assurance que la décision du

gouvernement n'aurait aucun effet préjudiciable sur les conclusions que la Commission pourrait tirer à l'égard des questions encore à l'étude.

524. Mme Brookfield a témoigné qu'elle avait remarqué un changement d'attitude de la part de l'employeur vers la fin de l'étude. Elle a signalé que, au début, les discussions entre les syndicats et l'employeur portaient sur l'apparence d'une partialité fondée sur le sexe. Après l'étude Wisner 222, cependant, le Conseil du Trésor a cessé de parler d'apparence de partialité fondée sur le sexe pour dire qu'il allait apporter des rajustements éventuels selon la partialité effective constatée.

525. Les syndicats étaient très préoccupés par ce revirement du Conseil du Trésor après la Wisner 222. Mme Brookfield a témoigné qu'il y avait eu échange de lettres à propos du rajustement des cotes. Elle a notamment mentionné une lettre (pièce HR-41) en date du 26 janvier 1990 -- datant d'après l'arrêt de l'étude -- que le président du Conseil du Trésor a adressée à M. Max Yalden, président de la Commission, pour lui expliquer que les rajustements paritaires avaient été calculés en fonction des rajustements que le Conseil du Trésor avait faits pour tenir compte de la partialité fondée sur le sexe.

130

526. Dans la lettre susmentionnée, le président du Conseil du Trésor, M. Robert de Cotret, a fait part à M. Yalden de la décision du gouvernement de mettre en oeuvre, à la grandeur de la fonction publique, des mesures fondées sur les résultats des évaluations effectuées sous la gouverne du Comité mixte. M. de Cotret ne parle pas de l'apparente partialité fondée sur le sexe relevée dans la Wisner 222, mais plutôt de «l'ampleur de la partialité fondée sur le sexe». Voici un extrait de la lettre de M. de Cotret :

[TRADUCTION]

Je suis convaincu qu'une étude de cette envergure -- qui est sans précédent -- doit être juste, statistiquement solide et crédible, compte tenu de ses ramifications importantes. Cette autre analyse était nécessaire pour déterminer l'ampleur de la partialité fondée sur le sexe et pour rajuster en conséquence les résultats des évaluations menées par le Comité mixte. Je suis donc reconnaissant envers la Commission d'avoir accepté de mener cette analyse afin de déterminer l'ampleur de la partialité fondée sur le sexe. [c'est nous qui soulignons]

(pièce HR-41)

527. L'extrait ci-dessus semble confirmer le changement d'orientation que le syndicat a dit avoir perçu chez l'employeur qui, au départ préoccupé par l'apparence de partialité fondée sur le sexe soulevée dans la Wisner 222, cherchait maintenant à rajuster les résultats en fonction de la partialité effective. Il a semblé à Mme Brookfield que le Conseil du Trésor avait décidé que la Wisner 222 avait définitivement établi la présence de partialité fondée sur le sexe et que tout ce qu'il restait à faire était de rajuster les cotes en conséquence.

528. Au début de 1990, la Commission a communiqué avec M. Willis. Ce contact a eu lieu après que l'Alliance eut annoncé son retrait de l'étude sur la parité salariale. M. Durber a informé M. Willis que la Commission avait décidé qu'une autre analyse s'imposait à partir de réévaluations qu'effectuerait Willis & Associates. Toutefois, la Commission analyserait elle-même les résultats des réévaluations Willis.

529. Selon M. Willis, à défaut d'une autre analyse il n'y avait qu'une autre solution qui s'offrait : utiliser un autre système d'évaluation, ce qui, dans les faits, aurait eu pour effet de reprendre une large partie de l'étude. Cette autre solution aurait été extrêmement coûteuse. M. Willis a par ailleurs exprimé son avis sur ce qu'on devait faire avec les résultats de l'étude. Le Tribunal, a-t-il dit, a trois possibilités : (i) mettre l'étude en oeuvre telle quelle; (ii) rajuster les résultats; (iii) rejeter l'étude. Pour sa part, il ne rejeterait pas l'étude, mais ajusterait les résultats en fonction de toute partialité fondée sur le sexe.

## E. LA COMMISSION

131

530. Lorsque, en avril 1985, la Commission a répondu à l'invitation du président du Conseil du Trésor d'appuyer le Comité mixte syndical-patronal sur la mise en oeuvre de la parité salariale, elle a accepté de suspendre l'enquête relative aux plaintes fondées sur l'article 11 qui avaient été déposées avant l'annonce du Comité mixte, et de suspendre également l'examen des plaintes déposées après cette annonce. La Commission a déclaré qu'elle attendrait les résultats de l'étude avant de prendre quelque mesure que ce soit. Cela dépendait aussi des circonstances qui existaient lors du dépôt des plaintes.

531. La réponse de la Commission à l'invitation était contenue dans une lettre datée du 17 avril 1985 (pièce HR-18, onglet 18) que M. Gordon

Fairweather a envoyée à M. de Cotret. M. Fairweather y affirmait que, si la Commission était convaincue que la méthodologie employée pour la réalisation de l'étude était conforme à l'article 11 de la Loi, elle publierait une ligne directrice spéciale portant que l'étude avait été menée en conformité avec la Loi. Elle publierait en outre des lignes directrices pour la mise en oeuvre de mesures correctives conformément à l'article 11.

532. La Commission a participé aux travaux du Comité mixte à titre de simple observateur. Des représentants de la Commission ont assisté aux réunions du Comité et, à la demande de ses membres, ont fourni éclaircissements et conseils à propos de l'étude sur la parité salariale. La participation de la Commission a surtout été d'ordre technique; celle-ci a notamment pris part à la sélection des échantillons pour les tests de fiabilité inter-comités et éclairé les membres sur l'interprétation de la Loi et de l'Ordonnance. Des employés de la Commission ont aussi siégé à titre d'observateurs durant les travaux des comités d'évaluation (les cinq du début et les neuf subséquents).

533. La Commission n'avait pas l'intention d'être partie aux ententes que concluraient éventuellement les parties réunies au sein du Comité mixte. Par contre, elle comptait examiner toute entente éventuelle afin de déterminer si elle satisfaisait aux exigences de l'article 11 de la Loi.

534. Au début de mai 1989, M. Durber est entré au service de la Commission à titre de chef, section de la parité salariale, titre qui a plus tard été changé à directeur de la Direction de la parité salariale. Le 12 juin 1989, M. Durber a rencontré les coprésidents du Comité mixte et leur a dit craindre que si les parties n'arrivaient pas à décider quoi faire à propos de la Wisner 222, l'étude sur la parité salariale pourrait facilement s'effondrer. M. Durber a témoigné que les coprésidents ont convenu à cette réunion que toutes les parties, y compris la Commission, devraient avoir accès aux résultats de l'évaluation des emplois découlant de l'étude sur la parité salariale.

535. M. Durber a informé les coprésidents à ce moment-là que pour la Commission la question qui se posait était comment interpréter les évaluations des postes qui avaient été faites. Il a souligné que s'il y avait de la partialité fondée sur le sexe la Commission voudrait avoir son mot à dire, car elle avait besoin de savoir si les évaluations étaient

admissibles en preuve, dans l'éventualité où elle poursuivrait l'examen des plaintes déposées par l'Alliance.

536. La Commission n'a procédé à aucune enquête formelle concernant les plaintes avant le 6 mars 1990. A cette date, à la demande de la Commission, les participants au Comité mixte se sont réunis avec celle-ci en vue d'examiner les questions en suspens. A ce moment-là, le Comité mixte avait été dissout de façon permanente.

537. Le 6 mars 1990 est la date importante suivante. A ce moment-là, la Commission a rencontré les participants au Comité mixte. Elle voulait réduire le nombre de questions en litige découlant de l'étude sur la parité salariale au cas où les plaintes seraient renvoyées à un Tribunal. Dans le communiqué qu'elle a émis après la réunion, la Commission a annoncé qu'elle devait être convaincue que toutes les exigences de la Loi avaient été remplies. Elle a en outre précisé que le Conseil du Trésor lui avait remis les calculs qu'il avait utilisés pour prévoir ses rajustements, calculs qu'elle examinerait au cours de son enquête.

(i) L'enquête de la Commission

538. Lorsque l'étude sur la parité salariale a pris fin au début de 1990, il est devenu évident à la Commission que son rôle d'observateur au sein du Comité mixte avait également pris fin et qu'il était temps de reprendre le processus normal d'examen des plaintes là où il avait été interrompu. La question de l'apparence de partialité fondée sur le sexe soulevée par l'étude Wisner 222 faisait partie de l'enquête relative aux plaintes. La Commission a décidé qu'elle s'attaquerait en tout premier lieu à cette question dans son enquête visant à déterminer si la discrimination salariale persistait dans la fonction publique fédérale. Le gouvernement avait effectué des rajustements paritaires en janvier 1990, et l'Alliance maintenait que ces paiements n'avaient pas fermé l'écart salarial, ce qui fait qu'il y avait toujours de la discrimination salariale.

539. La partialité fondée sur le sexe était une considération lorsque le président du Conseil du Trésor a annoncé les rajustements paritaires en janvier 1990. Le président du Conseil du Trésor n'avait pas précisé dans quelle mesure les rajustements tenaient compte de la partialité, mais il avait affirmé dans son annonce que la Commission examinerait la question.

540. La Commission, tel qu'elle l'a expliqué dans la pièce HR-55 («Notes for Presentation on Alleged Gender Bias in Job Evaluation of the Joint Initiative»), a envisagé de façon conservatrice, dans son enquête, l'ampleur de la preuve qu'elle rechercherait dans le but de trancher la question de l'apparence de partialité fondée sur le sexe.

541. M. Durber a témoigné que la Commission avait examiné les cinq plaintes de l'Alliance et de l'Institut simultanément. C'était

probablement l'enquête la plus rapide jamais menée par la Commission du fait qu'elle avait déjà en main les données sur l'évaluation des emplois obtenues de l'étude sur la parité salariale. La Commission n'avait donc pas à faire ses propres évaluations.

133

542. La Commission devait faire porter son enquête sur quatre grandes questions. La première était la partialité fondée sur le sexe. La Commission devait décider si elle pouvait se fier aux informations sur l'évaluation des emplois découlant de l'étude sur la parité salariale. Dans un deuxième temps, elle devait examiner tout écart salarial qu'on aurait décelé. Il lui fallait notamment élaborer une méthode pour calculer les écarts salariaux. Troisièmement, la Commission devait prendre en considération et évaluer les avantages. Enfin, elle devait examiner les parties de deux plaintes particulières qui soulevaient les limitations, sur le plan de l'avancement professionnel, qui découlaient des pratiques de rémunération. (L'examen de ce quatrième volet n'est pas encore terminé.)

543. Un aperçu de la chronologie commence avec l'enquête amorcée en mars 1990 par la Commission, qui a produit ses conclusions provisoires concernant la partialité fondée sur le sexe en juillet de la même année. Au cours du même mois, la Commission a fait part aux parties de ses constatations relatives à «l'apparence de partialité fondée sur le sexe» dans les évaluations des comités. En août 1990, elle a produit un rapport préliminaire sur l'écart salarial et informé les parties de ses conclusions provisoires.

544. Il y a également eu en août une réunion avec parties pour faire le point sur l'enquête de la Commission concernant l'évaluation des avantages. En septembre 1990, le Conseil du Trésor a répondu par écrit au rapport préliminaire que la Commission avait présenté en août. Le rapport final de l'enquête a été soumis aux commissaires à la fin de septembre 1990. En octobre, la Commission a rendu sa décision sur l'écart salarial qui était l'objet des cinq plaintes et a demandé au président du Comité du Tribunal des droits de la personne de constituer un tribunal.

545. La Commission a présenté les résultats de son enquête relative aux plaintes fondées sur l'article 11 dans un rapport (pièce HR-250) intitulé «Investigator's Report: Wage Adjustment in the Federal Public Service - Possible Gender Bias in Job Evaluation Data». M. Durber a distribué ce rapport aux parties en septembre 1990. Le rapport contient les constatations et conclusions de la Commission concernant la question de l'apparence de partialité fondée sur le sexe dans les évaluations des

comités. Les conclusions de la Commission figurent au paragraphe 51 du rapport; en voici la teneur :

[TRADUCTION]

#### 51. Conclusions

Le personnel de la Commission a constaté que les vérifications de M. Willis révélèrent certains écarts entre les évaluations des consultants et celles effectuées par le Comité mixte. Les enquêteurs n'ont pas trouvé que ces écarts révélèrent des tendances caractérisées par une corrélation constante avec le sexe ou la profession dans les évaluations du Comité mixte. L'ampleur de la possible «sous-estimation» des emplois masculins est inférieure à 3 %, mais elle peut probablement s'expliquer par le fait que chacun comprend un peu différemment le travail décrit,

134

ainsi que par la signification des postes-repères et l'application du plan Willis. L'écart ne serait apparemment pas lié à un parti pris relié au sexe. De plus, ce 3 % n'est pas distribué également parmi les professions. De toute évidence, le fait que deux groupes indépendants (les consultants de M. Willis et le comité d'analyse de la qualité) aient pu produire des résultats variant par une marge de 2 % à 3 % indique que de tels écarts sont à prévoir et peuvent s'expliquer par des facteurs autres que la partialité.

(pièce HR-250, partie I)

546. Le témoignage oral de M. Durber corrobore et confirme le contenu de ce rapport et souligne les étapes que la Commission a suivies dans son enquête sur la possibilité que les évaluations des comités soient empreintes de partialité fondée sur le sexe. Il ressort notamment du témoignage de M. Durber que la Commission a fait d'autres tests après le début de l'audience. Le Tribunal examinera tant les méthodes d'enquête que la Commission a appliquées durant son enquête initiale que les tests qui ont subséquentement été effectués à la demande de celle-ci.

547. Dans son rapport, l'enquêteur affirme qu'il n'a trouvé aucune preuve claire d'un parti pris fondé sur le sexe dans les résultats d'évaluation. Il recommande des formules pour calculer les rajustements salariaux à verser pour établir la parité salariale entre les hommes et les femmes, lesquels rajustements ne devraient pas être modifiés pour tenir

compte d'une partialité possible fondée sur le sexe. Enfin, il propose que la Commission accepte ses constatations concernant les plaintes connexes déposées en vertu de l'article 11.

548. A l'été de 1990, une ébauche du rapport d'enquête (pièce HR-250) a été distribuée aux parties pour qu'elles la commentent. Le Conseil du Trésor a répondu par l'intermédiaire de Mme Ouimet, en sa qualité de secrétaire adjointe, Division de la classification, des systèmes d'information sur les ressources humaines et de la paye, qui a adressé lettre et rapport à M. Durber. Mme Ouimet a témoigné durant le voir-dire tenu par le Tribunal, mais elle n'a pas été citée à la reconvoction de l'audience. Dans le dernier paragraphe de sa lettre elle conclut que l'enquête de la Commission était déficiente et ne démontrait pas clairement l'absence de partialité fondée sur le sexe. Par contre, elle exprime le point de vue qu'il est peu probable qu'une des parties parvienne à démontrer l'existence de partialité fondée sur le sexe dans les résultats. Nous reproduisons ci-dessous le paragraphe en question :

135

[TRADUCTION]

En revanche, il est peu probable que quiconque arrive à démontrer l'existence effective de partialité fondée sur le sexe, étant donné que la firme de M. Willis n'a fourni aucune référence à laquelle on pourrait comparer les résultats des évaluations d'une étude à l'autre. Il n'est pas possible de mesurer adéquatement l'application du plan de façon à conclure définitivement qu'il y a ou qu'il n'y a pas de partialité. N'allez pas conclure, cependant, que nous ne devrions pas examiner très attentivement toutes les incohérences dans les cotes qui ont été soulevées par les divers comités travaillant sous l'égide du Comité mixte, par votre propre recherche et par la nôtre. Il est maintenant vital que nous mettions de côté le «pourquoi» des anomalies de cotation et que nous tournions plutôt notre attention vers les moyens de les corriger. Nous sommes disposés à prendre part à la conception d'une étude qui permettrait de résoudre les incohérences dans les évaluations. [c'est nous qui soulignons]

(pièce HR-46, à la p. 2)

549. Dans les observations détaillées jointes à la réponse du Conseil du Trésor, Mme Ouimet pose la question rhétorique suivante : [TRADUCTION] «Est-il possible de distinguer entre la partialité fondée sur le sexe dans les évaluations et l'application globale du plan Willis de façon à

attribuer une pondération à chacune?» Selon le rapport, la réponse doit indiquer la mesure dans laquelle la question de la partialité fondée sur le sexe est une question purement statistique ou une question de fond. Dans la deuxième éventualité, de soutenir Mme Ouimet, les statistiques ne peuvent jeter que peu de lumière sur la question.

550. Dans la réponse écrite qu'il a faite au rapport final de la Commission intitulé Possible Gender Bias in the Evaluation Data, laquelle réponse est contenue dans une lettre de Mme Ouimet en date du 7 septembre 1990 adressée à M. Durber, le Conseil du Trésor exprime clairement l'opinion qu'une étude statistique n'est pas la meilleure façon de déterminer s'il existe une partialité fondée sur le sexe. Les extraits suivants de ses commentaires, à la page 1 du rapport, aident à comprendre la réponse de l'employeur :

[TRADUCTION]

Essentiellement, on peut résumer notre désaccord comme suit : l'enquêteur a entrepris un examen très restreint de la partialité fondée sur le sexe par des recherches statistiques mal exécutées. Même si cette approche eût été appropriée, la nature restreinte de l'étude globale est telle qu'on ne peut rien dire au sujet de la partialité fondée sur le sexe, puisque les questions importantes qui étaient en cause n'ont jamais été examinées.

La Commission cite abondamment la position de l'Alliance de la Fonction publique du Canada (AFPC) selon laquelle un grand nombre des questions ne sont pas de nature statistique. Nous sommes du

même avis et avons soutenu que l'analyse statistique en ce domaine n'est utile que dans la mesure où elle peut soulever la possibilité d'un problème dont la solution nécessite une approche non statistique. Nonobstant cette objection, nous estimons qu'une étude statistique, aussi solide soit-elle, n'est pas la meilleure approche en l'occurrence. Il y a tellement de jugement en cause dans la cotation d'un questionnaire relatif à un emploi qu'il est difficile, dans le meilleur des mondes, de déterminer la partialité fondée sur le sexe de façon statistique, car cela nécessite qu'on attribue à la cote elle-même une pondération pour chaque facteur de jugement, de partialité, d'incohérence, etc. Mais puisque vous avez décidé de restreindre votre étude à une analyse statistique des données de M. Willis sur les évaluations,

nous nous estimons obligés, néanmoins, de critiquer votre étude sur le plan statistique.

La longue critique que nous vous avons envoyée visait à démontrer, par des arguments statistiques, que l'approche adoptée et les résultats empiriques ne vous permettent sous aucune considération de conclure avec certitude qu'il n'y a pas de partialité fondée sur le sexe dans les évaluations issues de l'étude sur la parité salariale. Le plus que vous pouvez conclure, c'est qu'il n'y a pas assez de preuves pour conclure dans un sens ou dans l'autre. Vous n'avez examiné aucun de nos arguments de façon systématique, si ce n'est que pour dire -- sous forme de commentaire -- que nos «critiques statistiques font des distinctions trop subtiles».

(pièce HR-250, onglet J, aux pp. 1-2)

551. Le Conseil du Trésor avait apparemment employé une autre méthode pour étudier la possibilité d'une partialité fondée sur le sexe; celle-ci est décrite dans les commentaires détaillés de Mme Ouimet en date du 7 septembre 1990. A l'aide d'une approche différente, écrit-elle, le Conseil du Trésor est arrivé à la même conclusion que M. Sunter, mais selon le Conseil la conclusion est déroutante puisqu'elle ne représente que la moitié de l'histoire et ne dit rien au sujet de la fréquence de la sous-estimation ou de la surestimation des questionnaires. La conclusion générale du Conseil du Trésor se trouve à la page 10 et dit ce qui suit :

[TRADUCTION]

En utilisant les critères fournis par le cabinet de M. Willis, il n'est pas possible de conclure que, bien qu'il puisse y avoir des différences statistiques significatives dans les tendances des évaluations, celles-ci ne sont pas fondamentalement importantes. Tel que nous l'avons montré plus haut, la question du niveau de l'écart ne tient pas compte de la dimension fréquence et les différences dans les tendances sont effectivement significatives. Nous avons tenté de tenir compte des mauvaises évaluations pour voir s'il y avait une tendance quelconque liée au sexe qui s'y rattacherait, et il semble que ce soit le cas.

Nous avons analysé les mêmes données et utilisé la même mesure que M. Sunter, et pourtant nous arrivons à des conclusions différentes. Nous sommes convaincus que les données révèlent de sérieux problèmes concernant les évaluations et que ces problèmes

ressemblent beaucoup à de la partialité fondée sur le sexe; quoi qu'il en soit, une analyse plus poussée s'impose. Nous demeurons persuadés qu'il y a lieu de rajuster les cotes, mais nous sommes prêts à discuter d'une stratégie de rajustement différente de celle utilisée initialement. Tout rajustement sera difficile à calculer, compte tenu des différences significatives entre les deux études Willis. [c'est nous qui soulignons]

(pièce HR-250, onglet J)

552. Nous décrivons et examinerons maintenant certaines informations concrètes qu'on trouve dans l'enquête de la Commission fournie par M. Durber. Le 8 mars 1990, la Commission a reçu du Conseil du Trésor un document (pièce HR-185) expliquant la méthode utilisée par l'employeur pour faire ses rajustements paritaires. Selon M. Durber, le Conseil du Trésor, dans le document qu'il a publié en mars 1990, a estimé que la partialité moyenne s'établissait à +3 % pour les évaluations des postes faisant partie des groupes professionnels à prédominance féminine et à -4 % dans le cas des évaluations des postes faisant partie des groupes à prédominance masculine. Les rajustements paritaires avaient donc fait l'objet d'une correction générale lors du calcul de l'égalité de rémunération pour fonctions équivalentes. Ainsi, les fonctionnaires faisant partie des groupes professionnels à prédominance féminine ont touché des rajustements inférieurs à ceux qu'ils ou elles auraient reçus en l'absence de ces corrections.

553. La révision des cotes est expliquée ainsi dans le document sur la méthode suivie :

[TRADUCTION]

Le Conseil du Trésor a calculé le facteur de révision des cotes à l'aide de techniques statistiques simples. Tous les questionnaires, sauf ceux qui avaient été cotés par le Comité directeur et le consultant de M. Willis, ont été révisés : les cotes des questionnaires féminins ont été globalement réduites d'environ 3 % et celles des questionnaires masculins ont été haussées d'environ 4 % dans l'ensemble. Toutes les analyses de politiques présentées dans le reste de ce rapport utilisent les cotes d'évaluation révisées telles que décrites.

(pièce HR-185, aux pp. 6-7)

554. Pour bien comprendre la pièce HR-185, qui contient beaucoup de jargon et de données statistiques détaillées, M. Durber a fait parvenir le rapport à sept experts indépendants. Il s'agit des spécialistes en équité

salariale Weiner, Morley Gunderson, Lois Haignere, Willis & Associates, Roberta Rob et Judith Davidson-Palmer ainsi que du statisticien Sunter. M. Durber considérait ces spécialistes comme d'éventuels participants à un atelier que la Commission avait organisé pour avril 1990 dans le but

138

d'examiner la méthodologie du Conseil du Trésor (pièce HR-185) et de le conseiller sur la façon dont il devrait la traiter.

555. La Commission a eu de la difficulté à obtenir des données du Conseil du Trésor durant son enquête sur les plaintes. M. Durber a témoigné que les données réelles que le Conseil avait utilisées pour arriver à ses conclusions dans le document HR-185 n'ont jamais été produites. La Commission a dû effectuer des projections salariales et créer ses propres bases de données salariales à cause du temps que prenait le Conseil du Trésor à fournir des informations salariales. Un ensemble complet de données salariales a finalement été fourni à la Commission au cours des présentes audiences.

556. Le 9 avril 1990, la Commission a tenu son atelier et certains des individus mentionnés plus haut y ont assisté, à savoir M. Sunter, Mme Roberta Rob, Mme Judith Davidson-Palmer et un représentant de Willis & Associates. Les autres, qui n'étaient pas présents à la réunion, ont présenté des observations écrites. M. Durber voulait être «le mieux renseigné possible par certains des meilleurs esprits du Canada en matière d'équité salariale» (volume 147, à la p. 18197). Après des consultations assez approfondies auprès de ces experts, M. Durber a regroupé leurs avis et formulé un plan d'enquête et une hypothèse.

557. A la suite de la réunion du 9 avril 1990, M. Durber a regroupé les conseils qu'il avait recueillis dans ses discussions avec ces individus afin de clarifier les questions sur lesquelles devait se pencher la Commission. La décision a été prise de contester la méthodologie du Conseil du Trésor par des questions détaillées.

558. La Commission voulait aussi savoir si, eu égard aux facteurs du plan Willis, les résultats des groupes professionnels à prédominance masculine différaient de ceux des groupes professionnels à majorité féminine. M. Durber a engagé la société Wyatt, une firme internationale de consultants de gestion qui jouit d'une grande réputation dans le domaine de l'évaluation des emplois. Il a demandé à la société Wyatt d'utiliser la base de données de l'étude sur la parité salariale pour toutes les évaluations d'emplois effectuées. La société Wyatt a examiné les données afin de déterminer si le rapport entre les facteurs était le même quel que

soit le sexe du groupe et quelle que soit la profession dont était issu le questionnaire. Elle a présenté son rapport à la Commission au début de juin 1990. L'analyse Wyatt démontre qu'il y avait des corrélations entre divers facteurs, par exemple, la mesure dans laquelle une cote sur les exigences intellectuelles était en corrélation avec les connaissances. Les chercheurs ont conclu qu'il ne semblait pas y avoir de différences significatives dans les corrélations entre les facteurs pour les emplois masculins et féminins ou entre les tendances générales. Ils ont par ailleurs indiqué qu'il y avait un certain écart entre les emplois masculins et féminins en ce qui concerne les cotes sur les conditions de travail. M. Durber croyait que cela s'expliquait par la nature du travail (volume 147, à la p. 18208).

559. La Commission, dans son approche, ne visait pas à «prouver l'absence de partialité [fondée sur le sexe]», mais simplement à déterminer

139

si une personne raisonnable conclurait à l'existence de partialité (volume 149, à la p. 18521). Selon M. Durber, le fait qu'il se dégage une tendance différente pour les emplois masculins par opposition aux emplois féminins, comme l'a montré l'étude Wisner 222, ne dit rien au chercheur si ce n'est, peut-être, qu'il «doit chercher davantage».

560. Durant l'enquête initiale, un représentant du Conseil du Trésor a livré à la Commission une lettre datée du 20 juin 1990 accompagnée d'un cartable contenant des informations utiles à la Commission pour l'évaluation de la partialité fondée sur le sexe. Le cartable contenait les documents suivants : des documents du sous-comité FIE, les études FIC, les recommandations de changements formulées par la consultante de M. Willis, Mme Drury, relativement aux évaluations du Comité directeur, le rapport Tristat, le rapport de M. Willis en date de juillet 1988 sur les travaux du Comité directeur, les questions que le côté patronal a renvoyées à M. Willis en août 1988 concernant les évaluations du Comité directeur, les comptes rendus des réunions du Comité mixte, des copies de lettres écrites en juillet 1988 par l'Alliance et l'Institut à Mme Drury à propos des justifications raisonnées et de l'interprétation des facteurs suivant le plan Willis, la réponse de M. Willis à la contestation des évaluations du Comité directeur par les comités, une copie de la lettre que M. Willis a écrite le 17 août 1989 au sujet du comité no 4 ainsi que des copies de lettres entre les parties portant sur la Wisner 222.

561. M. Durber a affirmé que la documentation contenue dans le cartable fourni par l'employeur n'avait pas particulièrement trait à la partialité fondée sur le sexe. Dans les documents concernant le

sous-comité FIC, M. Durber a cherché des preuves spécifiques de partialité fondée sur le sexe. En ce qui concerne le rapport Tristat, il a témoigné qu'il recherchait des conclusions claires parce qu'il voulait savoir si, en fait, il y avait eu des indications ou des preuves solides de partialité fondée sur le sexe. Quant aux études FIC, vu le nombre restreint de tests, 25, il n'était pas possible, a-t-il dit, de détecter de tendance.

562. M. Durber s'est entretenu avec un des observateurs de la Commission, M. Brian Hargadon, au sujet de ses observations des tests FIC. M. Hargadon avait participé à tous les tests. Il avait aussi administré certains des tests. M. Durber a témoigné que, selon M. Hargadon, les comités d'évaluation ont fini, avec le temps, par ne plus prendre les tests aussi sérieusement qu'au départ. M. Durber a donc jugé que les tests FIC n'étaient pas concluants sur la question de la partialité fondée sur le sexe.

563. En ce qui concerne les modifications aux évaluations du Comité directeur formulées par la consultante Drury, M. Durber s'est principalement fié à l'opinion de M. Willis comme quoi les points que celle-ci avait soulevés avaient été résolus. Pour ce qui est du rapport préparé par Willis & Associates en juillet 1988 ainsi que de son analyse et de ses conclusions relativement au travail du Comité directeur, M. Durber considérait que, essentiellement, le rapport disait qu'il n'y avait pas lieu de s'inquiéter au sujet de la partialité fondée sur le sexe. M. Durber a par ailleurs affirmé qu'après avoir examiné la documentation que lui avait remise le Conseil du Trésor il ne comprenait pas plus comment

la partialité fondée sur le sexe pourrait se manifester dans les résultats de l'évaluation des emplois. La documentation du Conseil du Trésor ne lui a été d'aucune utilité, a-t-il ajouté, et il avait besoin de comprendre ce qui pouvait être en jeu par rapport à la soi-disant partialité fondée sur le sexe. Par conséquent, il a décidé de rechercher des réponses ailleurs.

564. M. Durber a signalé que les seules discussions qu'il avait eues avec le Conseil du Trésor à propos des documents contenus dans le cartable avaient eu lieu lors d'un exposé qu'il avait fait à l'employeur le 5 juillet 1990 au sujet des questions entourant la partialité fondée sur le sexe. La Commission a dû attendre au mois d'août 1990 avant de recevoir une analyse plus détaillée de la position du Conseil du Trésor sur cette question. C'est alors que le Conseil a remis à la Commission un document exposant plus en détail son point de vue sur la question.

565. Une partie de l'enquête de la Commission sur l'apparence de parti pris fondé sur le sexe a consisté à donner suite à la recommandation, contenue dans le rapport Saveland, d'analyser plus en profondeur les 27 emplois masculins «sous-estimés» (leur nombre a subséquemment été ramené à 25). Pour ces emplois, M. Saveland avait trouvé un écart de 10 % ou plus entre l'évaluation de M. Wisner et celle des comités d'évaluation. Au printemps de 1990, M. Durber a convoqué, sous la présidence de M. Ron Renaud, conseiller principal à la Section de la parité salariale de la Commission, un comité mixte composé de représentants de l'employeur et d'employés des syndicats. Ils se sont rencontrés pendant deux semaines à compter du 30 avril 1990. La Commission a envoyé aux membres du comité une lettre les informant de ce qui suit :

[TRADUCTION]

Le mandat du comité est d'effectuer un contrôle de la qualité de 27 postes qui ont été évalués, dans le cadre de l'étude sur la parité salariale, par des comités qui sont entrés en jeu après le Comité directeur. Dans une analyse de 222 évaluations de postes que Willis & Associates a exécutée en juin 1989, ce dernier a conclu que les évaluations s'écartaient de façon significative de la discipline du Comité directeur et avaient contribué le plus à la constatation d'une apparente partialité fondée sur le sexe.

(pièce PIPSC-135)

566. D'anciens évaluateurs du Comité directeur ont été choisis pour participer à ce comité, notamment deux représentants de l'employeur et trois représentants syndicaux dont les noms avaient été proposés par chacune des parties. M. Durber voulait des participants reconnus pour la largeur de leurs vues. A son avis, cet objectif a été atteint. Ce comité, qu'on a appelé le comité de l'analyse de la qualité («comité de l'AQ»), a produit un rapport intitulé The Quality Analysis Report.

567. M. Durber a témoigné que, dans le contexte du comité de l'AQ, il s'intéressait moins au fait qu'il y avait des écarts entre le consultant et les comités d'évaluation multiples qu'aux facteurs qui expliquaient ces écarts. Il voulait savoir si, aux yeux des membres du comité de l'AQ, les

différences en question avaient trait au fait qu'il s'agissait d'emplois masculins, ou si les membres percevaient de la partialité de la part des comités d'évaluation multiples. Il considérait que les cinq anciens membres du Comité directeur avaient une compréhension spéciale du plan

Willis et de la discipline du Comité directeur. Il s'attendait à ce qu'ils comprennent «les mécanismes sous-jacents à leurs propres différences par rapport aux comités».

568. M. Durber a témoigné que la Commission cherchait à savoir s'il y avait une raison, un motif, ou un effort conscient ou inconscient de la part des comités d'évaluation multiples tendant à défavoriser les emplois masculins en question. S'il y avait un endroit où la partialité fondée sur le sexe serait évidente, s'est-il dit, ce serait dans les évaluations de ces 25 emplois.

569. Pour s'acquitter de leur tâche, les membres du comité de l'AQ procédaient comme suit : chacun lisait le questionnaire, l'évaluait de façon indépendante, examinait les postes-repères du Comité directeur utilisés par le Comité mixte et ceux utilisés par le consultant, puis sélectionnait d'autres postes-repères appropriés du Comité directeur.

570. La preuve qui a été présentée au Tribunal est contradictoire quant à savoir si le comité de l'AQ était tenu d'en arriver à un consensus dans ses évaluations. Selon M. Durber, on n'avait pas demandé au comité de l'AQ de former un consensus. Il a témoigné que la Commission avait demandé à chaque membre du comité de déclarer ses évaluations au président et d'en discuter, mais non d'en arriver à un consensus. M. Durber a par ailleurs expliqué que la Commission ne cherchait pas à valider les cotes des 25 emplois; elle voulait simplement savoir si les membres du comité de l'AQ décèleraient, durant le processus, des facteurs quelconques liés au sexe dans leurs propres évaluations ou dans celles des comités d'évaluation multiples.

571. En revanche, deux membres syndicaux du comité de l'AQ ont témoigné qu'on avait demandé au comité d'en arriver à un consensus et qu'il n'avait pas réussi à le faire. D'après leur témoignage, le comité de l'AQ a suivi la même procédure Willis que les comités d'évaluation. La seule différence, selon ces témoins, c'est que le consensus devait être unanime pour chaque sous-facteur dans le plan Willis, par opposition à la majorité aux deux tiers qui était exigée pour le consensus dans les comités d'évaluation. Une pièce qui était jointe à la lettre datée du 23 avril 1990 adressée par le président du comité de l'AQ aux membres de ce comité corrobore et confirme le fait qu'on exigeait l'accord unanime pour un consensus. La partie pertinente de ce document dit ce qui suit :

[TRADUCTION]

Les résultats des évaluations doivent être déterminés par consensus. Autrement dit, chaque membre du comité doit approuver

l'évaluation pour chaque facteur et sous-facteur ainsi que les points attribués.

(pièce PIPSC-135, à la p. 3)

142

572. M. Durber a témoigné qu'à la fin des travaux du comité de l'AQ, son président, M. Ron Renaud, lui a signalé que les écarts dans les cotations entre le comité de l'AQ et les comités d'évaluation étaient attribuables à des «perceptions du travail», mais que le comité de l'AQ avait conclu que le sexe des titulaires des emplois n'était aucunement entré en ligne de compte dans les évaluations finales. L'examen du rapport écrit ne contient aucune mention de ce rapport verbal que M. Renaud a fait à M. Durber.

573. M. Durber a conclu des travaux du comité de l'AQ qu'il ne serait pas exceptionnel de trouver parmi les évaluateurs un éventail de vues qui se refléteraient dans un éventail de cotes. Il a interprété les écarts entre les évaluations de M. Wisner et celles des comités d'évaluation comme traduisant «des désaccords normaux et honnêtes au sujet du travail plutôt qu'un parti pris quelconque contre l'un ou l'autre sexe».

574. Dans son témoignage, M. Durber a fait remarquer «que tout ce qui entoure la question de la partialité fondée sur le sexe dont le Tribunal est saisi repose sur l'opinion d'une seule personne [M. Wisner] concernant 25 questionnaires» (volume 149, à la p. 18581).

575. M. Durber a utilisé le rapport du comité de l'AQ pour comparer la moyenne des cotes attribuées par l'évaluateur du comité de l'AQ à la cote globale donnée par les comités d'évaluation dans le cadre de l'étude sur la parité salariale et par M. Wisner. Selon M. Durber, cette comparaison lui a permis de constater que le comité de l'AQ était aussi souvent en désaccord avec M. Wisner qu'avec les comités d'évaluation. Voici d'ailleurs ce qu'il a dit à ce propos dans son témoignage (volume 149, de la p. 18573, ligne 14, jusqu'à la p. 18574, ligne 22) :

[TRADUCTION]

D'après les tendances remarquées, les évaluateurs qui donnaient de faibles cotes étaient essentiellement aussi souvent d'accord qu'en désaccord avec les comités.

Les évaluateurs qui attribuaient de fortes cotes n'étaient d'accord avec les comités que le tiers du temps, bien qu'un tiers soit toujours un nombre raisonnable.

Je conclus de cette analyse qu'il est en fait normal de trouver un éventail de points de vue, un éventail de cotations des emplois, qu'il n'était pas exceptionnel de constater une variation dans les cotes et qu'il ne serait certainement pas exceptionnel de trouver des différences entre n'importe quels évaluateurs.

Ces résultats m'ont amené à croire... à interpréter les différences par rapport aux comités relevées par M. Wisner comme étant des désaccords normaux et honnêtes au sujet du travail plutôt qu'un parti pris quelconque contre l'un ou l'autre sexe.

Le fait qu'il s'agisse d'emplois masculins peut avoir été une coïncidence ou non, mais je ne voyais aucune raison de penser

143

qu'un parti pris était en jeu, étant donné les écarts entre M. Wisner et les comités.

Je n'ai pas, par exemple, conclu que M. Wisner avait un parti pris en faveur des emplois masculins, ce qui aurait pu être une interprétation du rapport. Puisqu'il était un homme, on aurait bien conclure cela. Mais la question de savoir s'il était un consultant professionnel et objectif était d'un autre ordre.

Mais nous avons en fait constaté que ces cinq individus du Comité directeur étaient aussi moins souvent en désaccord que M. Wisner, mais probablement aussi souvent ou un peu plus souvent que M. Willis lorsque celui-ci et ses trois autres consultants avaient examiné les postes masculins.

576. En interrogatoire principal, on a demandé à l'un des représentants syndicaux faisant partie du comité de l'AQ, M. Tim Yates, ce qu'il comprenait de la raison d'être du comité de l'AQ. Le comité, a-t-il répondu, avait pour mandat d'examiner les évaluations des comités et de vérifier si ces derniers avaient choisi les bons postes-repères et les avaient appliqués correctement. Pour autant qu'il se souvienne, le comité n'a relevé aucun cas où l'on avait choisi un poste-repère inapproprié. A propos des écarts entre les évaluations du consultant et celles des comités, M. Yates a fait les observations suivantes (volume 175, à la p. 22226, lignes 4-22) :

[TRADUCTION]

Q. Eh bien, si on veut faire une supposition énorme, que nous étions des experts du domaine, parfois nous étions plus haut que le consultant, parfois nous étions plus bas que le comité. Je pense que c'est M. Willis qui a dit à maintes reprises, «ce n'est pas une science exacte, ça».

Je dirais quant à moi, c'est quoi le problème? Tout me semble à l'intérieur de la tolérance.

Q. Qu'est-ce que vous voulez dire quand vous dites que tout vous semblait à l'intérieur de la tolérance? D'où vient cette expression?

R. Le plus petit écart possible est un échelon. Un échelon, c'est 15 %. Autrement dit, la nuance la plus petite qu'on peut établir pour un facteur donné est de l'ordre de 15 %.

577. L'autre représentante syndicale qui a témoigné concernant le comité de l'AQ a été Mme Mary Crich, qui avait été suppléante au Comité directeur et qui avait participé aux évaluations des comités en tant que membre du comité no 5. On lui a demandé de parler de ses observations. A bien y penser, elle trouvait que sa participation au comité de l'AQ avait été une bonne expérience parce qu'elle y avait compris que ce qu'elle avait accompli en tant que membre d'un comité était exactement ce que les comités d'évaluation étaient censés faire. Elle a constaté que, au sein du comité

144

de l'AQ, on décidait des évaluations au bout d'exactly les mêmes discussions sur les mêmes points et avec plus ou moins les mêmes accords ou désaccords que ce qu'elle avait connu dans son comité d'évaluation.

578. En ce qui concerne la compréhension que Mme Crich avait du travail du comité de l'AQ, celle-ci a témoigné que les personnes choisies pour faire partie de ce comité connaissaient la discipline du Comité directeur et pouvaient donc décider si les cotes des comités d'évaluation respectaient cette discipline ou si elles s'en écartaient sensiblement. Lorsque le comité de l'AQ eut terminé ses travaux, tous s'accordaient pour dire que les évaluations n'étaient pas biaisées. Lorsque l'écart était important, c'était, selon Mme Crich, parce que le poste était réellement difficile à évaluer et qu'il n'y avait pas de poste-repère comparable. Elle a qualifié les 25 postes en question de «très difficiles». En contre-interrogatoire, on a demandé à Mme Crich ce qu'elle pensait qu'on

entendait par «biaisées». Elle a répondu ce qui suit (volume 192, à la p. 24830, lignes 5-15) :

[TRADUCTION]

R. Ce que je me souviens d'avoir entendu les autres participants dire, c'est qu'il y avait eu des allégations dans les médias comme quoi les résultats étaient biaisés, et qu'on voulait dire par là que les évaluations n'avaient pas été justes de manière égale envers tous les postes et que les postes féminins avaient été cotés trop haut. Je ne sais pas si, aussi, les postes masculins avaient été cotés trop bas. C'était peut-être les deux. Peut-être que c'était juste un les deux, ou peut-être... en tout cas, ce qui faisait que c'était biaisé, c'était que les postes féminins avaient été cotés trop haut.

579. Mme Crich a par la suite apporté l'éclaircissement suivant à sa réponse (volume 192, à la p. 24841, lignes 2-12) :

[TRADUCTION]

Q. Monsieur [sic] Crich, je n'ai qu'une question, en somme, et j'essaierai de la formuler le plus clairement possible.

Lorsque votre comité d'assurance de la qualité a convenu qu'il n'y avait pas de partialité dans ces questionnaires, ces 27 questionnaires que vous avez évalués, cherchiez-vous les causes de la différence, et cela vous a-t-il amenés à conclure que le sexe du répondant n'expliquait pas les écarts?

R. C'est exact.

580. M. Willis a témoigné qu'il avait de nombreuses réserves à propos du comité de l'AQ. Il était déçu de la composition du comité et aurait préféré qu'on reconvoque tout le Comité directeur au lieu de nommer les seuls cinq individus en question. Un autre point qui le troublait était le fait que, bien que trois des membres fissent partie du Comité directeur, deux d'entre eux avaient agi uniquement à titre de suppléants. De plus, le

rapport Tristat avait identifié un des membres comme étant un évaluateur marginal (c'est-à-dire qui cotait différemment de l'ensemble du groupe). M. Willis estimait par ailleurs que, puisque deux des membres avaient

participé aux comités d'évaluation, leur opinion sur les résultats des comités devenait suspecte.

581. Un autre point qui préoccupait M. Willis était le fait que ces cinq individus n'aient pas effectué d'évaluations depuis au moins deux ans. Ce comité, a-t-il dit, aurait dû à tout le moins bénéficier d'une journée ou deux de recyclage par les consultants. A son avis, il était difficile de revenir après une interruption de deux ans faire des évaluations, surtout des évaluations qui allaient être critiquées. Il a exprimé sa plus grande crainte dans les termes suivants (volume 208, à la p. 26950, lignes 3-8) :

[TRADUCTION]

Toutefois, je suppose que ma plus grande préoccupation concernant le comité de l'AQ était que, c'est ce que j'ai compris, il n'y avait pas de règle de consensus. Pour moi, la phase du consensus fait en quelque sorte partie de la collecte des données.

582. M. Willis a souligné que dans le processus Willis la phase du consensus est très importante parce qu'elle permet aux membres du comité de discuter des faits de l'emploi et de prendre le temps d'examiner les informations qui en découlent. C'est l'«affinage de l'information» qui est important, selon M. Willis, à ce stade du processus où les membres du comité changent leur évaluation. Il est approprié de changer sa cote, a-t-il poursuivi, si on se base sur des faits issus des discussions. Pour cette raison, M. Willis n'accorde pas beaucoup de valeur aux résultats du comité de l'AQ parce qu'une étape essentielle et critique a été omise. Il accepterait jusqu'à un certain point de revenir sur son avis si le comité de l'AQ avait en fait inclus le processus de consensus dans ses délibérations.

583. M. Durber a témoigné que, dans le cours normal d'une enquête, la Commission s'attend à ce que l'employeur produise des preuves pour «sa défense». Or la Commission a reçu de l'employeur une défense qui dit, en réalité, qu'il devrait être dispensé de devoir accepter les résultats de sa propre étude. Il reste que l'employeur a été dûment représenté, mais qu'il n'a présenté aucune preuve à l'appui de sa conclusion.

584. Selon M. Durber, il est inévitable qu'il y ait des différences entre un comité et un consultant, mais la Commission se doit d'être vigilante lorsqu'elle cherche à comprendre ces différences et leur lien avec le sexe des titulaires des postes évalués.

585. La Commission a décidé de faire faire une autre analyse de l'uniformité des évaluations effectuées par les neuf comités par

comparaison à celles réalisées par le Comité directeur. M. Durber estimait qu'il n'avait d'autre choix que de commander une autre étude pour compléter le portrait. La solution de rechange ne lui plaisait pas car, à son avis, il était impossible de reproduire les évaluations d'emplois effectuées par

146

les comités. M. Durber doutait de la validité du processus qui consistait en quelque sorte, a-t-il expliqué, à demander à des gens de refaire ce qu'un nombre assez important de personnes avait accompli sur un certain laps de temps. M. Durber aurait préféré que les parties à l'étude sur la parité salariale règlent à leur façon la question de l'apparence de partialité fondée sur le sexe. Dans son témoignage, il a précisé ainsi sa pensée (volume 149, à la p. 18599, lignes 1-11) :

[TRADUCTION]

Mais il n'aurait pas été impensable de demander aux comités d'expliquer leurs résultats, d'examiner les écarts entre eux-mêmes et M. Wisner. On aurait très bien pu soulever des jugements ou avancer des idées sur les tendances elles-mêmes et sur les différences entre les comités et M. Wisner.

On aurait pu avoir de bonnes séances de justification, si vous voulez. Mais à la fin, cela n'a pas été possible. Une fois que les comités ont été dissous, il n'y en avait plus de comités.

586. M. Durber a expliqué que, au cours de son enquête, il n'a pas communiqué avec M. Wisner parce qu'il préférerait avoir comme point de référence ce qu'il considérerait comme «des critères raisonnables pour juger de la qualité de l'évaluation des emplois». La question qui se posait, d'après M. Durber, concernait les écarts entre les comités et les consultants ainsi que le processus suivi par les comités. M. Durber ne voyait pas pourquoi il aurait cru un consultant de préférence aux comités d'évaluation. S'il avait à choisir entre accepter le jugement d'un groupe de personnes bien informées et suivre la discipline d'un individu, M. Durber préférerait croire le groupe de personnes. Ce fut là une des «mesures indirectes» qu'il a utilisées pour tirer ses conclusions au sujet de la partialité fondée sur le sexe. M. Durber estime que s'il avait contacté M. Wisner, il aurait alors été obligé d'appeler chacun des membres des comités.

587. M. Durber a approché M. Willis au début de 1990 pour lui demander de faire une autre évaluation. M. Willis lui a confirmé son acceptation

par une lettre en date du 12 février 1990 dont nous reproduisons ci-dessous des extraits :

[TRADUCTION]

L'objet sera de déterminer l'ampleur de tout biais systématique qui pourrait exister dans les résultats des comités d'évaluation.

[...] l'échantillon devrait comprendre 300 postes, dont au moins 131 devraient provenir de groupes professionnels à prédominance masculine, les autres provenant des groupes professionnels à prédominance féminine.

Pour ce qui est de la sélection de l'échantillon, la méthode du choix au hasard que nous avons utilisée pour la première analyse serait, je crois, acceptable aux syndicats et à l'employeur. La

147

Commission des droits de la personne devrait avoir son mot à dire dans le choix de cette méthode [...]

La méthode d'analyse sera la même que celle utilisée dans notre analyse antérieure. Le consultant examinera chaque questionnaire choisi et décidera s'il existe un poste semblable parmi les évaluations du Comité directeur. Lorsqu'un poste-repère semblable sera trouvé, l'évaluation du Comité directeur sera adoptée comme l'évaluation du consultant. S'il n'existe aucun poste-repère semblable, le consultant effectuera une évaluation indépendante du poste en s'appuyant sur les postes-repères appropriés du Comité directeur. Il comparera ensuite son évaluation à l'évaluation type du comité ainsi qu'à la justification raisonnée pour ce poste. S'il y a une différence entre l'évaluation du consultant et celle du comité, le consultant rédigera une justification raisonnée expliquant son évaluation.

(pièce HR-93)

588. M. Durber a expliqué que son objectif en chargeant M. Willis de réévaluer les 300 postes additionnels était essentiellement de poursuivre l'examen de la question qui avait été soulevée à la suite de la Wisner 222 relativement à la partialité fondée sur le sexe. Compte tenu de l'article 9 de l'Ordonnance, supra, M. Durber voulait s'assurer qu'il n'y avait aucune partialité fondée sur le sexe. Il a par ailleurs fait remarquer qu'il ne voyait pas comment on aurait pu ne pas utiliser la même

approche que M. Wisner, étant donné que c'était par cette approche que la question avait été soulevée au départ.

589. M. Durber aurait préféré confier les deuxièmes réévaluations à M. Wisner mais, entre temps, ce dernier avait quitté le cabinet de M. Willis. Par conséquent, M. Willis a été autorisé à former un comité de quatre consultants (collectivement connus comme le «groupe des quatre»), qui allait effectuer les 300 réévaluations (l'étude «Willis 300»).

590. M. Willis a témoigné avoir compris que l'on craignait que les quatre consultants travaillant ensemble obtiennent des résultats légèrement différents de ceux de M. Wisner. Par conséquent, son groupe avait comme tâche additionnelle de sélectionner des emplois parmi la Wisner 222 et de les évaluer de façon indépendante sans porter de jugement sur les écarts entre les réévaluations du groupe des quatre et celles de M. Wisner. Avec le groupe des quatre, M. Willis devait réviser environ 20 % des évaluations de la Wisner 222, soit 44 questionnaires, pour contre-vérifier l'interprétation que M. Wisner avait faite des emplois.

591. Le groupe des quatre a tâché de reproduire le plus fidèlement possible la méthodologie qui avait été utilisée dans la Wisner 222. C'est la Commission qui a sélectionné l'échantillon à partir de l'échantillon total des évaluations, mais excluant les évaluations du Comité directeur et les réévaluations comprises dans la Wisner 222. On n'a pas demandé à M. Willis d'analyser ces réévaluations. Après que le groupe des quatre eut terminé les 300 réévaluations, il a transmis les résultats à la Commission pour analyse.

148

592. Le groupe des quatre comprenait, en plus de M. Willis, ses deux associés Owen et Davis ainsi qu'une consultante externe bilingue, Mme Esther Brunet. Des questionnaires étaient assignés à chaque consultant et un deuxième consultant revoyait chacune des évaluations, de sorte qu'il y avait toujours deux consultants en cause. Le travail a pris environ deux mois. En mars 1990, M. Willis a présenté à la Commission un rapport intitulé Report to the CHRC Equal Pay, Quality Analysis of Sampled Committee Evaluations, Joint Initiative Equal Pay Study.

593. Bien que cette étude eût pour objet d'évaluer la qualité de la Wisner 222, la Commission a donné instruction à Willis & Associates de ne pas tirer de conclusion sur la qualité de son travail ou des réévaluations Wisner. Au cours de ces audiences, et dans le contexte de cette étude, on a demandé à M. Willis ce qu'il pensait de la qualité des réévaluations Wisner. Il a répondu ce qui suit (volume 59, à la p. 7337, lignes 11-24) :

[TRADUCTION]

LE TÉMOIN : J'étais satisfait de la qualité des évaluations Wisner six ou sept mois auparavant lorsque j'ai examiné ses justifications raisonnées et ses évaluations proprement dites. J'avais énormément confiance dans les habiletés de M. Wisner en tant qu'évaluateur d'emplois consciencieux.

Je n'ai pas, à ce stade, additionné les 44 évaluations effectuées par les consultants de notre équipe pour les comparer globalement aux évaluations de M. Wisner. Elles n'étaient pas identiques, il y avait certaines différences. Mais j'estimais qu'il revenait à M. Durber d'analyser ces différences et, en fait, de décider si la même qualité s'observait chez les deux équipes de consultants.

594. Pour ce qui est d'analyser les résultats des 300 évaluations, M. Willis a déclaré qu'il aurait été approprié, selon lui, de faire une analyse statistique pour déceler l'existence ou la non-existence d'une partialité systématique fondée sur le sexe. Si la Commission le lui avait demandé, il aurait retenu les services de M. Milczarek, le statisticien qui accomplit normalement ce genre d'analyse pour lui.

595. La dernière communication entre Willis & Associates et la Commission concernait les 44 réévaluations. Il s'est agi d'une lettre datée du 1er mai 1990 que le consultant de M. Willis, Keith Davis, a envoyée à la Commission. Pendant la réévaluation des 300 postes et l'examen des 22 réévaluations Wisner, le groupe des quatre s'est par inadvertance reporté à une liste ayant trait au facteur conditions de travail dans le plan Willis. Le Comité mixte avait apporté à ce facteur des modifications que le groupe des quatre n'a pas pris en considération. M. Davis a informé la Commission que lorsqu'on utiliserait les réévaluations il faudrait modifier le facteur conditions de travail. Au bout du compte, une seule réévaluation effectuée par un consultant a dû être modifiée.

596. Le Tribunal a pu entendre le témoignage de Mme Esther Brunet concernant sa participation aux réévaluations à titre de membre du groupe

des quatre. Mme Brunet était la seule membre du groupe qui était une fonctionnaire fédérale. Elle avait pris part à l'étude sur la parité salariale en tant que présidente de la première version du comité no 4. A l'époque pertinente, elle était directrice du personnel, Finances et Administration, Condition féminine Canada. M. Willis a témoigné qu'il

avait besoin d'un consultant francophone pour participer à la Willis 300 et que, puisque lui et son personnel avaient énormément confiance dans la capacité de Mme Brunet, il l'avait engagé pour qu'elle évalue les questionnaires français.

597. Mme Brunet a évalué environ 100 des 300 questionnaires. Environ 70 % de ceux-ci étaient des questionnaires français. Elle a d'abord évalué les questionnaires de façon indépendante. Si le comité d'évaluation avait utilisé un seul poste-repère, elle essayait d'en trouver d'autres. Une fois son évaluation terminée, elle regardait les cotes du comité et ses justifications raisonnées, et si l'explication de l'écart avait du sens, elle accordait le bénéfice du doute au comité, sinon, elle rédigeait sa propre justification qu'elle présentait aux trois autres consultants. Durant son exposé, elle cherchait à convaincre ses trois collègues de la nécessité de retenir le changement qu'elle proposait. Si elle ne réussissait pas à les convaincre, les cotes du comité demeuraient telles quelles. Mme Brunet a expliqué que le groupe des quatre ne rédigeait pas les justifications raisonnées de la même façon que les comités parce que la raison pour laquelle il les rédigeait était simplement de justifier la différence entre le consultant et le comité.

598. On peut comparer les évaluations que Mme Brunet a faites des questionnaires français aux cotes des comités puisqu'elle était la seule consultante du groupe des quatre à évaluer les questionnaires français. Ces derniers sont résumés dans la pièce PIPSC-162, qui confirme que, pour les postes à prédominance féminine, la cote moyenne de Mme Brunet était 157,1 comparativement à 157,9, la cote moyenne des comités. Dans le cas des postes à prédominance masculine, sa cote moyenne a été de 250,7, celle des comités, de 249,7. Mme Brunet a attribué la même cote que les comités sauf dans huit cas, cinq étant des postes féminins, trois des postes masculins.

599. Les présentes enquêtes menées en vertu de l'article 11 différaient quelque peu de celles que la Commission avait l'habitude de mener, en ce sens qu'en l'espèce elle connaissait les faits pertinents pour avoir participé au processus à titre d'observateur depuis le tout début. Les observateurs de la Commission avaient assisté aux réunions du Comité mixte et observé de manière suivie les comités d'évaluation durant leurs travaux depuis le début de l'étude. La Commission n'avait pas assez d'observateurs pour assister à toutes les séances des comités, et le nombre des observateurs a diminué au fil des ans.

600. Les observateurs prenaient quotidiennement des notes lorsqu'ils assistaient aux travaux des comités d'évaluation (pièce R-142). Ces notes, qui étaient assez détaillées, ont été produites en preuve durant le contre-interrogatoire de M. Durber. Ce dernier n'avait pas lui-même lu les

notes. Il a demandé à M. Brian Hargadon, un des observateurs de la Commission, s'il y avait dans ces notes quelque chose concernant la façon

150

de travailler des comités, notamment, que la Commission devait explorer dans son enquête. M. Durber a témoigné avoir reçu de M. Hargadon un aperçu faisant état des difficultés rencontrées, notamment en ce qui concerne le processus proprement dit d'évaluation des emplois, la réalisation d'un consensus et l'examen des points litigieux. En dernière analyse, cependant, il ne s'y trouvait rien qui aurait été de nature à remettre en question la fiabilité des résultats. Par conséquent, M. Durber était d'avis qu'il n'était pas nécessaire de produire les notes des observateurs en preuve à la présente audience du Tribunal.

601. En contre-interrogatoire, on a lu à M. Durber des extraits des notes des observateurs et on lui a demandé s'il avait obtenu de M. Hargadon ou dans un autre contexte les informations sur lesquelles il avait fondé ses conclusions concernant l'étude sur la parité salariale. Voici certains des extraits qui ont été lus :

[TRADUCTION]

Comité no 5 :

[...] Il semble se développer un parti pris (sexe) dans ce comité.

Il y a une femme (Sherry) qui donne des cotes plus élevées que le reste du groupe pour les postes à prédominance féminine et des cotes plus basses pour les emplois à majorité masculine. Elle prétend aussi avoir une connaissance directe de la plupart des postes et, lorsqu'elle les décrit, elle fait des commentaires très subjectifs qui reflètent son parti pris. Elle change rarement sa cote même si elle a adopté une position extrême.

Il y a un homme dans le groupe (Paul) dont les cotes reflètent le parti pris opposé. Toutefois, ses cotes ont tendance à se rapprocher davantage de la cote de consensus.

Il y a une autre femme dans le groupe (Mary) qui, dans les discussions, semble avoir une forte alliance avec Sherry. Les cotes de Mary, cependant, ne semblent pas refléter de partialité.

Les discussions tendent à s'éterniser dans ce groupe parce qu'il y a constamment des vues qui s'opposent [...]

(pièce R-142, volume I, page 6)

Fonctionnement des comités :

En général, les comités ont adopté des routines qui sont efficaces et qui reflètent également le caractère unique de chaque groupe. Compte tenu du fait que les conditions de travail ne sont pas idéales (l'horaire est serré et des individus aux personnalités et vues très différentes doivent travailler ensemble pendant de longues heures), les comités travaillent bien.

151

Toutefois, il y a certains problèmes qui doivent être surveillés. Je ne connais pas assez bien le comité no 3 pour commenter. Le comité no 5 a lui aussi ses problèmes qui nuisent à sa productivité, encore que ce ne soit pas aussi grave que dans le comité no 3.

Les membres du comité no 5 ont de la difficulté à écouter les vues des autres. Ils s'interrompent constamment et souvent la charge émotive est très élevée au sein du comité.

Le comité no 5 a besoin d'un président qui a de la «poigne» à cause des personnalités opposées et fortes. Le président actuel ne semble pas avoir cette capacité [...]

(pièce R-142, volume I, page 90)

Comité no 3 :

Scissions entre syndicat et employeur. Emploi bien écrit et complet. Louise a changé sa cote pour qu'elle soit conforme à celle de Jake et Al sur le facteur connaissances et qualifications. Aucune amélioration dans le fonctionnement du comité. Climat tendu.

Comité no 2 :

Le comité travaille bien.

(pièce R-142, volume II, page 125)

Comité no 4 :

Il a fallu 5 heures pour évaluer ce poste (simple). Les nouveaux ont prolongé le processus, s'obstinaient, même après que le consultant eut clarifié les choses.

(pièce R-142, volume II, page 130)

Comité no 4 :

[...] On tient compte du sexe du répondant. Le président essaie d'influencer les évaluateurs.

(pièce R-142, volume II, page 176)

Comité no 5 :

[...] Pierre Collard a noté une engueulade dans le comité no 5. Il pense qu'elle a pu avoir été influencée par le fait que certains membres du comité n'auraient plus d'emploi une fois ce processus terminé [...]

152

Mercredi - la Section de la parité salariale, CCDP, a reçu un appel du CT pour intervenir dans une dispute entre 2 membres du comité no 5.

Jeudi - Brian H. et moi avons circulé dans le comité et tout était tranquille.

(pièce R-142, volume II, page 203)

Réunion hebdomadaire, le lundi 31 octobre 1988 :

Ron [Renaud] a soulevé le point qu'on ne suivait pas les règles de base concernant le consensus, ce qui fait que, une fois que tout sera fini, une partie pourrait dire que l'accord n'était pas valide parce qu'on n'avait pas suivi la règle, telle qu'elle est formulée dans les lignes directrices.

(pièce R-142, volume I, page 36)

Notes additionnelles de l'observateur datées du 24 novembre 1988 :

3. Vote majoritaire. Les comités 3 et 5 ont de la difficulté avec ça. Apparemment, ils ne suivent pas les règles de consensus

énoncées à la page 2 des procédures. Le comité 2 suit les instructions sans exception [...]

On discute également de la possibilité d'obtenir un consensus en utilisant la médiane. C'est Fred qui a suggéré cela. Par exemple, supposons que pour le facteur conditions de travail vous avez les cotes 13, 13, 15, 17, 17. On devrait choisir 15 comme cote. Est-ce une solution?

(pièce R-142, volume I, page 79)

Comités de l'étude sur la parité salariale - Observations :

Aujourd'hui, lors de ma visite au comité no 3, j'ai remarqué qu'on n'observait pas la règle des deux tiers pour obtenir un consensus. Le comité avait décidé de prendre une valeur moyenne comme consensus, mais on m'a consulté sur la question et les membres ont décidé de suivre mon conseil. Par ailleurs, j'ai entendu le commentaire suivant : «Nous ne suivons pas cette règle à moins qu'il y ait quelqu'un ici qui nous observe.»

(pièce R-142, volume II, page 102)

Comité no 6 :

[...] Aussi, le comité a fait une supposition pour le facteur conditions de travail parce qu'il estimait que le titulaire n'avait pas fourni toutes les informations voulues dans le questionnaire.

153

(pièce R-142, volume II, p. 187)

Notes de Brian Hargadon à Ted Ulch :

Je vois un problème ou deux, au moins, dans le comité no 2.

[...]

On n'utilise pas les postes-repères comme on le devrait. On nous dit qu'en tant que comité nous ne sommes pas obligés de les suivre. Est-ce le cas?

(pièce R-142, volume II, page 111)

Comité mixte :

[...] Keith et Sharon ont commenté l'analyse qu'ils faisaient dans leurs comités respectifs. Tous les représentants de la CCDP ont fait remarquer qu'il était évident que certaines personnes cotaient toujours les questionnaires trop haut ou trop bas et que, si les résultats des tests n'étaient pas analysés assez rapidement, il pourrait être trop tard pour rectifier la situation.

[...] on nous a incités à ne pas intervenir sur le plan personnel, par exemple, en pointant du doigt quelqu'un qui ne serait pas à la hauteur, parce que cela risque de nous hanter par la suite. Certains membres des comités, pense-t-on, surtout ceux des syndicats, reçoivent des instructions sur la façon de traiter les évaluations de manière à servir le mieux possible les intérêts d'un groupe de syndiqués particulier.

(pièce R-142, volume I, page 56)

Comité no 2 :

Pour le poste 2317, le comité n'a pas suivi le poste-repère du Comité directeur et semble avoir surestimé sa cote [...]  
Commentaire entendu : «Cela fait réellement une différence quand vous êtes ici pendant les évaluations.» Les gens ici ne discutent pas du tout des postes.

(pièce R-142, volume II, page 180)

Activités hebdomadaires :

[...] Le problème est que le comité 5 a plus de 100 postes à récapituler et on se demande comment il a pu laisser monter ce nombre à ce point.

(pièce R-142, volume II, page 208)

Réunion hebdomadaire, le 8 novembre 1988 :

[...] Les membres du comité ont soulevé de nombreuses incohérences qui ont été notées dans les divers comités. Ils s'inquiètent de ne pouvoir faire plus que poser des questions lorsqu'il est

évident que, dans d'autres comités, on ne respecte pas une norme donnée.

Par exemple, un comité a décidé que le niveau D sous connaissance de l'emploi ne pouvait être demandé de nouveau que si l'emploi exigeait un grade universitaire. Ron a demandé au consultant si c'était le cas; réponse : ce n'était pas correct.

Ron rédigera une note plus détaillée qu'il remettra à Ted sous pli séparé. On craint vraiment que ces incohérences se poursuivent et s'accroissent, et que, au bout du compte, il s'en trouve pour contester la crédibilité des comités, voire la nôtre [...]

Le comité no 3 continue d'avoir des problèmes. Le comité arrête ses cotes, puis il cherche ensuite un poste-repère qui correspond à son évaluation, au lieu de confronter sa cote avec un poste-repère approprié.

(pièce R-142, volume I, page 45)

La question de l'uniformité - Comité mixte :

J'aimerais signaler à votre attention une question que je trouve importante à ce stade-ci de l'étude et qu'on devrait à mon avis porter à la connaissance du Comité mixte.

Essentiellement, nous devrions confirmer notre position que l'uniformité est importante; uniformité dans le respect de la discipline du Comité directeur et uniformité dans l'application du plan Willis par les cinq comités d'évaluation.

[...]

Il y a eu de nombreux cas, en plus de ceux mentionnés plus haut, où des comités ont suivi pour certains postes un processus d'évaluation incompatible avec [la discipline du] Comité directeur et ce que faisaient les divers comités d'évaluation [...]

- Il y a des situations comme celles-là qui nous inquiètent à cause des incohérences et nous nous demandons comment faire en sorte qu'elles soient corrigées le plus tôt possible sans compromettre notre rôle.

En résumé, je recommande qu'on informe le Comité mixte de notre opinion quant à la façon de traiter les situations de «suppléance». De plus, il serait opportun de confirmer notre

position sur l'importance de l'uniformité : uniformité par rapport à la discipline du Comité directeur et uniformité dans l'application du plan par les cinq comités d'évaluation.

155

(pièce R-142, volume I, page 47)

Mise à jour des remarques des observateurs, le 7 décembre 1988 :

Les observateurs ont décidé qu'ils voulaient examiner un certain nombre de points qui les préoccupaient, alors ils se sont réunis ce matin.

Avant d'aborder des points précis, je veux confirmer que certains comités ont commencé à manifester des craintes durant les tests [...]

La raison pour laquelle nos cotes étaient diminuées dans les comités a été discutée avec les observateurs pour que nous puissions donner la même raison : a) d'autres engagements et b) les comités ont maintenant besoin de moins d'observation parce que cela fait un certain temps qu'ils fonctionnent.

Les comités 1, 2 et 4 marchent pas mal bien. Le comité 5 a toujours des problèmes, qui vont probablement se tasser, toutefois.

Le comité 3 ne fonctionne toujours pas comme il le devrait. On se demande si les observateurs qui restent, Sharon et Keith, devraient consacrer un temps disproportionné au comité 3 à cause du problème. Alors la question se pose toujours, accordons-nous la préférence au comité no 3?

Lorsque nous retournons en arrière et que nous examinons la raison pour laquelle les observateurs de la CCDP sont entrés en jeu, nous nous inquiétons que nos efforts n'aient servi à rien si a) le Comité mixte est dissout, ou b) nous devons attester de la crédibilité tant du Comité directeur que des cinq comités courants.

Dans l'état actuel des choses, aucun observateur n'affirmerait que les évaluations sont justes, équilibrées et objectives. Il y a trop d'irrégularités à l'intérieur des comités et d'un comité à l'autre.

(pièce R-142, volume I, pages 82-87)

602. Il ne faut pas oublier que le rôle de l'observateur était celui d'un «chien de garde» dans le processus d'évaluation en comité. Il devait observer, critiquer et, lorsqu'on le lui demandait, proposer des améliorations au fonctionnement du comité. On doit lire les notes des observateurs dans ce contexte.

603. M. Durber a accepté la conclusion de l'observateur de la Commission, M. Hargadon, et a décidé de ne pas se fier aux notes comme preuve de la fiabilité des résultats d'évaluation.

156

604. Le Tribunal a entendu le témoignage d'évaluateurs faisant partie des comités qui ont témoigné en réponse à certaines notes des observateurs ayant trait à leur comité en particulier. Après avoir étudié les réponses de M. Durber aux questions soulevées durant son témoignage, et compte tenu du caractère vague et non spécifique de ces notes ainsi que des réponses des évaluateurs qui ont témoigné à l'audience, le Tribunal conclut qu'en fait les notes n'ont pas un effet négatif important sur la question plus générale de la fiabilité.

605. Un autre aspect de l'enquête de la Commission consistait à faire revoir, par un comité de trois personnes établi par M. Durber, les réévaluations que le Conseil du Trésor avait effectuées des postes-repères pour les groupes Sciences infirmières, Sciences domestiques, Ergothérapie et physiothérapie et Gestion des systèmes d'ordinateur. Ces réévaluations sont contenues dans les deux rapports que le Conseil a présentés à la Commission en juillet 1990, en réponse à l'enquête de celle-ci sur l'apparence de partialité fondée sur le sexe dans les résultats d'évaluation. Les rapports sont intitulés Evaluation of CS Benchmarks and Corrected Version of NU, Annex B (pièce HR-252) et Final Report on Evaluation of Equal Pay Study Questionnaire (pièce HR-253).

606. La Commission avait demandé au Conseil du Trésor si l'employeur souscrivait aux observations avancées dans ces rapports et qui soulevaient des questions sur les évaluations des emplois effectuées par les comités d'évaluation multiples. La Commission n'a pas reçu de réponse à ses questions. M. Durber a conclu qu'on pouvait considérer ces rapports comme preuve éventuelle dans l'enquête, mais que, dans l'immédiat, il les excluait comme élément de preuve valable dans l'enquête de la Commission. Il se réservait toutefois l'option d'exposer ces documents au Tribunal plus en détail. Quoi qu'il en soit, M. Durber a décidé de demander à un comité

(le «comité d'examen des postes-repères») d'explorer les rapports sur le fond.

607. Le comité d'examen des postes-repères regroupait Mme Esther Brunet, Mme Christine Roberge, une employée de la Commission, et M. Brian Hargadon, un enquêteur de la Commission. M. Hargadon et Mme Roberge ont été formés par M. Willis. Au début de septembre 1990, les trois participants, à l'aide du processus Willis, ont commencé à réévaluer chacune des évaluations contenues dans les rapports du Conseil du Trésor. Il y avait 65 questionnaires de postes-repères en tout. Ils ont aussi examiné 203 évaluations des comités multiples concernant les groupes OP, HE, NU et CS. Selon le processus défini par M. Durber, les trois membres du comité devaient se mettre d'accord sur l'évaluation de chaque poste réévalué. Une fois le consensus établi, le comité comparait sa cote à celle du consultant du Conseil du Trésor et à celle des comités d'évaluation multiples.

608. Si la cote du comité d'examen des postes-repères était différente de celles du Conseil du Trésor et des comités d'évaluation multiples, on tâchait de trouver la raison qui expliquait l'écart. Puis le comité d'examen accordait le bénéfice du doute aux consultants du Conseil du Trésor ou aux comités d'évaluation, ou il décidait de maintenir sa propre

157

cote en justifiant pourquoi elle différait de celles du Conseil et des comités.

609. Comme M. Durber n'a pas été informé par le Conseil du Trésor de l'objet des rapports présentés en juillet 1990, ses conclusions étaient principalement fondées sur les conclusions contenues dans le rapport du comité d'examen des postes-repères.

610. Mme Brunet n'a pas participé à la rédaction du rapport final du comité (pièce HR-254), qui a été rédigé par les membres Roberge et Hargadon, puis révisé par M. Durber. La conclusion du rapport, attestée par M. Durber, est qu'aucun poids ne devrait être accordé aux rapports du Conseil du Trésor. A quelques très rares exceptions, le comité d'examen des postes-repères a confirmé les évaluations qui avaient été faites dans le cadre de l'étude sur la parité salariale.

611. Une première ébauche du rapport HR-254, datée de juin 1991, avait été rédigée par les deux membres du comité. Cette première version a été admise en preuve sous la cote R-140 lorsqu'on a contre-interrogé M. Durber. Il y a deux passages, aux pages 26 et 27, qui n'ont pas été inclus dans la

version finale. Ces pages parlent des séances de récapitulation et des difficultés rencontrées par les comités d'évaluation dans l'utilisation des postes-repères. C'est M. Durber qui a retiré ces pages du rapport. A son avis, ces pages n'avaient pas beaucoup «rapport à ce qu'il [le comité d'examen des postes-repères] faisait [...]» (volume 159, à la p. 19790). M. Durber a donné l'instruction qu'on retire ces pages du rapport final. Selon lui, on y trouvait des observations intéressantes sur les difficultés liées aux postes-repères, mais rien que la Commission ne savait pas déjà. D'après lui, ces commentaires allaient être plus utiles pour des initiatives futures en matière de parité salariale.

612. M. Durber a témoigné avoir parlé à Mme Roberge et à M. Hargadon des questions soulevées aux pages 26 et 27 de la première ébauche du rapport (pièce R-140). Ces derniers lui ont expliqué qu'en rédigeant ces deux pages ils avaient voulu commenter «les leçons apprises, et faire part de leurs propres perceptions des obstacles auxquels la Commission pourrait se heurter en s'acquittant de son rôle d'observateur dans le contexte d'initiatives futures». La Commission aurait été au fait, ainsi, des problèmes qui s'étaient présentés durant l'étude sur la parité salariale, notamment en ce qui a trait aux justifications raisonnées. Aux yeux de M. Durber, ces commentaires ne constituaient pas une preuve solide, mais plutôt des observations utiles pour le travail futur de la Commission.

613. En ce qui a trait au rapport du comité d'examen des postes-repères, M. Durber estimait que les questions commentées aux pages 26 et 27 seraient soulevées par M. Willis durant les audiences du Tribunal. Il a fait valoir que la Commission n'avait ni les ressources ni le temps d'amorcer une enquête sur les méthodes du Comité directeur lorsqu'elle se préparait en vue des présentes audiences.

614. Le Tribunal a pu entendre le témoignage suivant de Mme Brunet au sujet des pages 26 et 27 de la pièce R-140 (volume 214, à la p. 27852, lignes 5-15) :

[TRADUCTION]

J'ai souri quand j'ai lu les pages 26 et 27 parce que lorsque je travaillais avec Christine et Brian Hargadon, Jim Sadler dirigeait l'étude dans les Territoires du Nord-Ouest. Il venait souvent voir comment ça allait, etc. Une fois, quand il nous a dit qu'il devait se rendre dans les Territoires, nous lui avons dit : «Que dirais-tu qu'on partage avec eux certaines informations que nous avons?»

En lisant les pages 26 et 27, j'ai vu qu'il y avait beaucoup de moi là-dedans.

615. Mme Brunet avait l'impression qu'on lui demanderait de revoir et de signer le rapport. En fait, ce n'est pas ce qui est arrivé, mais elle a toutefois reçu une copie du rapport. Elle a signalé que pendant que le comité accomplissait son travail, M. Jim Sadler, un employé de la Commission qui dirigeait une étude sur la parité salariale dans les Territoires du Nord-Ouest, venait souvent voir comment le comité d'examen des postes-repères fonctionnait. Le comité lui a proposé de partager des informations avec lui pour qu'il en tienne compte dans l'étude en cours dans les Territoires.

616. L'interprétation de Mme Brunet des observations figurant aux pages 26 et 27 et l'opinion exprimée par M. Durber quant à leur utilité sont corroborées dans la pièce R-141, une lettre que M. Sadler a adressée à un représentant syndical qui participait à une étude sur l'équité salariale dans les Territoires du Nord-Ouest. Cette étude conjointe sur la parité salariale -- connue sous son sigle anglais JEPS, Joint Equal Pay Study -- utilisait une version plus récente du plan Willis. Certains des commentaires que M. Sadler faisait dans cette lettre étaient basés sur les entretiens qu'il avait eus avec des membres du comité de la Commission. Ces entretiens corroborent le témoignage de M. Durber et de Mme Brunet sur la question de savoir comment le comité envisageait le partage de cette information avec la Commission.

617. Les opinions et conclusions de M. Durber concernant les pièces R-140 (rapport préliminaire) et R-142 (notes des observateurs) l'ont amené à décider de ne pas inclure ces documents dans la preuve de la Commission. L'audience que tient le Tribunal consiste en une enquête publique et le rôle de la Commission est de représenter l'intérêt public. Toute décision relative à la pertinence de la documentation réunie par la Commission durant son enquête relative à la plainte fondée sur l'article 11 est du ressort de celle-ci. Dans les circonstances que nous connaissons, cependant, la décision de la Commission d'exclure ces documents de sa preuve peut être critiquée s'il est établi que les documents sont pertinents et jettent une lumière particulière sur la question de la fiabilité.

618. Avant de poursuivre, le Tribunal estime que la Commission aurait dû produire en preuve, avec des notes explicatives, les rapports en question, à savoir les pièces R-140 et R-142. C'est le Tribunal qui aurait

dû décider de leur utilité. La cause de la Commission aurait été mieux servie si toutes les pièces avaient été produites au départ.

619. En contre-interrogatoire, M. Durber a fourni d'autres explications pour justifier le fait qu'il n'avait pas interrogé M. Wisner. Il a fait un examen ex post facto des justifications raisonnées de M. Wisner dans un but de clarification. M. Durber a utilisé à la fois les justifications du comité et celles de M. Wisner pour cette analyse. Il s'agissait d'examiner chaque écart et de déterminer dans quelle mesure les écarts s'annulaient les uns les autres. Après avoir catégorisé les écarts entre le comité et le consultant, M. Durber vérifiait les nombres en vue de déterminer si la distribution des écarts présentait une tendance ou si elle était aléatoire.

620. Le Tribunal a demandé à M. Willis de commenter l'analyse que M. Durber avait faite à partir de l'examen des justifications raisonnées. M. Willis a répondu (volume 208, à la p. 26939) qu'il avait des réserves à propos des conclusions de M. Durber. Il doute beaucoup qu'on puisse déceler la partialité en examinant les justifications. A son avis, la partialité est quelque chose de très subtil qu'il faut examiner poste par poste. Voici ce qu'il avait à dire à ce sujet (volume 208, à la p. 26939, lignes 8-13) :

[TRADUCTION]

Il faut examiner l'ensemble de la situation. Selon moi, il est tout à fait inapproprié de prendre uniquement certaines de ces réévaluations et de dire «Nous ne tiendrons pas compte de celles-ci». Ou bien on les prend toutes et on les examine pour ce qu'elles sont, ou bien on ne fait rien du tout.

621. Selon M. Willis, si ses consultants font une évaluation durant l'étude, les raisons des écarts sont très importantes car elles fournissent aux consultants des éléments utiles pour le recyclage des comités. M. Willis reconnaît qu'il y a toujours une certaine variation et qu'il y aura toujours des écarts aléatoires une fois l'étude terminée; il ne se préoccupe donc pas, à ce stade-ci, des raisons. Par rapport à l'analyse de M. Durber, M. Willis a fait remarquer qu'il s'attend toujours à trouver des différences entre les consultants et les comités, mais qu'il ne voit aucune utilité à se servir de ces différences pour chercher à déceler l'existence d'un problème éventuel. Il a apporté les précisions qui suivent à cet égard (volume 208, à la p. 26944, lignes 17-23) :

[TRADUCTION]

Ce que j'ai dit, ou du moins ce que j'ai voulu dire, c'est que puisque la partialité est quelque chose de très subtil, la seule façon que nous avons de savoir dans quelle mesure on a pu interpréter différemment les postes masculins et féminins consiste à examiner l'ensemble des résultats une fois l'étude terminée.

622. M. Durber a présenté son analyse dans un format mathématique -- sous forme de nombres et de tableaux -- et il y tire des conclusions sur

160

la symétrie et les tendances qui se dégagent de ces nombres. Cette analyse, selon le Tribunal, a une composante statistique à cause de la méthode particulière utilisée par M. Durber. Sans l'aide d'un expert en statistique, nous sommes incapables d'interpréter l'analyse de M. Durber, que nous avons donc décidé de ne pas prendre en considération.

623. En 1992, lorsque M. Willis a comparu devant le Tribunal, M. Durber a décidé d'étudier plus en profondeur la qualité des informations sur les postes contenues dans les questionnaires. Il a donc retenu les services d'un chercheur n'ayant aucune expérience de l'évaluation d'emploi, mais ayant «une capacité d'analyse assez solide», pour qu'il examine un échantillon de questionnaires. L'échantillon comprenait 650 questionnaires : 63 portaient sur des postes-repères et 587 sur les autres postes. Le chercheur n'a pas comparu devant le Tribunal.

624. Le chercheur avait comme mandat de vérifier les informations pour déterminer si elles étaient complètes, cohérentes et lisibles, et de déterminer si on avait appliqué les mesures de protection ou d'appoint. Enfin, il devait vérifier si chaque questionnaire avait été validé par le superviseur de l'employeur.

625. M. Durber a témoigné avoir discuté avec le chercheur de certaines des caractéristiques à surveiller pour décider si un questionnaire était complet ou non. A cet égard, il a rédigé à l'intention du chercheur certaines procédures et questions. A titre d'information, il lui a communiqué l'objet des informations sur l'emploi, le processus suivi durant l'étude pour recueillir et trier les informations, ainsi que des renseignements lui permettant de relever les données de base, notamment le ministère, le numéro du questionnaire et le groupe professionnel.

626. Le chercheur a mis deux mois pour s'acquitter de sa tâche. Il rencontrait M. Durber toutes les semaines pour discuter des problèmes rencontrés. M. Durber a lui-même revu tous les questionnaires problématiques, soit environ 5 % des questionnaires. Il a témoigné avoir

surveillé attentivement le chercheur pendant qu'il examinait les questionnaires des postes-repères.

627. Voici la liste des critères que le chercheur a utilisés dans sa vérification :

1. Lisibilité - Peut-on lire le questionnaire?
2. Langue - Le questionnaire a-t-il été rempli en français ou en anglais?
3. Présentation - Le questionnaire a-t-il été rempli à la main ou à la machine?
4. Signature - Le répondant a-t-il signé son questionnaire?
5. Commentaires - Le surveillant a-t-il fait des commentaires?
6. Exhaustivité - Toutes les parties du questionnaire ont-elles été remplies?
7. Concordance - Le surveillant était-il du même avis que le titulaire?
8. Notes - Y a-t-il une indication de notes provenant d'un réviseur?
9. Faits - Le répondant présente-t-il des faits ou un commentaire?

161

628. Le rapport intitulé *An Examination of the Quality of Questionnaire Information used by the Federal Pay Equity Study* (pièce HR-245) contenait des constatations et des conclusions au sujet de l'exhaustivité et de l'exactitude des informations sur les emplois. De l'avis du Tribunal, M. Durber exprime, dans le rapport et dans son témoignage oral, les opinions de son chercheur, qui peuvent être justifiées ou non. Étant donné le manque d'expérience du chercheur en évaluation d'emplois dans un contexte de parité salariale, le Tribunal estime qu'il doit rejeter toute opinion contenue dans ce rapport. Par contre, il accepte les constatations de fait que le rapport contient, lesquelles sont à notre avis utiles. En voici la teneur :

[TRADUCTION]

Constatations :

- .. On a répondu aux questions requises dans 95 % des cas.
- .. Le surveillant a signé le questionnaire dans un peu plus de 99 % des cas. Sur 96 % questionnaires, le surveillant a inscrit des commentaires où, dans environ 9 % des cas, il contredisait le titulaire. Dans 95 % de ces contradictions, une entrevue subséquente a permis de clarifier le travail.

- .. Dans les deux tiers des dossiers, une entrevue a été menée et des informations supplémentaires ont été fournies. L'enquêteur a noté que des ces renseignements étaient souvent détaillés [...]
- .. La lisibilité des questionnaires était bonne dans tous les cas.

Conclusions :

- .. Le Comité mixte avait mis en place un système pour réviser les questionnaires et veiller à ce qu'ils soient complets.
- .. Il y avait en place un système pour assurer l'exactitude des informations sur les emplois [...] au moyen d'une contre-vérification par le surveillant.
- .. Les personnes chargées de lire les questionnaires ont tâché [...] d'obtenir plus de renseignements pour mieux comprendre [...] lorsque le surveillant et le titulaire ne semblaient pas d'accord sur le travail.

(pièce HR-245)

(ii) Analyse de M. Sunter

629. La Commission a demandé à un ancien directeur de Statistique Canada, M. Alan Sunter, d'examiner l'ensemble des données provenant des études Wisner 222 et Willis 300 et de vérifier les tendances relatives à la

composition des groupes suivant le sexe. La Commission a en outre demandé à M. Sunter d'évaluer la signification statistique des formules relatives à la partialité fondée sur le sexe employées par le Conseil du Trésor dans son document de méthodologie de mars 1990.

630. M. Sunter, un expert en statistique qualifié, n'avait pas de connaissances du domaine de l'équité salariale avant de prendre part à l'analyse des résultats de l'étude sur la parité salariale. Son intervention dans le dossier remonte au 6 avril 1990, soit la date à laquelle M. Durber lui a demandé d'assister à un atelier prévu pour le 9 avril 1990. L'atelier devait porter sur le document de méthodologie du Conseil du Trésor (pièce HR-185). M. Sunter a témoigné qu'il avait été incapable de contribuer de façon constructive à l'atelier et qu'il avait

simplement écouté les discussions. Après l'atelier, il a rencontré M. Durber et a commencé à prendre conscience du fait qu'on avait mené une vaste étude sur la parité salariale entre les groupes professionnels à prédominance masculine et féminine. Il a aussi appris qu'on avait subséquemment réévalué des échantillons tirés des évaluations. On en était ainsi venu à se demander si les évaluations étaient entachées de partialité fondée sur le sexe. Cette question préoccupait la Commission.

631. Les preuves statistiques concernant la partialité fondée sur le sexe dans les résultats des évaluations ont été fournies par MM. Sunter et Shillington, deux experts en statistique. M. Shillington n'a pas été engagé par la Commission pour faire quelque analyse statistique que ce soit des résultats. Cependant, à cause de sa participation aux tests de fiabilité inter-évaluateurs (FIE) et à d'autres aspects de l'étude sur la parité salariale, il a témoigné devant le Tribunal. Durant sa comparution, on lui a demandé de donner son avis sur l'analyse de M. Sunter.

632. M. Durber a demandé spécifiquement à M. Sunter d'effectuer trois analyses. Premièrement, M. Sunter devait examiner la question de la partialité fondée sur le sexe dans les réévaluations; à cette fin, on lui a remis deux ensembles de données, les réévaluations de la Wisner 222 et celles de la Willis 300. Deuxièmement, on lui a donné l'ensemble des données de l'étude sur la parité salariale et demandé d'examiner la question de l'égalité de rémunération pour fonctions équivalentes entre les groupes professionnels à prédominance masculine et ceux à majorité féminine. Troisièmement, on lui a remis le document de méthodologie du Conseil du Trésor (pièce HR-185) et demandé d'examiner spécifiquement les méthodes employées par le Conseil du Trésor et de présenter les critiques qu'il jugerait appropriées.

633. M. Sunter, dans son témoignage, a expliqué dans les termes suivants la façon dont il avait interprété la notion de «partialité fondée sur le sexe» dans son analyse des données (volume 102, à la p. 12275, lignes 3-17) :

[TRADUCTION]

R. J'ai supposé que par partialité fondée sur le sexe on voulait dire qu'il y aurait une tendance systématique de la part des comités d'évaluation à sous-estimer les postes des professions

à prédominance masculine ou à surestimer les postes des professions à prédominance féminine, ou peut-être les deux.

Q. Que voulez-vous dire par «tendance systématique»?

R. A ce moment-là, bien sûr, je l'ignorais, mais puisqu'on avait employé le terme partialité, alors j'ai supposé que la partialité devait se traduire par une tendance qui se refléterait d'une manière décelable dans les données, tendance que je pourrais découvrir en examinant les données et en les analysant d'une façon quelconque.

634. M. Willis a témoigné qu'un consultant formé et expérimenté dans l'application du système d'évaluation et possédant un point de vue objectif s'acquitte normalement de sa tâche d'évaluation de façon uniforme et sans favoriser les emplois à prédominance masculine ou féminine, ou encore l'une ou l'autre des parties, patronale ou syndicale. Selon lui, les évaluations des consultants sont utiles pour examiner l'uniformité des évaluations effectuées par les comités et, ce qui est plus important, pour détecter toute partialité qui a pu se manifester. M. Willis estime que ses consultants -- vu leur expérience, leurs antécédents, leur volonté et leur philosophie -- ont toujours comme ligne de conduite de ne pas favoriser un côté ou l'autre mais de rester neutres. En effectuant les réévaluations, M. Willis visait à déterminer si la façon de traiter les emplois à prédominance masculine ou féminine reflétait un parti pris fondé sur le sexe. Il emploie le terme «disparités» pour désigner les différences entre le consultant et le comité. C'est dans ce contexte que M. Sunter a amorcé l'examen des données.

635. M. Sunter a expliqué qu'un statisticien recueille et analyse des données en établissant une distinction très nette entre les notions de «description» et d'«analyse». A son avis, il est important de bien distinguer ces deux grands champs d'enquête pour comprendre le travail qu'il a accompli et son interprétation des données de l'étude sur la parité salariale.

636. M. Sunter a comparé les deux ensembles de données des réévaluations, la Wisner 222 et la Willis 300, aux évaluations qu'avaient faites les comités des mêmes emplois. Il voulait d'abord savoir si, statistiquement, il y avait une différence systématique dans la manière dont les évaluateurs traitaient divers genres de poste, puis il voulait mesurer ces écarts le cas échéant.

637. M. Sunter a appliqué un test statistique qu'on appelle un «test t» pour vérifier s'il existait une différence entre les consultants et les comités dans le traitement des questionnaires des répondants masculins et féminins. Pour ce test, il a d'abord utilisé la Wisner 222 seule, ensuite la Willis seule, puis enfin la combinaison des deux études.

638. Selon M. Shillington, qui a lui aussi fait des tests t dans son analyse de la FIE, le test t est une épreuve statistique qui résume l'information indiquant de combien deux moyennes s'écartent l'une par rapport à l'autre. En l'occurrence, le statisticien examine la moyenne

164

masculine et la moyenne féminine pour voir s'il y a une indication comme quoi on aurait traité différemment les questionnaires des titulaires masculins et féminins. Il a fait remarquer, dans son témoignage, qu'il y a trois choses qui importent lorsqu'on utilise le test t :

[TRADUCTION]

1. Quelles est l'importance de l'écart entre les deux moyennes? Plus l'écart est grand, plus grande est la probabilité qu'elles proviennent de populations différentes; autrement dit, plus grande est la probabilité qu'un tel ou une telle traite les questionnaires masculins et féminins différemment.
2. Plus la taille de l'échantillon est importante, plus la preuve indiquant que l'on traite les deux populations différemment est significative.
3. Plus la concentration des valeurs est grande, plus on peut affirmer avec certitude que la tendance en question est réelle.

639. Si la différence entre les cotes moyennes est importante, d'expliquer M. Shillington, il est plus probable qu'on obtiendra un résultat «significatif» dans une analyse statistique. Un écart significatif reflète une différence réelle entre deux groupes et démontre que le résultat, en toute probabilité, n'aurait pas pu se produire au hasard. La notion de signification statistique, dans ce contexte, a trait aux probabilités mathématiques et à la question de savoir si les nombres auraient pu se produire ou non par pur hasard (volume 87, à la p. 10673).

640. M. Sunter a témoigné à propos des limitations du test t. Une de ces limitations est le fait que si l'échantillon est très gros, un test t révélera comme significative même une différence minuscule. Autrement dit, le test t rejette l'hypothèse nulle (il n'y a aucune différence entre les variables) lorsque l'échantillon est assez important. Autre limitation : le test t ne tient pas compte des différences ayant une importance pratique; il applique simplement une méthode mathématique pour vérifier l'hypothèse nulle de la non-existence d'une différence par opposition à l'hypothèse contraire selon laquelle il existe une différence.

641. M. Sunter a trouvé qu'il y avait un écart de 2,3 % entre les évaluations que les comités et les consultants avaient faites des postes à prédominance masculine et féminine. Il a effectué d'autres tests t pour déterminer si les consultants et les comités traitaient différemment les postes à majorité féminine. Les résultats ont montré que, pour les postes à prédominance féminine, il n'y avait aucune différence statistiquement significative entre les consultants et les comités quant à la façon dont ils cotaient ces postes. Pour les postes provenant des groupes professionnels essentiellement féminins, les cotes des consultants et des comités ne sont pas significativement différentes, que l'on compare les comités à la Wisner 222, à la Willis 300 ou aux réévaluations combinées des consultants (522). La taille de la différence non significative dans le traitement des postes à prédominance féminine pour les données combinées

165

était de 0,5 %. Dans le cas de la Wisner 222, cet écart était de 0,02 % et, pour la Willis 300, il était de 0,07 % (pièce HR-191).

642. M. Sunter a ensuite appliqué le même test t aux postes à prédominance masculine. Il a constaté que les cotes des consultants et des comités étaient significativement différentes pour les groupes professionnels à prédominance masculine. L'écart entre le comité et le consultant dans le traitement accordé à ces groupes était plus ou moins important selon qu'on utilisait comme base de comparaison avec les résultats des comités les réévaluations de la Wisner 222 ou celles de la Willis 300. Il était aussi fonction du fait qu'on avait ou non placé le comité ou les consultants dans le dénominateur de l'équation. M. Sunter a témoigné que puisqu'il n'y avait pas de «valeur réelle» pour un questionnaire donné, il devait y avoir une norme à laquelle il pouvait comparer les évaluations des comités et des consultants. Si l'on soutient que le comité est biaisé par rapport au consultant, alors on se trouve à prendre le consultant comme point de référence ou norme de comparaison et à placer les cotes de ce dernier dans le dénominateur de l'équation pour déterminer toute différence dans le traitement.

643. La taille de la différence dans le traitement des postes à prédominance masculine pour les réévaluations combinées des consultants (522) était de 1,8 % en utilisant les évaluations des consultants comme dénominateur. Dans le cas de la Wisner 222, cet écart était de 2,5 %; pour la Wisner 300, il était de 1,3 % (pièce HR-191).

644. M. Sunter a donc trouvé une différence significative entre les cotes des consultants et des comités pour les groupes professionnels à

prédominance masculine. Il a témoigné que cet écart était deux fois plus important dans la Wisner 222 (2,5 %) que dans la Willis 300 (1,3 %).

645. M. Sunter préférait utiliser le résultat combiné des études Wisner et Willis (522), qu'il trouvait plus fiable, pour établir la taille de l'écart entre les comités et les consultants, plutôt que d'utiliser de façon indépendante la Wisner 222 ou la Willis 300. Cet écart, a-t-il signalé, était de 2,3 %.

646. Quant à savoir s'il se dégageait une tendance dans les différences constatées entre les comités et les consultants, M. Sunter a constaté que pour plus de la moitié des évaluations, il n'y avait en fait aucun écart entre consultants et comités. Après avoir séparé les données, il a trouvé que pour environ le tiers des comparaisons entre la Wisner 222 et les comités il n'y avait aucune différence; il n'a pas trouvé d'écart non plus entre la Willis 300 et les comités pour environ les deux tiers des comparaisons. A ses yeux il était inconcevable que, à la lumière de ce nombre d'accords, il puisse y avoir une discrimination systématique.

647. M. Sunter a témoigné qu'après avoir trouvé des différences entre le comité et le consultant dans le traitement des questionnaires masculins, il ne conclurait pas que le comité était partial ou que le consultant était partial. A son avis, la seule conclusion qu'on puisse tirer, c'est que le comité et le consultant semblent avoir un parti pris l'un par rapport à l'autre en ce qui concerne les évaluations des postes masculins. On peut

166

parler d'une partialité relative, a-t-il poursuivi, ou qualifier ce parti pris de partialité fondée sur le sexe. Il n'était cependant pas à l'aise d'employer l'expression «partialité fondée sur le sexe» parce qu'en l'absence d'autres tests il était impossible de conclure qui avait un tel parti pris et de savoir si le parti pris n'était qu'accessoirement lié au sexe ou s'il tenait à un autre facteur qui, lui, serait lié accessoirement au sexe.

648. La question cruciale à ce stade du témoignage de M. Sunter était de savoir si les résultats des tests t indiquaient une tendance systématique dans les disparités ou si les différences étaient simplement aléatoires. Selon la Commission, les écarts systématiques liés au sexe doivent, par définition, être le reflet d'un système -- quelque chose de régulier ou de méthodique -- qui est en jeu (par. 199 des arguments écrits).

649. Selon l'employeur, une différence de traitement des questionnaires masculins et féminins est indiquée par une tendance dans les disparités telle que l'évaluation des emplois féminins diffère systématiquement de l'évaluation des emplois masculins (par. 289 des arguments écrits). On peut mieux comprendre l'interprétation que l'employeur fait de la notion de «tendance» à la lumière de l'échange suivant avec M. Sunter (volume 217, de la p. 28243, ligne 8, jusqu'à la p. 28244, ligne 1) :

[TRADUCTION]

Q. Monsieur Sunter, je parle seulement du chi-carré et du test t lorsque vous séparez les questionnaires par sexe. Il y avait une tendance là.

R. Il y avait une différence dans la tendance. Je n'utiliserais pas le terme «tendance». Il y a une différence. Nous avons constaté la différence. Nous cherchons à expliquer la différence.

Q. Mais il y a une différence de traitement, disons.

R. Il y a une différence dans la moyenne -- je n'aime pas le terme «traitement», dois-je dire, parce qu'il implique un processus physique quelconque. Il y a un écart dans les différences entre les cotes des consultants et des comités. Vous pouvez employer le mot «traitement» pour désigner cette réalité si vous le voulez, mais je préfère quant à moi ne pas parler de «traitement».

650. M. Sunter a ensuite tenté d'expliquer et de comprendre les différences entre les comités et les consultants en appliquant aux données des modèles qui, a-t-il dit, étaient nécessaires pour donner un sens à la «notion de partialité fondée sur le sexe». C'est par rapport à cet aspect de son travail que M. Sunter souligne la distinction à faire entre l'utilisation descriptive des statistiques par opposition à leur interprétation à des fins analytiques. Pour les besoins de l'analyse, il a

dû adapter des modèles aux données. S'il y a de la partialité fondée sur le sexe dans les résultats, un statisticien s'attend à trouver dans les évaluations une certaine uniformité qui est d'une façon ou d'une autre reliée au sexe. Par conséquent, il a vérifié s'il y avait une telle

uniformité dans les données en utilisant des modèles pour illustrer comment la partialité fondée sur le sexe avait pu agir.

651. M. Sunter a examiné trois modèles plausibles pour expliquer comment la partialité fondée sur le sexe avait pu influencer sur les résultats des comités. Il a décrit, par exemple, le modèle «additif», qui consiste pour un comité à ajouter constamment quelque chose aux cotes du consultant ou pour le consultant à soustraire constamment quelque chose des cotes du comité. Il a finalement rejeté les trois modèles parce que les données ne se prêtaient pas à de telles manipulations.

652. M. Sunter a aussi étudié les différences entre le comité et le consultant à l'aide du test chi-carré, qu'il a appliqué aux données de la Wisner 222, de la Willis 300 et des deux études combinées. Tous ces tests ont donné des résultats statistiquement significatifs. M. Sunter a critiqué l'utilité de l'analyse de chi-carré dans ces circonstances. A son avis, le test du chi-carré n'est pas utile pour comprendre les différences entre les consultants et les comités dans leur façon de traiter les emplois à prédominance masculine et féminine. Sa réserve au sujet du chi-carré tient au fait que ce test mesure la fréquence plutôt que la taille de l'écart, ce que mesure en revanche le test t. Par conséquent, un résultat significatif obtenu au moyen du chi-carré peut être trompeur quant à la différence réelle entre les nombres. C'est pourquoi il préfère employer le test t, qui a indiqué un écart de 2,3 % par rapport aux données regroupées, comme représentant l'importance de la différence entre les comités et les consultants.

653. N'ayant trouvé aucune différence entre les consultants et les comités, en moyenne, pour les postes à prédominance féminine, M. Sunter a ensuite exploré l'idée que la partialité fondée sur le sexe pût s'exprimer par une discrimination inconsciente en faveur ou contre des groupes professionnels suivant le sexe. Il a postulé que la partialité fondée sur le sexe, dans ce contexte, pourrait se manifester ainsi : il y a des caractéristiques masculines et féminines sous-jacentes et les groupes professionnels comptant plus d'hommes auront tendance à manifester ce genre de discrimination assez fortement. Les tests qu'il a appliqués n'ont révélé aucune corrélation entre la proportion d'hommes faisant partie d'un groupe professionnel et une tendance de différences relatives. M. Sunter a conclu de son analyse qu'il était incapable de trouver quelque tendance systématique que ce soit dans les écarts, et qu'il n'y avait aucune explication plausible ou concluante des écarts existant entre les comités et les consultants.

654. M. Sunter a conclu de son analyse que, en l'absence d'un certain degré d'uniformité ou de régularité dans les différences entre les comités et les consultants suivant le sexe, il était incapable de conclure que

celles-ci étaient attribuables à la partialité fondée sur le sexe. Il a notamment fait les observations suivantes à ce propos (volume 102, de la p. 12277, ligne 25, jusqu'à la p. 12279, ligne 1) :

168

[TRADUCTION]

R. Ma conclusion générale sur la partialité fondée sur le sexe -- mais je vous rappelle que je ne sais toujours pas ce qu'est la partialité fondée sur le sexe --, ma conclusion générale, dis-je, est la suivante. Il y avait une légère différence entre -- il n'y avait pratiquement pas de différence, en fait -- entre les consultants et les comités pour ce qui est des professions où les femmes prédominent. On peut oublier ça.

Pour ce qui est des professions à majorité masculine, il y a une différence réelle, pas très grande mais tout de même statistiquement significative, entre les évaluations des comités et celles des consultants. Cela ne m'amène pas à conclure, toutefois, qu'il y a de la partialité fondée sur le sexe pour l'instant, en oubliant pour le moment que je ne sais toujours pas exactement ce que j'entends par partialité fondée sur le sexe, puisqu'il y a d'autres explications possibles [...]

R. Pour conclure qu'il s'agissait de partialité fondée sur le sexe, il me faudrait trouver une certaine uniformité dans les observations. Je n'ai pu trouver ce genre d'uniformité qui m'aurait permis d'arriver à une telle conclusion.

655. Par ailleurs, M. Sunter était d'avis que le manque d'uniformité dans les différences et l'absence d'un modèle de rechange plausible pour expliquer la partialité fondée sur le sexe ne justifiaient pas le rajustement des cotes des comités suivant la méthode adoptée par le Conseil du Trésor dans son document de méthodologie de 1990.

656. M. Sunter est retourné à la question de la partialité fondée sur le sexe et a exploré d'autres facteurs qui lui sont venus à l'esprit et qu'il n'avait pas examinés dans son enquête initiale. Son exploration a porté sur des facteurs qui pouvaient être associés d'une façon ou d'une autre au sexe et qui, par conséquent, pourraient -- au lieu de la partialité fondée sur le sexe -- expliquer l'écart dans les cotes. Il a examiné d'autres caractéristiques, notamment le salaire perçu, la nature du travail et la taille du groupe, qui, pensait-il, pouvaient avoir un lien de corrélation avec le sexe. En bref, pour employer des termes simples et

ordinaires, M. Sunter a exploré le degré d'association entre les écarts et certaines autres caractéristiques des données.

657. Une caractéristique que M. Sunter a remarqué entre les questionnaires masculins et féminins est le fait que les données indiquaient que les questionnaires féminins provenaient d'un petit nombre de groupes professionnels relativement importants en taille. Les questionnaires masculins, en revanche, provenaient d'un grand nombre de groupes professionnels relativement petits. M. Sunter a postulé que les évaluateurs connaissaient peut-être mieux les emplois à prédominance féminine -- qui incluaient des postes tels que commis, secrétaire et infirmière -- que les emplois où les hommes sont majoritaires, notamment les contrôleurs de la circulation aérienne, les chercheurs scientifiques de la défense, et les examinateurs de brevets. Il a divisé les bases de

169

données selon la taille du groupe, et en utilisant le groupe comme un indicateur de la familiarité avec le genre de travail, il a comparé les différences entre les consultants et les comités pour les données de la Wisner 222 et de la Wisner 300. Bien que les résultats de cette analyse statistique ne fussent pas statistiquement significatifs, M. Sunter croit qu'ils ont démontré une forte association entre la taille du groupe et les écarts constatés entre les comités et les consultants.

658. M. Sunter a remarqué une autre caractéristique qui différenciait les questionnaires masculins et féminins, à savoir la distribution relative des groupes à prédominance masculine et féminine suivant l'échelle des points d'évaluation. Il a constaté que 75 % des questionnaires provenant des groupes à majorité féminine se situaient au-dessous d'une certaine cote, tandis que seulement 25 % des groupes à prédominance masculine étaient en-deça de la même valeur. Il a donc formulé l'hypothèse que toute partialité qui a trait à la distribution des cotes, par exemple un parti pris en faveur du placement dans la hiérarchie des emplois, ou un parti pris en faveur ou contre les postes de gestionnaire ou de surveillant, ou encore un parti pris en faveur des compétences acquises au moyen d'études postsecondaires, pouvait ressembler à de la partialité fondée sur le sexe.

659. M. Sunter a ensuite effectué plusieurs comparaisons pour vérifier si les différences entre les évaluations des comités et les réévaluations des consultants étaient associées à la distribution relative des questionnaires dans la fourchette supérieure ou inférieure de cotation. Ici encore, la comparaison n'a révélé aucune différence statistiquement significative. Il a conclu, toutefois, que les comparaisons montraient qu'il existait effectivement une association entre les cotes élevées et les

cotes faibles, et que les écarts entre les comités et les consultants étaient liés au sexe. M. Sunter a attribué cette partialité à un «effet valeur»; autrement dit, plus la valeur d'un emploi est élevée, plus il est probable qu'il y aura un écart entre la cote attribuée par le comité et la cote donnée par le consultant.

660. M. Willis a répondu au témoignage de M. Sunter sur l'effet valeur durant sa deuxième comparution devant le Tribunal, qui a suivi le témoignage de M. Sunter. Il a dit qu'il aimerait qu'on effectue une autre analyse pour voir si les différences en les comités et les consultants sont associées à l'effet valeur. M. Willis voudrait savoir si en enlevant 10 % des cotes élevées de la base de données l'importance des écarts entre les consultants et les comités serait moindre. A ce propos, il a fait les observations que voici (volume 211, de la p. 27491, ligne 19, jusqu'à la p. 27492, ligne 4) :

[TRADUCTION]

J'ai dit que je ferais appel à un statisticien. On ne m'a pas donné cette tâche, mais si on l'avait fait et si mon statisticien avait dit qu'il y avait apparence de partialité ici et que ce n'était pas forcément de la partialité, je lui dirais : «OK, enlevons les plus hautes cotes et voyons ce que cela donnera.» Peut-être que ce sera moins de 1,8 %, peut-être que non. Comme nous avons affaire à plusieurs millions de dollars, ma suggestion

170

serait que si cela ne change pas ce pourcentage, alors je serais enclin à rajuster.

661. A la suite des commentaires de M. Willis, M. Sunter a fait une autre analyse pour déterminer si les écarts entre les consultants et les comités pouvaient être réduits par l'«effet valeur». Son analyse, qu'il a désignée l'effet valeur, a été produite par la Commission en réponse à la question soulevée par M. Willis. M. Sunter a défini comme suit l'effet valeur (volume 216, de la p. 28049, ligne 23, jusqu'à la p. 28050, ligne 1) :

[TRADUCTION]

R. L'effet valeur serait la tendance systématique qu'ont les écarts entre consultant et comité à se manifester en association avec l'augmentation de la valeur de l'emploi.

662. Dans ses autres travaux statistiques, M. Sunter a voulu savoir dans quelle mesure les écarts entre les évaluations des comités et les réévaluations des consultants étaient attribuables à cette partialité liée à la valeur du poste. Par là il entendait la différence dans la façon dont les comités et les consultants traitaient les questionnaires à cote élevée et à cote faible. Son analyse faisait appel à des méthodes statistiques pour normaliser les données à cause de ce qu'il a décrit comme un problème de distribution. A cause de ce problème, il ne pouvait pas simplement rejeter 10 ou 20 % des cotes les plus élevées comme l'avait proposé M. Willis. L'analyse que M. Sunter a réalisée lui a permis de conclure qu'au moins la moitié des différences apparemment liées au sexe constatées entre les comités et les consultants s'expliquaient immédiatement par des écarts au niveau de la distribution des cotes suivant la valeur de l'emploi.

663. Dans son témoignage, M. Sunter a expliqué que, à partir de l'analyse qu'il avait effectuée (pièce HR-265) et une fois l'«effet valeur» enlevé, l'écart global de 2,3 % entre les consultants et les comités se trouvait ramené à 1,2 %.

664. M. Shillington, pour sa part, doute qu'on puisse séparer -- statistiquement ou autrement -- les deux questions d'analyse de données en cause, à savoir si, d'une part, une tendance est reliée ou non au sexe, et si, d'autre part, la tendance est associée au fait que les cotes soient élevées ou faibles. Il a expliqué ainsi ce problème dans son témoignage (volume 131, de la p. 16045, ligne 23, jusqu'à la p. 16046, ligne 15) :

[TRADUCTION]

Les régressions ont été calculées de façon à vérifier s'il y avait une relation entre les différences entre les consultants et les comités et le facteur sexe.

Il est aussi possible que tout écart qui ait pu exister entre les cotes des consultants et des comités n'ait pas été relié

171

directement au sexe mais plutôt aux valeurs élevées par opposition aux valeurs faibles. Il a été question de cela ici.

Le facteur de confusion entre en jeu parce qu'il y a dans les données une forte tendance pour les questionnaires masculins à avoir tous des cotes élevées par rapport aux questionnaires féminins, et une tendance modérée pour les questionnaires féminins

à provenir de la fourchette inférieure des cotes, ce qui veut dire que vous ne pouvez pas séparer ces deux questions d'analyse des données, ou qu'il est difficile de les séparer à tout le moins.

M. Shillington a poursuivi sa pensée dans les termes suivants (volume 131, de la p. 16048, ligne 16, jusqu'à la p. 16049, ligne 11) :

[TRADUCTION]

Dans cette circonstance, si on se reporte à l'analyse des cotes Willis et de leur rajustement éventuel, on a une situation qui -- dans la mesure où il y a une tendance ici, si quelqu'un arrivait et disait cela n'est peut-être pas relié au sexe, au fait d'être un homme ou une femme, mais pourrait plutôt être dû à la professionnalisation ou au fait que certains questionnaires ont des cotes beaucoup plus élevées que d'autres, vous auriez de la difficulté à extraire ces deux hypothèses distinctes de l'analyse parce que vous avez une situation où les hommes ont majoritairement des cotes élevées, et les femmes ont majoritairement des cotes faibles. Alors le fait d'être un homme se mêle avec des valeurs élevées et faibles.

Cela se reflète dans la distribution. C'est pourquoi il s'agit d'une question de distribution. La distribution des cotes Willis pour les hommes avait tendance à être pas mal plus élevée que la distribution des cotes Willis pour les femmes. C'est une variable confusionnelle. Voilà pourquoi en l'interprétant vous devez faire très attention à ce phénomène.

Et plus loin, sur ce point, il fait remarquer ce qui suit (volume 131, de la p. 16051, ligne 12, jusqu'à la p. 16052, ligne 5) :

[TRADUCTION]

LA PRÉSIDENTE : [...] Mais juste en les regardant que pouvez-vous dire qu'elles décrivent quant à leur distribution, ce que vous pouvez en interpréter, c'est que les hommes ont tendance à avoir des cotes élevées et les femmes, des cotes faibles, mais que vous ne pouvez pas, à cause de cet effet de confusion, vous ne pouvez pas interpréter quoi que ce soit avec certitude. Est-ce cela que...

LE TÉMOIN : C'est exact. Vous devez faire très attention en interprétant les résultats parce que vous ne devez pas oublier que si quelqu'un arrive avec une autre explication pour les données, par exemple, que cela n'a rien à voir avec le fait d'être un homme

ou une femme, mais que c'était plutôt l'effet de la valeur élevée ou faible des cotes, que vous avez recueilli vos données de telle sorte que la plupart des cotes élevées sont celles d'hommes et que la plupart des cotes faibles sont celles de femmes. Ce sont donc deux explications également valables des mêmes données.

665. M. Sunter a reconnu qu'il était difficile d'extraire un facteur confusionnel des données. Il a néanmoins affirmé avoir pu isoler, ou distinguer parmi les disparités, un élément qui était attribuable à des distributions différentes des valeurs selon que les questionnaires visaient des emplois masculins ou féminins. M. Sunter a déclaré qu'il n'avait pas trouvé cela difficile de différencier entre les variables sexe et valeur, et qu'il était à même de «démêler» les données dans cette mesure. Lorsqu'il a été contre-interrogé par l'avocat de l'intimé, il a affirmé qu'il n'était pas prêt à souscrire à l'idée que les deux variables en cause, le sexe et la valeur, pouvaient se substituer l'une pour l'autre. Il a reconnu, cependant, qu'il y avait de nombreux facteurs qui étaient en corrélation avec le sexe, et que, si les écarts entre les comités et les consultants découlaient d'un autre facteur déterminant lui-même associé au sexe, il ne pourrait jamais déterminer quelle partie de l'écart serait attribuable à la partialité fondée sur le sexe (volume 217, à la p. 28247).

666. M. Sunter estime que l'association des écarts dans les cotes à d'autres caractéristiques des données devient une question importante si l'on doit rajuster les résultats des comités afin d'éliminer la partialité fondée sur le sexe. Dans ce contexte, il croit qu'il est important d'établir l'ampleur de cette partialité, de démontrer comment elle s'exerce et d'expliquer comment on peut rajuster en conséquence les données réelles. M. Sunter pense qu'à ce stade-ci l'association des écarts dans les cotes à une partialité liée à la valeur des emplois revêt une importance primordiale.

667. M. Sunter conclut que toute la question de l'association des écarts à d'autres caractéristiques est étroitement liée au processus de rajustement. Par conséquent, il lui paraît difficile de séparer la façon dont on doit analyser les données de ce que l'on compte faire des résultats.

668. M. Sunter connaissait le document de méthodologie du Conseil du Trésor dans lequel ce dernier avait utilisé et rajusté les données de la Wisner 222 pour calculer les rajustements salariaux paritaires en janvier 1990. Aux yeux de M. Sunter, il s'est agi en l'occurrence d'un rajustement

«général». Dans son témoignage, il a expliqué ce qu'il entendait par un rajustement général des cotes (volume 103, à la p. 12426, lignes 16-22) :

[TRADUCTION]

Ce que je ferais si j'étais sur le point de faire un rajustement général, disons, des valeurs attribuées aux questionnaires portant sur les emplois à prédominance masculine, je dirais «Augmentons toutes ces cotes, la totalité d'entre elles, de 4 % sans exception». C'est cela que je veux dire quand je parle d'un rajustement général.

173

669. Selon M. Sunter, un rajustement général exige une certaine uniformité dans la partialité fondée sur le sexe décelée, et un rajustement général ne peut se faire que s'il y a une partialité générale. Il a expliqué ce qu'il voulait dire comme suit (volume 103, à la p. 12427, lignes 8-10) :

[TRADUCTION]

[...] ce sont les deux côtés de la même médaille. Sans l'un, comment pourrais-je être fondé à faire l'autre?

670. Selon M. Sunter, l'employeur a effectué une analyse de régression linéaire, une autre forme de mesure statistique, sur les données de la Wisner 222, tel qu'il l'a décrit dans son document de méthodologie (pièce HR-185). Dans son analyse de régression, l'employeur a évalué les différences de traitement entre les comités et la Wisner 222. C'est à partir de cette analyse qu'il a calculé les corrections qu'il a unilatéralement apportées aux cotes en janvier 1990. M. Sunter, dans la critique qu'il a faite de la méthode employée par le Conseil du Trésor, a analysé les «régions de confiance chevauchantes» des lignes de régression représentant les cotes attribuées aux postes à prédominance masculine et féminine.

671. Selon M. Sunter, le Conseil du Trésor n'aurait absolument pas dû utiliser les régressions pour rajuster les cotes des groupes professionnels essentiellement féminins. En ce qui concerne les données masculines, la ligne de régression comparant les réévaluations de la Wisner 222 et les cotes des comités était significativement différente pour la seconde portion de la gamme des cotes. M. Sunter a trouvé que le chevauchement des régions de confiance des hommes et des femmes, jusqu'à la cote Willis 250, ne constituait pas une forte indication du fait que les consultants et les

comités différaient de façon significative ou uniforme au-dessous de 250 points Willis.

672. M. Sunter a conclu de son analyse des lignes de régression qu'il ne semblait y avoir aucune différence entre les consultants et les comités pour au moins les trois quarts des questionnaires féminins. Par conséquent, il estimait que le Conseil du Trésor n'était pas fondé à se servir des lignes de régression pour apporter des rajustements relatifs à tous les questionnaires masculins et féminins.

673. M. Shillington, lorsqu'il a été contre-interrogé par l'avocat de l'intimé, a affirmé qu'il n'avait rien à redire au sujet de la façon dont M. Sunter avait analysé la méthode de rajustement appliquée par le Conseil du Trésor. Il était d'avis que M. Sunter avait tiré une conclusion raisonnable de son analyse (volume 136, aux pp. 16741-42).

674. Le Tribunal n'a pas entendu de témoignage expert concernant la méthode de rajustement des cotes utilisée par le Conseil du Trésor, exception faite des explications avancées par MM. Sunter et Shillington quant à leur compréhension de la méthodologie présentée dans la pièce HR-185.

174

675. M. Sunter a témoigné que l'analyse de régression n'était pas un outil statistique approprié pour identifier les différences dans les cotes d'évaluation entre M. Wisner et les comités. Les équations de régression, estime-t-il, ne donnent pas raison au Conseil du Trésor d'avoir rajusté les cotes des questionnaires féminins à la baisse, d'en moyenne 3 % dans l'ensemble, et les questionnaires masculins à la hausse, d'environ 4 % dans l'ensemble. Selon M. Sunter, qui s'appuie sur la pièce HR-213, les régressions permettent d'établir des prédictions pour les trois premiers quarts des questionnaires féminins; en l'occurrence cela représente une augmentation de la cote ou ne se traduit par aucun changement du tout.

676. En ce qui concerne les trois questions que M. Sunter devait étudier à la demande de la Commission, ses conclusions sur l'analyse de la partialité fondée sur le sexe relativement aux deux premières questions sont les suivantes : (i) la régularité ou l'uniformité des écarts constatés suivant les sexes était de loin inférieure au niveau qui aurait pu lui permettre de conclure à l'existence d'une partialité fondée sur le sexe. Cela ne veut pas dire qu'il n'y a pas de partialité fondée sur le sexe, d'expliquer M. Sunter, mais seulement qu'on ne peut pas conclure qu'il y en a. Par ailleurs, l'analyse qu'il a faite de la méthode du Conseil du Trésor l'amène à conclure que le Conseil n'avait aucun fondement à partir

duquel justifier quelque rajustement que ce soit des cotes des comités. La troisième question soumise à M. Sunter, qui concerne l'analyse des écarts salariaux entre les groupes professionnels à prédominance masculine et féminine, n'est pas mise en cause à ce stade-ci de notre décision.

#### F. ROLE DES CONSULTANTS DANS LES RÉÉVALUATIONS

677. Les deux experts en statistique ont déclaré en contre-interrogatoire qu'on pouvait utiliser les cotes des consultants comme point de référence pour comparer les cotes des comités à celles des consultants, en tenant pour acquis que les cotes de ces derniers sont exemptes de «partialité reliée au sexe». C'est l'avocat de l'intimé qui a introduit cette expression pour désigner une partialité qui n'est pas reliée au sexe mais qui est reliée à un autre facteur qui, lui, est relié au sexe.

678. Les deux experts en statistique ont exprimé l'opinion qu'ils préféreraient les cotes des comités à celles des consultants. M. Shillington, en particulier, trouvait difficile à accepter qu'un individu quelconque puisse être libre de partialité reliée au sexe, réticence qu'il a exprimée ainsi (volume 139, de la p. 17084, ligne 4, jusqu'à la p. 17085, ligne 2) :

[TRADUCTION]

R. Je pense qu'il s'agit plus là d'une décision que vous pourriez prendre. Vous avez affirmé que c'est de la partialité reliée au sexe qu'il faut se préoccuper, et encore qu'il importe peu que le lien soit direct ou pas. Alors je pense que le fait de décider de ne pas vous préoccuper de la raison pour laquelle il y a de partialité reliée au sexe, s'il est établi qu'elle existe -- cela est une décision.

175

Si on doit interpréter cette phrase comme voulant dire «si vous décidez que la raison ne vous importe pas, alors vous n'êtes pas obligé de la rechercher», vous avez raison. Mais il est certain que pour ma part -- plusieurs fois dans mon témoignage vous m'avez demandé de supposer que M. Wisner était libre de partialité reliée au sexe et à maintes reprises j'ai affirmé «Comment cela est-ce possible? Comment quelqu'un peut-il être libre à ce point de pensées sur les cotes élevées/faibles, le travail salissant/propre? Comment cette personne peut-elle avoir la même

connaissance de tous les emplois?», mais vous m'avez demandé de tenir cela pour acquis.

Donc, je ne suis pas sûr si la phrase telle qu'elle est présentée résume fidèlement ou complètement mon opinion sur le sujet, et je ne peux certainement pas parler pour M. Sunter.

679. Essentiellement, l'employeur soutient que les réévaluations des consultants ne sont utilisées dans l'analyse statistique qu'à titre de point de référence pour déterminer s'il existe, au sein des comités, une tendance à traiter différemment les questionnaires des répondants masculins et féminins. M. Willis a témoigné que les cotes des consultants ne devaient pas être substituées aux cotes des comités. L'employeur fait donc valoir que le fait d'utiliser les réévaluations des consultants comme point de référence ne signifie pas que ces réévaluations doivent être préférées à celles des comités, car il n'y a aucune substitution de cotes. Cependant, l'employeur soutient que, s'il s'agit d'employer les réévaluations des consultants pour déceler une tendance éventuelle à traiter différemment les questionnaires masculins et féminins, le Tribunal pourra préférer le traitement relatif appliqué par les consultants à ces questionnaires, sans pour autant préférer leurs cotes pour aucun questionnaire particulier (arguments écrits de l'intimé, par. 319 et 320).

680. M. Shillington a exprimé l'opinion que, pour utiliser les cotes du consultant comme point de référence, il fallait supposer en partant que les cotes du consultant étaient préférables à celles des comités. Il a donné la réponse suivante à ce propos (volume 136, de la p. 16692, ligne 16, jusqu'à la p. 16693, ligne 15) :

[TRADUCTION]

Q. Lorsque nous utilisons les consultants comme point de référence seulement, nous ne disons pas que, pour un questionnaire donné, nous préférons la cote du consultant à celle du comité. Nous faisons uniquement l'hypothèse que les cotes du consultant, en général, sont exemptes de partialité reliée au sexe.

R. Mais que vous ne les préférez pas?

Q. Mais que nous ne les préférons pas. Donc, nous ne prendrons pas la cote de n'importe quel questionnaire pour dire que les cotes du consultant sont meilleures. Ce n'est pas une hypothèse nécessaire.

R. Mais je pense toujours que vous finissez par supposer qu'elles sont meilleures et l'exemple encore est lorsque j'ai utilisé... supposons que les consultants n'ont même pas regardé les questionnaires et qu'ils n'ont écrit que les températures de jour, la tension artérielle, n'importe quoi. D'accord? On ne pourrait certainement pas parler de préférence, là, et les consultants ne manifesteraient pas de préférence pour un sexe ou pour l'autre s'ils ignoraient totalement les questionnaires. Donc, je pense que vous devez supposer que les cotes des consultants doivent être préférées.

681. D'après M. Sunter, il faut préférer les cotes des comités à celles du consultant pour quatre raisons. Sa première raison est fondée sur sa propre expérience dans le domaine de la statistique qui l'a amené à conclure que les comités appliquent souvent un système mieux que le consultant qui l'a conçu. Ses trois autres raisons pour appuyer les évaluations des comités de préférence à celles du consultant sont basées sur son analyse des données. Dans une de ces analyses, il a comparé les cotes attribuées par M. Wisner à celles du groupe des quatre pour vérifier le degré de compatibilité ou d'uniformité entre les deux.

682. M. Sunter a vérifié l'uniformité entre les cotes de M. Wisner et celles du groupe des quatre à l'aide de tests statistiques comme le test t et l'analyse du chi-carré. Les résultats qu'il a obtenus ont confirmé, à ses yeux, le fait que M. Wisner et le groupe des quatre différaient entre eux. M. Sunter a conclu que si les consultants ne pouvaient s'entendre entre eux il était impossible que le consultant ait toujours raison. Son analyse l'a amené à conclure que les consultants ne cotaient pas de façon uniforme et qu'il fallait donc préférer les comités.

683. Lorsque, en novembre 1984, M. Sunter a été cité en réplique par la Commission, il a témoigné qu'il avait effectué une autre analyse sur la fiabilité relative des comités et des consultants. M. Sunter a aussi utilisé des mesures statistiques standard, en l'occurrence des analyses de régression, pour soutenir l'utilisation des comités comme point de référence dans toute analyse de la partialité fondée sur le sexe. Il a fait des comparaisons de régression linéaire en utilisant deux ensembles de données, les cotes du Comité directeur et toutes les cotes sur lesquelles les comités et les consultants étaient d'accord, ce qui l'a amené à conclure qu'il était impossible de soutenir l'hypothèse de la partialité du comité en faveur des évaluations des postes masculins.

684. M. Shillington a commenté toutes les analyses de M. Sunter à l'exception de l'analyse produite en contre-preuve. Il a souscrit aux conclusions statistiques de M. Sunter, sauf dans le cas d'une analyse, à savoir l'analyse de variance et de covariance. M. Shillington a eu

l'occasion de rencontrer M. Sunter pour discuter de cette analyse. Or M. Shillington a néanmoins continué de maintenir qu'il avait de la difficulté à conclure de l'analyse de variance et de covariance qu'on devait préférer le consultant au comité. Il a formulé l'explication suivante dans son témoignage (volume 133, à la p. 16036, lignes 8-22) :

[TRADUCTION]

177

Alors j'ai de la difficulté à croire que les données peuvent vous aider à démêler ça, que les données peuvent réellement vous aider à décider qu'un évaluateur est préférable à un autre, à moins que vous ayez un troisième ensemble de nombres que vous estimez être les valeurs correctes.

Alors j'examine les modèles et je me dis que les modèles ont l'air raisonnables et, oui, il est clair que la matrice des corrélations dans un cas se rapproche davantage des données observées que la matrice des corrélations dans l'autre cas, mais même après avoir discuté de cela, je dois m'arrêter et poser la question : Cela est peut-être vrai, mais comment les données peuvent-elles vous aider à décider quel évaluateur est le meilleur si vous n'avez pas de troisième ensemble de nombres représentant les valeurs correctes?

685. M. Shillington a poursuivi en disant que son opinion sur cet aspect du témoignage de M. Sunter ne changeait rien au fait qu'il approuvait l'analyse que ce dernier avait faite de la partialité fondée sur le sexe. Il a répondu ce qui suit à ce propos (volume 133, de la p. 16036, ligne 23, jusqu'à la p. 16037, ligne 23) :

[TRADUCTION]

Q. Après avoir discuté de cette question avec M. Sunter et décidé de maintenir votre opinion, quel effet cela a-t-il sur votre opinion concernant ses approches qui ont été résumées dans le document HR-184 portant sur la question de la partialité fondée sur le sexe?

R. Il s'agissait là d'une partie de la preuve de M. Sunter, des arguments pour lesquels il ne préférerait pas les consultants aux comités, mais il y a d'autres éléments à cet argument. Je n'ai rien à redire sur les autres éléments que j'ai vus et j'ai déjà à ce sujet -- j'ai témoigné sur l'autre partie, alors je n'ai pas de problème concernant ces éléments-là de la preuve.

En général, en dépit du fait que je ne suis pas de son avis concernant cette partie de son témoignage, je n'ai rien à redire sur la façon dont il a traité les comités par opposition aux consultants, bien que je ne souscrive pas à cette étape précise de son argumentation.

Q. Je ne voulais pas simplement parler de la question des comités par opposition aux consultants, je voulais parler de toute la question de la partialité fondée sur le sexe, de tous les autres tests présentés dans le document HR-156.

R. Je vous répète que je n'ai pas de réserves au sujet de ces analyses, non.

686. On a demandé à M. Shillington de commenter l'inclination de M. Sunter à préférer les comités aux consultants, non pas pour des raisons statistiques, mais plutôt parce qu'une décision prise par un groupe

178

d'individus, aux yeux de M. Sunter, était préférable à une décision prise par quelqu'un qui a pu avoir reçu une formation technique plus poussée. M. Shillington était du même avis que M. Sunter, et il a affirmé qu'il préférerait lui aussi le consensus de sept personnes atteint d'une façon équilibrée à l'avis d'un seul expert technique, tout au moins lorsqu'il s'agit d'une question telle que la parité salariale.

687. Les deux statisticiens, MM. Sunter et Shillington, étaient d'accord et ont informé le Tribunal que s'il prévoyait utiliser les résultats des tests t de M. Sunter pour rajuster les cotes d'évaluation des comités, alors les cotes des consultants ne servaient plus uniquement de point de référence mais se trouvaient, en fait, à être retenues de préférence aux cotes des comités. Dans ce contexte, les statisticiens estiment qu'on doit pouvoir tenir pour acquis que les cotes des consultants sont exemptes de partialité fondée sur le sexe ou reliée au sexe, avant de rajuster d'une façon ou d'une autre les cotes des comités.

688. M. Shillington a témoigné qu'il fondait son opinion non pas sur la statistique, mais sur le raisonnement scientifique et la logique. Il a expliqué ses motifs comme suit (volume 136, de la p. 16706, ligne 14, jusqu'à la p. 16707, ligne 6) :

[TRADUCTION]

R. Je vous laisse à vous le soin de débattre de la question de savoir si c'est statistique ou pas. Ce qu'il faut savoir ici c'est si le fait d'utiliser comme point de référence pour évaluer la préférence reliée au sexe quelqu'un qui est cohérent et impartial signifie que vous préférez ces cotes. C'est là le noeud de la question je pense.

Q. C'est exact.

R. J'ai des problèmes à ce sujet parce que je ne comprends pas la logique sous-jacente. Je dis pour ma part que c'est un raisonnement scientifique. Quant à moi, c'est de la logique.

Q. Est-ce que je peux le formuler comme suit : Ce qui fait problème à vos yeux, c'est que vous ne voyez pas comment quelqu'un peut appliquer un plan de façon cohérente, uniforme, sans partialité reliée au sexe, et ne pas en même temps être préféré. Est-ce que c'est ça?

R. Oui.

689. M. Willis a défendu l'impartialité et l'objectivité de ses consultants, et a témoigné que les réévaluations des consultants pouvaient servir de point de référence pour déterminer s'il existait, parmi les comités, une tendance à traiter différemment les questionnaires masculins et féminins. Il fondait son opinion sur la conviction que ses consultants avaient toujours eu pour principe de ne pas favoriser un côté sur l'autre, qu'ils étaient plus expérimentés en évaluation de postes et qu'ils pouvaient s'acquitter de leur tâche de façon uniforme et impartiale.

179

Enfin, les consultants avaient plus d'expérience lorsqu'il s'agissait d'interpréter des questionnaires difficiles.

690. M. Fred Owen, un expert en parité salariale et un ancien consultant de Willis & Associates qui a participé à l'étude sur la parité salariale, a témoigné qu'il considérait comme très important pour déterminer la fiabilité des évaluations effectuées dans le cadre de cette étude que les consultants fournissent un cadre de référence permettant d'établir l'exactitude des évaluations. Il estimait que les évaluations des consultants pouvaient servir de norme de comparaison pour plusieurs raisons. Premièrement, non seulement les consultants ont énormément de connaissances et d'expérience par rapport au plan d'évaluation, mais ils ont été exposés à des évaluations dans un large éventail d'emplois.

Deuxièmement, a-t-il fait valoir, les consultants avaient accès à une gamme complète de postes en évaluation, alors que les comités n'avaient accès qu'à un groupe plus restreint d'emplois. Troisièmement, les consultants ne savaient rien du système de classification de l'employeur ou des échelles de rémunération des classes d'emplois, et ils n'avaient pas d'idées préconçues concernant le système de rémunération. M. Owen a témoigné que les consultants eux-mêmes consultaient presque quotidiennement des contrôles de qualité, non seulement pour déterminer l'uniformité avec laquelle on appliquait la discipline du Comité directeur, mais également afin de vérifier les évaluations faites par les consultants eux-mêmes en vue de savoir si ces derniers étaient corrects dans leur évaluation.

691. M. Owen, dans un document écrit (pièce R-167) dont il a confirmé la teneur lorsqu'il a témoigné devant le Tribunal, a énoncé des critères pour l'adoption des évaluations des comités. A son avis, si les comités manifestaient une bonne compréhension du plan d'évaluation -- ce que montrerait le caractère raisonnable de leurs évaluations -- et s'il n'y avait aucune tentative observable de la part de membres de manipuler les résultats de l'évaluation ou de favoriser de manière préjudiciable une profession ou un titulaire en particulier, il ne serait nullement nécessaire d'évaluer les évaluations des comités en fonction des réévaluations des consultants. Selon M. Owen, ces critères n'ont pas été respectés à cause notamment du fait qu'on ne disposait pas d'informations complètes sur les emplois et du fait que certains membres des comités -- on l'a observé -- manipulaient les évaluations de façon à surestimer les emplois à prédominance féminine et à sous-estimer les emplois traditionnels occupés majoritairement par des hommes.

692. Il y a amplement de preuves qui montrent que le Comité mixte, pendant que l'étude était en cours, était disposé à utiliser les consultants comme norme. Le Comité mixte avait convenu d'utiliser les cotes des consultants comme point de référence durant les tests de fiabilité inter-comités (FIC). Dans ce contexte, les consultants ont évalué les questionnaires qui avaient été remis aux comités, et les cotes des consultants ont servi de point de référence pour les tests FIC.

693. Tout au long de l'étude, M. Willis a recouru aux consultants, avec l'assentiment du Comité mixte, à titre de norme pour valider le travail des comités. Dans une lettre datée du 6 janvier 1989 qu'ils ont envoyée à M. Willis, les coprésidents du Comité mixte ont demandé à ce

dernier de fournir des cotes de référence pour les questionnaires devant faire l'objet des tests FIC. Voici un extrait de cette lettre :

[TRADUCTION]

[...] Votre retard à fournir des cotes de référence retarde le travail du sous-comité des tests de fiabilité inter-comités (FIC), qui a besoin de cette information pour analyser l'uniformité des cotes des comités en fonction d'une norme.

(pièce HR-82)

694. Il y a eu d'autres occasions, durant l'étude sur la parité salariale, où le côté patronal et le côté syndical ont conjointement et séparément demandé aux consultants de M. Willis de réviser les évaluations des comités. Bien que cela n'ait pas eu lieu dans le même contexte que les tests FIC, dans lesquels les cotes des consultants ont servi de point de référence pour fin de comparaison avec les cotes des comités, l'opinion des consultants, néanmoins, a été recherchée pour contrôler la qualité des cotes des comités. Les études menées par les consultants relativement aux évaluations des postes-repères du Comité directeur ont été décrites ailleurs dans cette décision. Il reste à aborder l'acceptation, par le Comité mixte, de mandater M. Willis pour qu'il confie à son consultant Wisner les 222 réévaluations des évaluations réalisées par les comités. Il y a en outre les études moins formelles que les consultants ont faites durant les travaux des deux générations de comités (les cinq premiers et les neuf subséquents) pour vérifier l'uniformité de leurs évaluations.

695. Les extraits suivants du témoignage de M. Willis font état d'autres exemples d'évaluations que les consultants ont faites des questionnaires cotés par les comités dans le but de valider les résultats (volume 60, à la p. 7435, lignes 3-23) :

[TRADUCTION]

Q. Pendant que le Comité directeur effectuait ses évaluations indépendantes, est-ce que vous révisiez vous aussi les questionnaires qu'ils examinaient?

R. Oui.

Q. Dans quel but?

R. Une partie du travail visait, en fait, à valider l'uniformité ou la cohérence de leurs évaluations. Mon rôle, en général, consistait à examiner les questionnaires avec le comité, à écouter leurs discussions et à faire ma propre évaluation de l'emploi à partir des informations qui étaient avancées. Puis, je faisais le suivi de tout ça.

Je ne communiquais pas mon évaluation au comité, mais je comparais le consensus à mon évaluation pour m'assurer que le comité interprétait toujours de façon uniforme les informations figurant dans les questionnaires et qu'il appliquait aussi de façon uniforme le système d'évaluation proprement dit.

181

Plus loin, il ajoute ce qui suit (volume 67, à la p. 8429, lignes 2-10) :

[TRADUCTION]

R. J'ai répondu à nombre de préoccupations exprimées et ré-exprimées par le Conseil du Trésor à partir de l'été de 1989 -- l'été de 1988. J'avais eu l'impression de régler une fois pour toutes la question de savoir si le Comité directeur évaluait de façon juste et équitable. C'est moi, en fait, qui a validé les résultats. J'ai affirmé qu'il avait accompli un travail estimable et crédible, et pourtant les problèmes refaisaient toujours surface.

696. La réaction du Comité mixte, durant l'étude, à la demande de M. Willis de faire faire la Wisner 222 au cours de l'étude n'a pas, à l'époque, remis en question l'impartialité de M. Wisner. Il est raisonnable de conclure que les parties elles-mêmes, à ce moment-là, tenaient pour acquis que les consultants remplissaient leur mandat d'une manière impartiale.

697. Les parties ont compris de M. Willis qu'il n'y avait pas de cote correcte pour aucun questionnaire. Avec le temps, la seule initiative que les parties et le consultant ont prise pour contrer l'influence d'une partialité éventuelle fondée sur le sexe fut de mettre en oeuvre des mesures pour améliorer le processus. Ces mesures d'appoint ou de protection ont été décrites antérieurement. C'est dans ce contexte que M. Willis a été appelé à conseiller les évaluateurs et à leur donner -- ainsi qu'aux comités -- de la formation supplémentaire.

698. M. Willis a témoigné qu'il y avait une grande différence entre le fait d'invoquer les réévaluations des consultants une fois le processus terminé et le fait de les utiliser pendant que le processus était en cours. Dans le premier cas, a-t-il expliqué, les réévaluations servent à déterminer si, dans les différences constatées, une partialité fondée sur le sexe est en jeu. Dans le second cas, M. Willis estime qu'il n'est pas particulièrement important de connaître les raisons des disparités entre

les consultants et les comités parce que seule l'existence d'une tendance est importante à ses yeux.

699. M. Willis croit fermement que lui et ses consultants ne manifestent aucune tendance dans leurs évaluations. Il a précisé ainsi sa pensée (volume 210, à la p. 27323, lignes 9-12) :

[TRADUCTION]

R. Je suis persuadé que les consultants expérimentés de Willis & Associates sont en général impartiaux ou aussi impartiaux qu'il est humainement possible de l'être.

700. Il a poursuivi en expliquant que par «impartiaux» il voulait dire qu'ils ne faisaient aucune distinction fondée sur le sexe entre les répondants masculins et féminins. On a demandé à M. Willis s'il croyait que ses consultants pouvaient avoir un parti pris relié au sexe, par exemple, en étant plus libéral dans la fourchette supérieure des cotes ou

182

plus conservateur dans la fourchette inférieure. Il a répondu dans les termes suivants (volume 210, de la p. 27323, ligne 23, jusqu'à la p. 27325, ligne 22) :

[TRADUCTION]

R. C'est un point intéressant.

Q. C'est un peu plus difficile à trancher, n'est-ce pas?

R. Eh bien, il y a certaines études que nous avons faites qui montrent qu'il est difficile de bien comprendre un emploi qui se situe deux ou trois niveaux au-dessus de votre propre poste.

M. Alan Sunter a fait remarquer que ce qui paraît d'abord comme de la partialité fondée sur le sexe pourrait être autre chose.

Q. Oui, c'est un bon point. Je vais vous parler beaucoup de ce point, alors il n'est pas nécessaire que nous -- ramenez-moi sur le sujet plus tard si je ne l'ai pas traité en détail. Vous dites qu'il y a des études qui montrent qu'il est difficile de bien saisir les emplois trois ou quatre niveaux au-dessus du vôtre, mais vos consultants étaient en général des gens qui occupaient de très hautes fonctions avant de se joindre à votre cabinet, n'est-ce pas?

R. Et ce sont des consultants expérimentés dans l'évaluation de postes de haut niveau. Un des problèmes qui s'ajoute au fait qu'il est difficile pour un membre de comité d'évaluer un emploi classifié plusieurs niveaux au-dessus du sien -- c'est-à-dire de bien comprendre les principes et théories en cause, l'importance de l'emploi, ce que veut dire la planification stratégique etc., autant de choses qui lui sont plutôt étrangères -- une autre chose que nous constatons en même temps, dis-je, c'est que les emplois complexes sont plus difficiles à décrire.

Alors il ne faut pas s'étonner -- je crois que c'est Alan Sunter qui a laissé entendre que peut-être les consultants avaient tendance à évaluer de façon plus libérale que les comités les postes de niveau élevé.

Q. Et cela irait dans le sens de ce que vous soutenez. Je suppose que ce que vous voulez dire, c'est que cela concorde avec l'expérience que vous avez acquise en observant des consultants et des comités évaluer des emplois?

R. Je dirais que ce phénomène n'est pas forcément exceptionnel.

Q. Bon, disons que les consultants sont donc mieux placés pour apprécier ces emplois. Je dirais que les autres facteurs en jeu par rapport aux emplois de haut niveau, je pense me rappeler vous avoir entendu dire que les gens ont tendance à surestimer leur propre emploi par comparaison aux emplois qu'ils connaissent moins bien. C'est vrai?

183

R. Je pense que nous avons un certain préjugé en ce sens.

701. Si M. Willis est d'avis qu'on ne peut pas examiner les raisons qui expliquent les écarts entre les comités et les consultants, c'est parce qu'il croit qu'il serait très difficile de choisir des évaluations individuelles dans le but d'expliquer la différence.

702. Cependant, il y a eu des occasions pendant l'étude où M. Willis a examiné les évaluations des consultants (Mme Drury et M. Wisner) en vue de comprendre les écarts entre les consultants et le Comité directeur. M. Willis a réalisé ce genre d'analyse avec 46 postes-repères du Comité directeur pour lesquels il y avait un écart de plus de 10 % entre M. Wisner et le Comité directeur. Dans cette analyse, il recherchait une tendance. Il voulait savoir s'il décèlerait dans les résultats une partialité -- ou

l'apparence d'une partialité -- fondée sur le sexe. Pour ce faire, il a étudié les écarts et les motifs de ces écarts tels qu'ils étaient formulés dans les justifications raisonnées des comités.

703. En contre-interrogatoire, M. Willis a reconnu que la difficulté à laquelle on se heurte en comparant les écarts entre les consultants et les comités réside dans la détermination du poids à attribuer à un facteur donné, car rien ne garantit qu'un seul facteur est en cause (volume 210, à la p. 27350).

704. Pour ce qui est des écarts entre les comités et les consultants dans le traitement des emplois de haut niveau, M. Willis a témoigné qu'il était prêt à accepter le fait que les consultants étaient probablement plus libéraux dans leur évaluation des postes de niveau élevé. D'après son expérience, les consultants comprenaient probablement mieux que les comités les emplois de niveau élevé. Il a par ailleurs tenu les propos suivants sur ce point (volume 210, de la p. 27355, ligne 18, jusqu'à la p. 27356, ligne 23) :

[TRADUCTION]

Q. Vous nous avez dit pourquoi ils pourraient avoir une meilleure compréhension de ces postes, mais vous conviendrez avec moi qu'il est possible que dans les situations où il y a moins de postes-repères -- c'est exact?

R. Oui.

Q. Et vous devez faire davantage preuve de jugement. C'est exact?

R. Oui.

Q. -- que le consultant peut être influencé dans sa façon de percevoir ces emplois par l'expérience qu'il a des postes de haut niveau à l'extérieur de la fonction publique fédérale.

R. Et par d'autres études qu'il a pu réaliser. Oui, c'est possible.

184

Q. Donc vous pouvez voir qu'il y a des choses qui pourraient les placer dans une meilleure position pour avoir une vue «préférable» de ces emplois de haut niveau. Exact?

R. Je pense qu'il n'y a aucun doute là-dessus.

Q. Vous nous avez dit qu'ils pouvaient être influencés entre autres par des choses extérieures, par ce qu'ils ont appris d'études extérieures.

R. Je dirais que lorsque nous parlons d'emplois complexes et de haut niveau, les consultants devraient pouvoir mieux en saisir le contenu que n'importe qui d'autre au sein des comités d'évaluation, qui n'ont pas forcément des gens de cette expérience dans leurs rangs.

#### G. OPPORTUNITÉ DE RAJUSTER LES RÉSULTATS - LES EXPERTS

705. Selon M. Willis, pour trancher la question de la fiabilité des résultats, le Tribunal a trois options : (i) appliquer l'étude telle quelle; (ii) rajuster les résultats; (iii) rejeter l'étude.

706. En ce qui concerne l'option (i), M. Willis a déclaré que sans analyse statistique et l'avis d'un statisticien il n'accepterait pas les résultats. Dans son témoignage, il a souligné ce qui suit (volume 78, de la p. 9576, ligne 19, jusqu'à la p. 9577, ligne 8) :

[TRADUCTION]

C'est vrai que je n'étais pas heureux des diverses mesures qui ont été prises, mais nous avons pu jusqu'à un certain point appliquer des mesures de soutien. Toutefois, sans aucune analyse, sans la possibilité de faire, ou de faire faire, des analyses statistiques ou d'obtenir l'avis d'un statisticien, je ne pense pas que j'aurais pu accepter les résultats.

Une fois l'étude terminée, il est possible d'examiner les résultats sans égard aux autres questions et de prendre une décision séparée : Avons-nous un résultat uniforme, ou avons-nous une certaine partialité et quelle en est l'ampleur? Dans un sens, votre perspective change une fois que l'étude est finie.

707. A propos de la troisième option, M. Willis a fait les observations suivantes (volume 78, de la p. 9574, ligne 15, jusqu'à la p. 9575, ligne 7) :

[TRADUCTION]

LE PRÉSIDENT : Pouvez-vous nous dire quand il y aurait lieu de retenir la troisième option?

LE TÉMOIN : Je voudrais m'asseoir et parler avec Milczarek et examiner tous les détails avec lui. Mais il est possible, je

185

suppose, que les résultats puissent être bizarres au point d'être incroyables. A ce moment-là, il faut les rejeter.

Si nous avons arrêté après les 222 évaluations, si rien n'était arrivé après cela, et si des décideurs m'avaient demandé quoi faire avec, sans possibilité d'analyser les résultats, à ce moment-là j'aurais dit qu'il n'y avait rien à faire avec le résultat. Nous ne pouvons pas utiliser ce que nous avons jusqu'ici pour obtenir des résultats valides. Le test sur les 220 postes était en soi trop restreint pour qu'on puisse porter un jugement. Donc, s'il n'est pas possible de faire quoi que ce soit de plus, il faut oublier l'étude.

708. Dans sa dernière comparution devant le Tribunal, en juin 1994, M. Willis a réitéré son avis à savoir qu'il n'y avait pas lieu de rejeter l'étude. L'étude sur la parité salariale, a-t-il fait remarquer, portait sur l'équité dans le traitement des employés. Or les différences entre les consultants et les comités mises en lumière par l'analyse de M. Sunter étaient si faibles, si on se reporte au salaire d'un seul employé, qu'«après avoir déduit l'impôt, il reste à peine de quoi se payer un café» (volume 211, à la p. 27489). En revanche, d'ajouter M. Willis, «ce sont des millions de dollars qui sont en cause, alors il y a peut-être plus que l'équité envers l'employé qui est en jeu» (volume 211, à la p. 27489).

709. Après avoir rencontré M. Sunter, M. Willis était intéressé de savoir dans quelle mesure la différence entre les comités et les consultants était véritablement attribuable à ce qu'il appelait l'«effet valeur». Selon M. Willis, si cette partialité avait pour effet de réduire l'écart entre consultants et comités au point où il devenait négligeable, alors il n'était guère nécessaire de rajuster les évaluations des comités. Par contre, si l'analyse supplémentaire n'avait pas pour effet de réduire cet écart, alors il «procéderait à des rajustements», vu les sommes d'argent en cause (volume 211, à la p. 27492).

710. Bien que personne n'ait témoigné pour l'employeur relativement au document de méthodologie du Conseil du Trésor (pièce HR-185), la preuve démontre que le Conseil a rajusté les cotes d'évaluation en prenant comme base les réévaluations de l'étude Wisner 222. Les corrections apportées ont précédé les rajustements paritaires que l'employeur a versés en janvier 1990. L'employeur a rajusté toutes les cotes, autres que celles des

postes-repères, pour lesquelles il existait une réévaluation effectuée par le consultant. Les questionnaires ont été rajustés en appliquant deux équations de régression contenues dans la pièce HR-185, à la p. 11, renvoi 7. On a demandé à M. Shillington ce qu'il pensait de ces équations de régression; voici sa réponse (volume 134, à la p. 16401, lignes 13-25) :

[TRADUCTION]

LE TÉMOIN : Je ne rajusterai pas. Je peux vous dire que lorsque j'ai vu ces équations pour la première fois et que je savais encore peu de choses sur les antécédents des données, je me suis tout de suite dit, et j'ai exprimé cette opinion à maintes reprises, que le fardeau repose sur la personne -- avant de

186

rajuster, je pense qu'il incombe à l'enquêteur de montrer que le rajustement est justifié, et la preuve ici est que le rajustement n'était pas justifié, pourtant on l'a quand même fait. Je suis arrivé à cette conclusion en tant que statisticien sans savoir grand-chose sur l'historique de l'étude et rien de ce que j'ai appris depuis ne m'a incité à changer d'avis.

711. M. Sunter a exprimé la même opinion que M. Shillington au sujet des rajustements du Conseil du Trésor. Voici la teneur de ses propos (volume 106, de la p. 12745, ligne 21, jusqu'à la p. 12747, ligne 10) :

[TRADUCTION]

Là où je veux en venir, c'est que les équations de régression que présente le Conseil du Trésor dans son document n'approchent même pas du niveau de certitude que je considérerais nécessaire pour rajuster de quelque manière que ce soit les évaluations des postes masculins et féminins.

LE PRÉSIDENT : Pouvez-vous expliquer cela un peu plus.

LE TÉMOIN : Parce qu'elles ne sont pas significativement différentes. Pour appliquer un rajustement général sur la base du sexe, il faudrait que je sois sûr d'un certain nombre de choses.

Premièrement, il faudrait que je sois certain que le consultant est préférable au comité, et je n'ai aucune certitude à ce sujet. Comme j'ai tenté de le montrer hier, il y a de bonnes raisons d'être sceptique sur ce point.

Deuxièmement, il faudrait que je sois sûr que la raison de l'écart est le sexe des titulaires, et non pas quelque chose d'autre qui est simplement relié au sexe d'une façon ou d'une autre.

Enfin, il faudrait que je sois sûr des nombres que j'utilise si je voulais faire un rajustement.

Nous avons vu que l'importance de l'écart entre les consultants et le comité, selon l'équation que vous utilisez et l'ensemble de données que vous retenez, est de l'ordre d'environ 2 à 2,5 %. Néanmoins, nous avons ici une méthode qui arrive à un rajustement de 7 %. Comment est-ce possible?

La réponse est que l'analyse de régression est un instrument très pauvre, rudimentaire pour estimer l'écart. Même si je devais croire toutes les autres choses, la régression demeure un très mauvais outil pour effectuer ce rajustement à cause de l'incertitude inhérente aux analyses de régression elles-mêmes.

712. M. Durber, au nom de la Commission, a appuyé comme tels les résultats de l'étude. Il a formulé ainsi sa conclusion au Tribunal (volume 154, à la p. 19167, lignes 4-24) :

187

[TRADUCTION]

R. Ma conclusion est que les parties ont très bien réussi à produire une masse d'informations excellentes sur les emplois. Elles ont consacré des coûts et des efforts énormes pour produire des résultats d'évaluation. Elles ont testé ces résultats, nous l'avons vu, de façon exhaustive; du moins elles ont été exhaustives et je ne suis pas sûr des résultats sur notre compte.

Je suis convaincu que les études que nous avons examinées n'atteignent pas la qualité de travail dont nous avons été témoins dans cette étude-ci. Je pense que les parties ont beaucoup de mérite d'avoir produit cette étude et j'ai pour ma part fait entièrement confiance à ces résultats pour recommander à la Commission de les utiliser pour examiner les preuves d'un écart salarial.

Je ne pense pas qu'il existe ce que je qualifierais de preuve de partialité. Au bout du compte, j'estime qu'on devrait accepter tels quels les résultats de l'évaluation des emplois et les

utiliser en toute confiance dans tout calcul des disparités salariales.

713. M. Sunter a maintenu tout au long de son témoignage qu'il ne fallait pas rajuster les évaluations des comités. Toutefois, en réponse à des questions soulevées par M. Willis et à la demande de la Commission, il a proposé quelques méthodes pouvant servir à rajuster les résultats des évaluations. Nous traiterons de ces méthodes plus en profondeur si nous concluons qu'il y a effectivement lieu de rajuster les évaluations.

## VII. DÉCISION ET ANALYSE

714. Tout au long de l'étude sur la parité salariale menée par le Comité mixte, le Conseil du Trésor et l'Alliance ont invoqué le témoignage expert de M. Willis pour faire valoir leurs positions. Cependant, durant l'audience et dans les argumentations écrites et verbales, il y a eu énormément de controverse entre le Conseil du Trésor et les autres parties à propos du rôle joué par ce consultant dans la réévaluation des questionnaires et de la question de savoir si on pouvait se fier à lui pour produire des évaluations sans partialité fondée sur le sexe.

715. Le Tribunal trouve la position de la Commission et de l'Alliance particulièrement difficile à comprendre. L'impartialité de M. Willis n'était pas une question litigieuse avant l'audience. Dans ses arguments, toutefois, l'Alliance a cité un certain nombre de raisons pour lesquelles on devait préférer les évaluations des comités à celles des consultants. Le «bagage» des consultants, a-t-elle soutenu, ainsi que d'autres facteurs tels que l'âge, le sexe, l'éducation et le manque de sensibilisation aux questions de partialité fondée sur le sexe contribueraient à un parti contre l'autre sexe.

188

716. A notre avis, l'Alliance cherchait à discréditer le témoin sur l'opinion experte duquel elle s'appuyait pour faire valoir le processus de collecte des données et d'évaluation des emplois suivi pendant l'étude. Pour mieux illustrer ce fait, nous nous reportons à l'échange suivant, entre l'avocat de l'Alliance et le Tribunal, qui figure au volume 224, de la p. 29495, ligne 17, à la p. 29500, ligne 19 :

[TRADUCTION]

LA PRÉSIDENTE : Avant que vous poursuiviez, Me Raven, j'aimerais réagir au mot «antagonisme» que vous employez à l'endroit du Tribunal. C'est un mot plutôt... c'est un mot qui n'est pas sans

connotation. Je pense que ce que le Tribunal cherche à faire, c'est de comprendre votre argument.

Les parties ont retenu ou engagé ces consultants pour les aider à mener une étude sur une période de cinq ans. Face à un argument selon lequel ces consultants pourraient avoir un parti pris contre l'autre sexe, je pense que le Tribunal a voulu chercher à comprendre la nature de vos arguments et à vous amener à les justifier. Je ne pense pas qu'il soit juste d'affirmer que nous ayons, ce faisant, voulu antagoniser qui que ce soit ou que nous ayons nous-mêmes été antagonisés. J'estime que c'est notre rôle d'agir ainsi et nous continuerons de le faire en vue de comprendre et de juger les arguments que vous tentez de faire valoir.

Me RAVEN : Je comprends cela. Je cherchais plutôt à préciser davantage mes arguments dans ce but.

MEMBRE FETTERLY : Avant que vous commenciez, j'aimerais faire quelques commentaires sur ce point.

Pour commencer, si c'était un procès civil et si M. Willis était votre témoin -- techniquement parlant, ce n'est pas le cas; il est le témoin de la Commission -- auriez-vous le droit de le discréditer après l'avoir cité comme témoin?

Me RAVEN : Monsieur Fetterly, je répondrai ainsi à votre question. Je ne cherche pas à discréditer M. Willis. C'est peut-être là que nous...

MEMBRE FETTERLY : C'est clairement l'impression que vous donnez, Me Raven. Permettez-moi d'ajouter ceci : M. Willis et ses collègues consultants ont été les seuls experts à participer effectivement aux travaux du Comité mixte. Vous vous appuyez sur les résultats du Comité directeur et il a défendu ces résultats non seulement devant ce Tribunal, mais également devant le Comité mixte. Il a défendu les résultats du sous-comité de fiabilité inter-comités. Et il l'a fait tant devant ce Tribunal que devant le Comité mixte. Il a défendu l'ensemble des résultats devant ce Tribunal. Essentiellement, il a dit qu'il ne fallait pas les rejeter. Et c'est son plan qu'on a adopté comme étant exempt de partialité fondée sur le sexe.

Le fait de vous entendre vous et, dans une certaine mesure, Me MacLean attaquer, dans un sens, sa neutralité place le Tribunal dans une

situation très inconfortable. Cela me préoccupe beaucoup. Ce n'est pas une question d'antagonisme.

Me RAVEN : J'espérais ce matin pouvoir clarifier où nous en étions à ce sujet. Votre commentaire, Monsieur Fetterly, arrive à point. Il nous permet, et il me permet à moi, de l'aborder concrètement.

Il ne s'agit nullement ici d'une attaque contre M. Willis comme quoi lui ou n'importe qui d'autre associé à son cabinet auraient été coupables de quelque faute de commission ou de conduite que ce soit au cours de l'étude ou dans la façon dont ils ont...

MEMBRE FETTERLY : Pas du tout. Pas du tout. M. Willis et ses associés se considèrent comme des experts en équité salariale. Ils font la promotion de leur plan comme étant libre de discrimination fondée sur le sexe. Ils forment des évaluateurs pour qu'ils puissent évaluer sans parti pris contre l'autre sexe. Maintenant vous affirmez que leur propre capacité d'évaluer sans un tel parti pris est douteuse. A mes yeux, il y a là une véritable contradiction.

Me RAVEN : Le fait que les parties ici présentes aient accepté le plan Willis comme étant libre de partialité fondée sur le sexe pour les besoins de l'étude est une chose. Mais je puis vous assurer, Monsieur Fetterly, que nous n'attaquons nullement M. Willis ou ses associés. Ce que nous cherchons à éclaircir ici, c'est un très léger écart entre la façon dont les consultants et les comités ont évalué le quartile supérieur des emplois masculins; nous cherchons à déterminer si nous devrions rajuster ces données afin que les cotes des comités concordent avec celles des consultants et s'il y a des raisons convaincantes de le faire ou de ne pas le faire.

Les arguments que j'avance ici et que je suis sur le point de développer ont pour but de soulever devant le Tribunal des considérations dont il devrait être tenu compte en vue de décider s'il y a lieu d'effectuer des rajustements dans les circonstances. Il ne s'agit pas du tout d'une attaque contre M. Willis.

MEMBRE FETTERLY : Cela, je le comprends. Je pense que c'est très légitime.

Comme je vous l'ai dit hier, est-il nécessaire, pour y arriver, d'affirmer que la capacité des consultants d'évaluer sans parti pris

fondé sur le sexe est douteuse? Faut-il que vous affirmiez cela afin d'établir ou de soutenir que les résultats des comités doivent être retenus de préférence à ceux des consultants? Je ne le pense pas.

Me RAVEN : Je suis plutôt de votre avis qu'il y a diverses raisons pour lesquelles on devrait préférer les cotes des comités, non seulement les questions qui ont été posées ici et qu'on soulève quant à la façon dont les consultants eux-mêmes ont effectué ces réévaluations.

Par exemple, si je comprends bien l'argument avancé par Me MacLean l'autre jour, c'est que les consultants avaient une discipline un peu

190

différente, une attitude plus libérale que les comités. M. Willis l'a reconnu, notamment dans ses rapports au Comité mixte patronal-syndical, et il n'a trouvé rien à redire à ce propos. En fait, pour reprendre ses propres propos, compte tenu du contexte, il n'y a pas lieu de préférer notre compréhension antérieure et notre application de la discipline Willis dans d'autres contextes à celles du Comité directeur».

Je ne sais pas si cela soulève la question de partialité, consciente ou non, ou de différences qui tiennent à la façon de faire. Cela confirme, cependant, que (1) il y a des différences entre la discipline que M. Willis a adoptée dans d'autres études et celle qui caractérise le travail du Comité directeur; (2) la discipline du Comité directeur était plus conservatrice; (3) les cotes des comités étaient plus conservatrices que celles des consultants en ce qui concerne les emplois de haut niveau à prédominance masculine. Donc je ne soulève pas nécessairement ce point pour alléguer qu'il y a eu partialité.

717. La Commission a produit des preuves statistiques qui, soutient-elle, montrent un manque de cohérence entre le consultant Wisner, qui a effectué les 222 réévaluations (l'«étude Wisner 222» ou «la Wisner 222»), et ce qu'il est convenu d'appeler le groupe des quatre, qui a fait les 300 réévaluations (l'«étude Willis 300» ou «la Willis 300»). Cette preuve a aussi été présentée dans le but de savoir s'il y avait lieu de préférer les évaluations des comités aux réévaluations des consultants. A notre avis, elle a le même effet qui consiste à discréditer l'expert même dont la Commission a retenu les services pour mener une autre étude et sur lequel elle s'est fiée durant son enquête. Nous nous reportons ici aux paragraphes 184 et 185 des arguments écrits de l'avocate de la Commission :

[TRADUCTION]

(184) On peut raisonnablement conclure que les consultants en tant que groupe n'évaluaient pas sans parti pris fondé sur le sexe ou avec une partialité relativement plus grande que les comités à la lumière du fait que Esther Brunet, une évaluatrice qui a pris part aux réévaluations Willis II et qui connaissait bien la fonction publique fédérale, et que M. Willis considérait comme une évaluatrice compétente sans parti pris contre l'autre sexe, a produit des évaluations qui concordaient presque à 100 % avec celles des comités (en ce qui concerne les questionnaires français).

(185) Si l'allégation de partialité fondée sur le sexe est étayée par l'application non uniforme du plan d'évaluation aux évaluations des hommes et des femmes, alors il importe de faire le point sur l'uniformité ou la cohérence relative des évaluations effectuées par les consultants par rapport à celles des comités. L'uniformité se mesure statistiquement. Les preuves statistiques sur l'uniformité ou la cohérence des évaluations produites par les uns et les autres -- les comités par opposition aux consultants -- montrent que ce sont les comités et non les consultants qui ont été les plus cohérents dans leurs évaluations. Le plus grand taux d'erreur constaté chez les consultants, de signaler M. Sunter, est la preuve concluante qu'il faut préférer le comité au consultant. Par conséquent, l'allégation de partialité fondée

191

sur le sexe dont auraient fait preuve les évaluateurs des comités n'est pas fondée -- si on s'en tient aux statistiques -- pas plus d'ailleurs que l'allégation selon laquelle les cotes des consultants seraient plus uniformes ou plus fiables.

718. A notre avis, il y a d'autres caractéristiques valables qui peuvent expliquer les écarts entre l'étude Wisner 222 et l'étude Willis 300 et dont on devrait tenir compte en dehors des analyses purement statistiques. Si les deux études ont suivi les mêmes procédures, elles sont toutefois très différentes à d'autres égards. La Wisner 222 a été entreprise afin de valider un processus qui avait mis M. Willis dans l'embarras. Il s'agit d'une étude plus restreinte menée par un seul consultant ayant fait preuve d'une discipline plus libérale que le Comité directeur. L'analyse de M. Wisner était de nature ponctuelle seulement et ne visait pas à dépendre l'ensemble de la situation. Non seulement la

période sur laquelle portaient les deux études -- la Wisner 222 et la Willis 300 -- différait dans les deux cas, mais la première utilisait un échantillon d'emplois provenant qu'une population plus petite que la Willis 300. La Wisner 222 a prélevé son échantillon parmi les évaluations effectuées par les comités d'évaluation multiples et a exclu les évaluations faites par le Comité directeur. Les comités d'évaluation multiples fonctionnaient depuis environ trois mois à l'époque de l'étude Willis 222.

719. La Willis 300 était une étude de plus grande envergure qu'on a entreprise à la fin du processus. Elle avait pour objet de confirmer ou de contester l'analyse contenue dans la Wisner 222. Quatre consultants ont mené l'étude Willis 300, avec deux (ou plus) qui travaillaient en tandem. Un des consultants était membre d'un comité d'évaluation. L'échantillon des postes a été prélevé parmi l'entière population des postes évalués par les comités multiples, à l'exception de ceux contenus dans la Wisner 222. Comme on pouvait s'y attendre, il y a eu une plus grande concordance avec les évaluations des comités dans cette dernière étude.

720. Le moment où le soi-disant «groupe des quatre» a effectué ses réévaluations, l'étendue de l'échantillon, le nombre de consultants en cause, le processus suivi et les circonstances qui existaient à ce moment-là expliquent davantage, à notre avis, toute divergence réelle entre les comités d'évaluation et les consultants.

721. Le Tribunal a eu amplement l'occasion d'observer M. Willis durant son témoignage, tant lors de sa première comparution, qui a duré 36 jours d'audience, qu'au moment de sa seconde comparution, qui s'est échelonnée sur 4 jours d'audience. A notre avis, M. Willis s'est révélé être un témoin digne de foi qui a fait preuve de patience, de coopération et, ce qui est plus important, d'une impartialité irréprochable. Le Tribunal a accepté M. Willis à titre d'expert en matière d'équité salariale. Avant l'étude du Comité mixte, M. Willis avait acquis son expérience entièrement de sa participation aux études américaines sur l'équivalence des fonctions («comparable worth»). Il avait de l'expérience et tous s'accordaient pour dire qu'il était un pionnier dans son domaine. Il était reconnu comme un expert qualifié en équité salariale dans le système judiciaire des États-Unis.

722. Nous avons examiné les nombreuses occasions où le Comité mixte a demandé à M. Willis et à ses consultants de revoir les évaluations des comités ou de fournir un point de référence pouvant servir à des comparaisons avec les évaluations des comités. Ce rôle a été bien établi

et soutenu avant la fin prématurée de l'étude. Nous n'avons pas l'intention maintenant de considérer le rôle de M. Willis différemment de celui qu'il a rempli auprès des parties durant l'étude sur la parité salariale. Le Tribunal tiendra compte de tous les facteurs indiqués si jamais la question du rajustement des cotes est soulevée.

723. Les difficultés rencontrées par les comités d'évaluation multiples n'étaient pas imprévues et, pour bien les reconnaître et les comprendre, il faut avoir le contexte à l'esprit, à savoir la taille énorme de la fonction publique fédérale, sa dispersion géographique et la diversité des professions et métiers exercés par son effectif, aux qualifications des plus variées. Ces complications combinées aux obstacles logistiques rencontrés ont imposé un lourd défi à tous les intéressés. Les experts Armstrong et Durber ont souligné les difficultés inhérentes au processus complexe d'évaluation des emplois en ce qui a trait à l'équité salariale.

724. Compte tenu de la nature des travaux du Comité mixte, des nombreux participants aux antécédents variés ainsi que des conditions de travail dans lesquelles les comités d'évaluation multiples ont accompli leur tâche, il est normal et inévitable, de soutenir la Commission et l'Alliance, que l'évaluation des emplois aux fins de l'équité salariale entraînent des conflits. Ces conflits, expliquent-elles, surgissent de l'incompatibilité des valeurs des évaluateurs qui tentent, dans une étude sur la parité salariale, de mettre en question les stéréotypes et les attitudes de personnes ayant des vues plus traditionnelles. Dans ce contexte, les conflits qui ont surgi, fait-on valoir, sont compréhensibles et en fait inévitables.

725. Selon l'avocat de l'intimé, ce ne sont pas tous les comités qui ont envisagé leur tâche dans un «esprit d'équipe», certains d'entre eux ayant plutôt adopté le mode antagoniste ou accusatoire. M. Willis a expliqué que certains comités [TRADUCTION] «avaient tendance à se considérer en situation de négociation plutôt que comme une équipe de six ou sept personnes cherchant à atteindre un but commun». L'avocat de l'intimé estime que le Tribunal aurait «tort» d'accepter l'idée que l'évaluation des emplois dans le contexte de l'équité salariale doit inévitablement s'accompagner de conflits et d'antagonisme. A son avis, les évaluateurs doivent attaquer cette tâche comme un problème à résoudre dans un esprit de coopération, chacun étant tourné vers l'objectif commun à atteindre et s'acquittant de sa tâche à la lumière des faits pertinents. Selon l'employeur, on doit pouvoir conclure du processus suivi que les évaluateurs ont analysé les faits pertinents et leur ont accordé le poids qui convenait. Lorsque ces conditions sont réunies, de conclure l'avocat, le Tribunal peut considérer que les résultats sont fiables.

726. De l'avis de Mme Weiner, l'application du plan est plus importante que le plan lui-même lorsqu'il s'agit d'assurer des évaluations libres de partialité fondée sur le sexe. Elle a décrit les

193

caractéristiques du processus qui permet d'éviter ou de réduire au minimum ce genre de partialité. Il importe, d'abord, de constituer des comités composés d'hommes et de femmes provenant de niveaux organisationnels différents. Puis, a-t-elle ajouté, il faut former les comités en conséquence, discuter de la façon dont la partialité fondée sur le sexe peut se manifester, disposer d'informations complètes et à jour pour les besoins de l'évaluation des emplois et songer au mode de fonctionnement du comité. Elle fait sur ce dernier point les commentaires suivants (volume 8, p. 1092, ligne 13, jusqu'à la p. 1093, ligne 3) :

[TRADUCTION]

Q. Maintenant, qu'en est-il de la façon dont le comité s'acquitte de sa tâche au jour le jour?

R. Traditionnellement, les comités d'évaluation des emplois cherchent à être très efficaces. Ils tâchent d'évaluer le plus d'emplois possible dans une journée.

Un comité sur la parité salariale doit adopter une approche différente. Ses membres doivent s'interroger plus ouvertement et poser plus de questions s'ils ne comprennent pas tout à fait quelque chose qu'ils lisent dans les informations sur le poste. Il faut qu'ils discutent de partialité fondée sur le sexe, qu'ils s'entendent dire des choses comme «C'est juste une secrétaire» et qu'ils se rendent compte du peu de valeur qu'ils accordent ainsi au travail de la femme.

Tout cela, ça prend du temps, se questionner, approfondir.

727. Mme Weiner parle de «se questionner, approfondir» dans le contexte des évaluations effectuées par les comités. Bien qu'elle n'ait pas commenté directement l'aspect des conflits au sein des comités, elle a insisté sur le fait qu'on doit contester les valeurs traditionnelles lorsqu'on évalue les emplois en vue d'atteindre la parité salariale.

728. Le Tribunal n'est pas persuadé, compte tenu de la question qu'il est appelé à trancher, qu'on devrait lui demander de définir la nature et le degré de ce qu'il est permis, acceptable et légitime de discuter au sein

d'un comité. De plus, il est très difficile d'en mesurer l'effet, en particulier lorsqu'on remet en question les valeurs traditionnelles et qu'on en débat dans le contexte de la parité salariale. Le Tribunal n'est pas non plus prêt à proposer des solutions aux conflits qui surgissent entre les membres des comités qui peuvent avoir des opinions très nettes d'un côté ou de l'autre de cette question très délicate. L'étude et la mise en oeuvre de la parité salariale au Canada constituent une discipline relativement nouvelle qui est toujours en voie de se définir. Néanmoins, nous estimons qu'il est nécessaire, compte tenu des réserves exprimées par M. Willis au sujet du fonctionnement des comités et de la conduite des évaluateurs eux-mêmes, de déterminer si le processus suivi a atteint son objectif, qui était de produire des évaluations exemptes de partialité fondée sur le sexe.

194

729. En ce qui a trait à l'efficacité des mesures de protection mises en place durant l'étude, et plus spécifiquement des procédures définies par M. Willis comme faisant partie du processus Willis, nous estimons que l'opinion experte de M. Willis a été très convaincante et informative. À cause de son importance par rapport à l'évaluation des résultats, nous avons décrit de façon assez détaillée les procédures et les mesures de protection qu'il a recommandé qu'on adopte pour cette évaluation.

730. Le Tribunal estime qu'il lui incombe de commenter le processus que le Comité mixte a élaboré pour s'acquitter de son mandat. Qu'il suffise de dire que dès sa formation le Comité mixte a éprouvé des difficultés dans ses rapports de travail. Pour des raisons incompréhensibles, il a choisi de priver tant M. Willis que la Commission de tout pouvoir décisionnel réel. Le Comité a agi ainsi nonobstant l'impartialité de M. Willis et de la Commission, leur compétence et leur vaste expérience en matière d'équité salariale comparativement aux parties elles-mêmes. Tant à l'étape de la collecte des informations qu'à celle de l'évaluation de l'étude sur la parité salariale, le Comité mixte n'a pas suivi les conseils de M. Willis et a souvent refusé d'appliquer ses recommandations. Certaines des recommandations de M. Willis n'ont pas été mises en oeuvre à cause de décisions de «faire ou faire faire» largement entre les mains de l'employeur et motivées par des considérations financières. Cependant, d'autres recommandations de M. Willis, qui n'étaient pas compliquées par ces considérations, ont également été ignorées.

731. Selon M. Willis, le Comité mixte a constitué une faiblesse majeure de l'étude et, à notre avis, son opinion est bien fondée. Le ton antagoniste établi par le Comité mixte reflétait les difficultés profondes

qui existaient depuis longtemps entre les parties patronale et syndicale, lesquelles ont marqué l'étude tout au long de sa réalisation.

732. D'après la preuve, le chef de l'Équité salariale au Conseil du Trésor considérait l'étude sur la parité salariale menée par le Comité mixte, pour reprendre les mots de M. Willis, comme un «tas d'âneries» (volume 210 à la p. 27280). En revanche, l'Alliance voulait suivre une stratégie cohésive, comme en fait foi la lettre que M. Millar, au nom de l'Alliance, a envoyée pour annoncer la réunion de Mont Sainte-Marie. Cet incident et d'autres ont menacé les fondements de l'étude dès le départ et contribué largement aux difficultés qui devaient s'ensuivre. Le fossé entre les parties syndicale et patronale était évident dans la façon dont elles ont tenté de résoudre les questions. Il était évident même dans la disposition des sièges aux réunions du Comité, les représentants de l'employeur et des syndicats prenant place les uns en face des autres. Les parties se sont opposées à la proposition de M. Willis, qui a voulu modifier cette disposition. «Ils m'ont regardé comme si j'étais fou», a-t-il dit (volume 60, à la p. 7459).

733. M. Willis était contre l'idée des réunions que l'Alliance tenait avec ses membres avant et pendant la réalisation de l'étude. Il y a d'abord eu la réunion de Mont Sainte-Marie avant le début de l'étude proprement dite, où les membres de l'Alliance ont discuté de la sous-estimation du travail des femmes en l'absence des consultants et des

autres parties à l'étude. Pendant l'étude, l'Alliance a aussi tenu en soirée des réunions au cours desquelles les participants discutaient non seulement de leurs problèmes logistiques, mais aussi des évaluations proprement dites. De plus, le représentant de l'Alliance au Comité directeur de l'évaluation assistait à ces réunions du soir et répondait au besoin aux questions concernant les postes-repères du Comité directeur. A l'occasion d'une réunion qui a eu lieu une fin de semaine à l'automne de 1988, l'Alliance a tenu une séance de formation sur l'évaluation des emplois dans le contexte de la parité salariale, et ce, à l'insu de M. Willis et des autres parties. A cette réunion, les membres ont examiné certains des postes-repères du Comité directeur. Durant la fin de semaine, les participants ont reçu une formation destinée à les sensibiliser à la partialité fondée sur le sexe telle que la concevait l'Alliance. L'Alliance a justifié cette mesure exceptionnelle en faisant valoir qu'il était nécessaire de corriger ce qu'elle concevait comme les injustices historiques faites aux femmes, qu'elle considérait comme des «victimes» dans la population active.

734. Dans le contexte de l'étude et plus spécifiquement durant son exécution, M. Willis estimait qu'il n'avait pas l'appui nécessaire de la part des autorités patronales et des dirigeants syndicaux. Bien que le sous-comité des communications eût conçu un plan stratégique pour communiquer l'étude sur la parité salariale aux employés, M. Willis estimait qu'on n'avait pas assez accordé d'importance au besoin de communication de la part de la haute direction. Il avait initialement proposé au moins 10 jours de consultation au cours desquelles les consultants auraient rencontré face à face les gestionnaires ministériels et les dirigeants syndicaux. Aucune réunion du genre n'a eu lieu et M. Willis estime que c'est probablement à cause de cette lacune qu'il a fallu attendre si longtemps que les fonctionnaires remplissent leurs questionnaires.

735. Dans son témoignage, M. Willis a expliqué qu'il avait conçu le processus pour qu'il permette d'obtenir des résultats solides; si les résultats sont bons, il est sans importance que le processus comporte des failles. Nous avons examiné le plan Willis proprement dit et nous estimons qu'il s'agit d'un outil approprié pour l'évaluation des emplois dans le cadre de l'étude sur la parité salariale. Durant les plaidoiries finales, on a informé le Tribunal que le plan Willis ne faisait l'objet d'aucun litige entre les parties. Nous nous reportons à l'observation suivante de l'avocat de l'intimé (par. 41) :

[TRADUCTION]

41. Néanmoins, pour les besoins de la présente affaire, l'employeur accepte le plan Willis comme étant un plan approprié pour évaluer les emplois dans la fonction publique fédérale. Par conséquent, le Tribunal n'a pas à décider s'il est valable d'appliquer une pondération au plan Willis.

736. Nous nous fions à l'opinion experte de M. Willis comme quoi le questionnaire Willis, avec quelques modifications mineures, était capable de recueillir suffisamment d'informations sur les emplois pour qu'on puisse

196

les évaluer dans le cadre de l'étude sur la parité salariale. A son avis, le questionnaire contenait assez d'informations à partir desquelles un comité d'évaluation des emplois bien formé et supervisé pût produire des évaluations fiables et impartiales.

737. Le degré d'efficacité des mesures de protection prévues à l'étape de la collecte des données a déçu M. Willis. C'est à cette étape qu'on a

tâché de faire en sorte que les questionnaires soient bien remplis. Les détails de ces efforts sont décrits dans la décision sous la rubrique «Le processus Willis».

738. En ce qui concerne le rôle des coordonnateurs, nous sommes venus à la conclusion qu'il aurait été extrêmement difficile pour Willis & Associates, compte tenu de l'envergure de l'étude, d'agir eux-mêmes à titre de coordonnateurs sans entraîner des retards importants et occasionner des dépenses supplémentaires non négligeables au Comité mixte. Les coordonnateurs étaient chargés de communiquer directement avec les fonctionnaires choisis pour remplir les questionnaires. De plus, ce sont les coordonnateurs qui montraient aux titulaires comment remplir leurs questionnaires. Les consultants ont aidé le Comité mixte à préparer le matériel devant servir à la formation des coordonnateurs. Si le nombre de questionnaires remplis est une indication de la qualité du travail des coordonnateurs, alors on peut dire que leur travail a été très satisfaisant. Le pourcentage de questionnaires retournés a été impressionnant : près de 100 pour cent des questionnaires ont en fait été retournés.

739. Ce qui a le plus préoccupé M. Willis, c'est le temps qu'il a fallu attendre pour recevoir les questionnaires. Selon lui, plus les questionnaires sont retournés tard, plus la qualité de l'information en souffre. Il y a peu de preuves qui expliquent ce qui a pu contribuer à ces délais ou les causer. La preuve ne montre pas que les titulaires aient négligé de remplir les questionnaires dans un délai raisonnable, soit dans le délai prescrit de 10 à 14 jours suivant la formation reçue. De plus, il y a peu de preuves concernant les dates auxquelles les coordonnateurs auraient donné les séances de formation aux titulaires. Dans une certaine mesure, le grand nombre de substitutions a certes retardé le processus.

740. Même si l'efficacité du rôle des coordonnateurs semble faible, cela n'a pas empêché M. Willis d'aller de l'avant avec les évaluations. Celui-ci souhaitait qu'on amorce l'étude même si les informations n'étaient pas celles qu'il aurait souhaiter avoir. M. Willis a institué d'autres mesures de protection pour assurer qu'on dispose d'informations complètes sur les postes, par exemple faire revoir les questionnaires par des réviseurs ou les évaluateurs eux-mêmes. A notre avis, les limitations associées au rôle des coordonnateurs n'ont pas influé de façon importante sur la question de la fiabilité.

741. Les réviseurs ont appliqué une technique raffinée de contre-vérification. Il leur incombait de faire en sorte que les questionnaires contiennent des informations complètes quant aux faits pour que les comités puissent procéder à leur évaluation.

742. La fonction de révision n'a pas été prise en charge par M. Willis. Pour déterminer si elle a été suffisante, on doit prendre en considération la formation donnée, le témoignage des personnes qui ont effectivement accompli ce travail, la recherche de la Commission (menée par un chercheur externe, pièce HR-245) ainsi que les propres observations et commentaires de M. Willis. Les réviseurs qui ont témoigné estimaient avoir bien fait leur travail. En téléphonant aux fonctionnaires qui avaient retourné leurs questionnaires, ils estiment avoir pu obtenir les informations nécessaires. Bien que M. Willis eût préféré plus d'entrevues en personne, dans l'ensemble il n'avait rien à redire à propos du rendement des réviseurs ou du rôle qu'ils avaient joué dans l'étude sur la parité salariale.

743. Les réviseurs ont reçu la même formation initiale sur le plan Willis que les évaluateurs du Comité directeur. Ils ont également bénéficié de «formation sur le tas» de la part du consultant au besoin. Nous estimons qu'ils se sont bien acquittés de leur tâche sans problème apparent, si on fait exception des quelques réviseurs qui, dans leurs comités, avaient tendance à produire des évaluations différentes des autres membres. Par contre, rien n'indique que ces réviseurs n'avaient pas bien accompli leur travail ou qu'ils avaient influencé indûment les autres. Ils n'étaient que six, finalement, ce qui ne représente qu'une faible proportion des évaluateurs ayant fait fonction de réviseur.

744. On peut comprendre pourquoi M. Willis aurait préféré participer directement à la révision des questionnaires. Cependant, il semble peu probable, compte tenu du volume de questionnaires, qu'un consultant ait pu accomplir cette tâche dans le délai fixé. Après avoir examiné attentivement la preuve concernant la collecte des informations, nous acceptons l'opinion de M. Willis et concluons que les informations sur les postes étaient de qualité suffisante si on tient compte de toutes les mesures de soutien qui ont été prises.

745. L'uniformité est une caractéristique importante dans le processus d'évaluation des emplois aux fins de la parité salariale. Le plan Willis devrait être appliqué de façon uniforme, spécialement lorsque des comités d'évaluation multiples sont en cause. Cette exigence, si elle est respectée par les participants, ne signifie pas forcément que le processus est exempt de partialité fondée sur le sexe. En revanche, le manque d'uniformité entre les comités n'implique pas nécessairement un parti pris, et il ne revêt pas non plus une importance critique par rapport à la question de la fiabilité. En dernière analyse, ce qui préoccupait M. Willis, c'était de savoir si les résultats étaient empreints de

partialité. Toutefois, compte tenu du contexte de cette étude et en vue d'évaluer jusqu'à quel point le processus a fonctionné, il nous apparaît prudent de commenter la question de savoir si les comités d'évaluation multiples ont appliqué de manière uniforme la discipline établie par le Comité directeur.

746. Quelques-uns des cinq comités initiaux -- à savoir les comités 1 et 2 et la première version du comité 4 -- ont bien fonctionné. Les neuf comités multiples créés à la suite de la restructuration des cinq comités initiaux semblent aussi avoir bien fonctionné dans l'ensemble. La plupart

198

des comités d'évaluation multiples ont en fait tâché de suivre les postes-repères du Comité directeur, de respecter la discipline établie par ce dernier et d'appliquer la même méthode d'évaluation des postes que celui-ci. La preuve -- les premiers résultats des tests effectués par le sous-comité de fiabilité inter-comités tout au moins -- indique que les comités ont interprété les facteurs Willis et appliqué le plan de façon uniforme. Dans une certaine mesure, les postes-repères du Comité directeur ont eu un effet stabilisateur sur le fonctionnement des comités d'évaluation multiples et sur l'étude dans son ensemble. Cela est surtout évident dans la réponse suivante que M. Willis a donnée à une question du Tribunal sur la première version du comité 3 (volume 69, p. 8676, lignes 8-18) :

[TRADUCTION]

Mais, finalement, une des choses qui ont peut-être aidé à stabiliser les valuations, c'est qu'on avait les postes-repères du Comité directeur de l'évaluation pour eux et que peut-être ils se sont fatigués de se battre chacun pour leur côté et qu'ils ont décidé de suivre les postes-repères du Comité directeur. Je n'étais pas du tout convaincu de pouvoir laisser les choses ainsi. Mais je n'ai pu observer aucun problème particulier dans les évaluations effectives que nous avons pu examiner.

747. Le Tribunal se penchera maintenant sur la formation que les membres des comités ont reçue en vue de pouvoir exercer convenablement leur fonction d'évaluateur. La méthode que M. Willis utilise à l'égard des stéréotypes sexuels et des valeurs traditionnelles consiste à donner instruction aux évaluateurs de décomposer un emploi en ses composantes et d'évaluer chacune d'entre elles séparément de façon à assurer une évaluation impartiale. L'opinion de M. Willis diffère de celle de Mme Armstrong quant à savoir si sa méthode aurait dû inclure une

sensibilisation plus structurée à ces questions, notamment en ce qui concerne la sous-estimation du travail des femmes. A notre avis, le fait que cette sensibilisation n'a pas formellement été incluse par M. Willis n'a pas accru le risque d'évaluations empreintes d'un parti pris contre l'autre sexe. M. Willis a préféré parler de «formation sur le tas», une formule qu'il a appliquée avec succès dans des études antérieures. De plus, le Comité mixte avait le pouvoir de décider ce que devait comprendre la formation et quelle formation on devait effectivement dispenser. M. Willis a été critiqué par l'Alliance, au cours de la présente audience, pour ne pas avoir donné une formation du type sensibilisation que préconisait Mme Armstrong et la Commission de l'équité salariale de l'Ontario à l'égard de la partialité fondée sur le sexe. Il y a lieu de signaler, cependant, que l'Alliance a approuvé la méthode de formation de M. Willis au début de l'étude sur la parité salariale lorsqu'elle faisait partie du Comité mixte. La critique de l'Alliance à l'endroit de M. Willis semble motivée par le fait que ce dernier avait désapprouvé le fait que l'Alliance ait dispensé ce genre de formation à l'une de ses réunions tenues en l'absence des consultants et des autres participants. Par ailleurs, M. Willis a commenté un autre aspect qui avait trait à la qualité de cette formation (volume 211, p. 27483, ligne 24, jusqu'à la p. 27484, ligne 20) :

199

[TRADUCTION]

Q. J'ai une autre question, celle-là à propos d'un point vous avez discuté assez longuement avec mon ami Me Raven. C'est à propos de la sensibilisation des participants à une étude aux questions touchant la partialité fondée sur le sexe. Vous vous souvenez de ce point?

R. Oui, je m'en souviens.

Q. Pour décider si une telle formation est bénéfique, est-il pertinent de savoir quelque chose sur la qualité de la formation?

R. Bien entendu.

Q. Pourriez-vous élaborer en peu, s'il vous plaît?

R. J'estime qu'il serait important que la personne donnant cette formation soit reconnue comme quelqu'un d'impartial ayant reçu une formation dans ce domaine.

Q. Si la formation n'est pas bien faite ou si elle est empreinte de partialité, pourrait-elle avoir un effet autre que celui de transmettre ses propres préjugés?

R. C'est possible qu'elle puisse avoir pour effet de créer encore plus de préjugés.

748. Nous sommes d'avis que, reconnaissant la vaste expérience pratique de M. Willis dans la conduite d'études sur la parité salariale, son approche pratique est valable et acceptable. Nous l'affirmons nonobstant l'opinion de Mme Armstrong qui repose, semble-t-il, entièrement sur la recherche et la documentation disponible.

749. En ce qui concerne le processus d'évaluation des postes proprement dit, il y a des preuves anecdotiques comme quoi le processus n'aurait pas fonctionné aussi bien qu'il aurait dû. M. Willis a témoigné à propos du malaise qu'il avait face au comportement de certains membres des comités, en particulier la première version du comité 3, qu'il a qualifié de «deux camps en guerre». Le Comité mixte l'a empêché de prendre les mesures correctives qu'il estimait nécessaires à l'endroit des évaluateurs du comité 3 qui manifestaient un parti pris contre l'autre sexe.

750. Le Tribunal a eu l'avantage d'observer et d'entendre des témoins qui avaient participé aux comités d'évaluation. En général, on peut dire qu'ils ont apporté un élément de «réalité» à ce qu'on peut le mieux décrire comme un processus long, ardu, compliqué, stressant et difficile. Dans l'ensemble, la question de savoir si les informations fournies dans les questionnaires étaient suffisantes n'a pas posé de difficulté à ces évaluateurs. Si et quand un comité avait besoin de renseignements supplémentaires pour achever une évaluation, il s'en remettait à la mesure prévue pour palier à la lacune, c'est-à-dire demander au réviseur de compléter ou de clarifier les informations, ou encore d'en obtenir de nouvelles.

751. M. Willis a témoigné au sujet de certaines des forces de l'étude sur la parité salariale. Il en a relevé trois : premièrement, le grand nombre d'individus ayant participé aux comités d'évaluation; deuxièmement, le nombre considérable d'emplois variés évalués; et troisièmement, le grand nombre de postes compris dans l'échantillon, ce qui lui a permis de faire face à une «disparité légèrement plus grande» dans l'information sur les emplois que si l'étude avait porté sur une population plus restreinte. M. Willis estimait que la composition des comités était «assez bien» équilibrée puisqu'on y trouvait des représentants syndicaux et patronaux de

milieux variés, encore que les syndicats aient eu de la difficulté à nommer des représentants masculins. Selon la preuve, certaines des évaluatrices étaient membres de syndicats à prédominance masculine, ce qui a contribué à diversifier davantage la composition des comités.

752. Un des problèmes que M. Willis a reconnus était la participation de cadres de gestion ayant reçu une formation en classification. Sept des évaluateurs nommés par la partie patronale avait une connaissance approfondie du système de classification dans la fonction publique fédérale. Ils ont fait partie de quatre des comités d'évaluation ainsi que du Comité directeur. Les problèmes liés à l'expérience de la classification ont surgi durant le processus d'évaluation. La preuve statistique, toutefois, n'a pas relevé l'expérience de la classification de ces individus comme ayant influé sur les cotes de consensus des comités d'évaluation multiples. La preuve anecdotique semble indiquer que ces individus n'ont exercé que très peu d'influence, voire aucune, et que les autres participants avaient tendance à les ignorer.

753. Un autre problème qui s'est présenté a été la participation de certains partisans de l'Alliance qui s'étaient donné pour mission d'accroître la valeur des emplois à prédominance féminine. Il y a eu des tentatives malavisées, qui ont pris la forme d'affrontements et d'intimidation, pour influencer les évaluations de certains des membres. Les écarts constatés entre les cotes des comités et celles résultant des réévaluations effectuées par les consultants montrent qu'on a sous-évalué certains emplois à prédominance masculine, mais ne montrent pas que ces individus malavisés aient atteint leur objectif de persuader les autres de surestimer les emplois à majorité féminine. Comme l'analyse de M. Sunter le révèle, c'est presque entièrement dans le traitement qui a été fait des questionnaires portant sur les postes à prédominance masculine qu'on relève des écarts significatifs entre les comités et les consultants. De plus, les résultats des tests sur la fiabilité inter-évaluateurs révèlent que la majorité des évaluateurs ayant produit des cotes aberrantes -- qu'ils fussent du côté patronal ou syndical -- ont manifesté une préférence masculine. Par conséquent, toute tentative volontaire de la part de certains membres de l'Alliance de surestimer les emplois à prédominance féminine a été infructueuse. On peut par ailleurs trouver un certain réconfort dans le témoignage de tous les évaluateurs de l'Alliance qui ont déclaré qu'à aucune réunion de l'Alliance on n'a dit aux membres de surestimer les emplois à prédominance féminine et de sous-évaluer les emplois à majorité masculine.

754. Certains des évaluateurs ont été identifiés tant par les consultants que par les résultats des tests sur la fiabilité

inter-évaluateurs comme ayant produit des cotes aberrantes. Au cours de l'étude sur la parité salariale, on a tâché de déterminer si ces évaluateurs exerçaient une influence sur le consensus final du comité. L'analyse statistique a démontré que leur influence a été négligeable. De plus, en les observant au travail dans les comités d'évaluation, M. Willis n'a pu déceler qu'ils exerçaient quelque influence que ce soit sur les autres membres.

755. Un des aspects qui a le plus racheté les défauts de l'étude sur la parité salariale fut le travail du Comité directeur de l'étude, auquel M. Willis a souscrit sans réserve. Lorsque le Comité directeur a terminé ses travaux, M. Willis était convaincu qu'il avait accompli un bon travail. Le consultant a procédé à plusieurs examens du travail du Comité directeur qui ont révélé certains écarts entre les évaluations de celui-ci et celles des consultants. M. Willis ne s'est pas préoccupé de ces écarts, puisqu'il n'y avait aucune indication de partialité fondée sur le sexe dans les évaluations du Comité directeur. A ses yeux, il est normal qu'il y ait des différences entre les comités et les consultants. De plus, la présence de telles disparités ne signifie pas nécessairement que le consultant a «toujours raison».

756. D'après M. Willis, pour déterminer s'il existe un problème réel ou non, il faut répondre aux quatre questions suivantes :

- (i) Quelle est l'étendue des disparités au niveau des cotes totales d'une évaluation particulière?
- (ii) A quelle fréquence les disparités se manifestent-elles?
- (iii) La justification : pourquoi les comités ont-ils fait ce qu'ils ont fait?
- (iv) Y a-t-il une tendance dans les disparités, et si oui laquelle?

757. Lorsque l'étude est terminée, M. Willis examine le score total afin de répondre à deux des quatre questions ci-dessus, à savoir quelle est l'étendue des disparités et à quelle fréquence elles se manifestent. Il effectue cet examen avec l'aide d'un statisticien, à qui il demande également de répondre à la quatrième question, à savoir s'il y a une tendance dans les disparités. Divers facteurs peuvent expliquer les disparités mentionnées à la question (iv), mais à ce stade-ci de l'étude M. Willis ne s'en préoccupe pas.

758. Selon M. Willis, lorsque l'étude est achevée, ce qu'il importe de déterminer c'est la mesure dans laquelle les comités se sont écartés des évaluations des consultants. D'après lui, les autres considérations sont sans importance à ce stade-ci. Il ne tient compte que des résultats essentiels parce que les comités d'évaluation ne fonctionnent plus. Il n'est plus utile au consultant de savoir si les comités ont appliqué le plan correctement puisqu'il n'est plus possible de conseiller et de former les membres.

202

759. A plusieurs reprises dans son témoignage, M. Willis a exprimé l'avis que les résultats étaient plus importants que le processus. Par résultats, il entendait les comparaisons entre les évaluations des comités et les réévaluations des consultants.

760. Cependant, étant donné notre interprétation de l'article 11 de la Loi, à savoir que la causalité est implicite dans la loi, nous devons examiner la question de savoir si les différences qui ont surgi durant le processus entre les consultants et les comités s'expliquent par la partialité fondée sur le sexe ou par d'autres facteurs. Il s'ensuit donc qu'il est non seulement nécessaire mais crucial d'examiner la preuve en détail afin de déterminer si les écarts entre les comités et les consultants sont dus à un parti pris contre l'un ou l'autre sexe.

761. L'Alliance a présenté une preuve concernant des analyses effectuées par deux témoins qu'elle a assignés à comparaître. Ces derniers ont examiné la justification raisonnée des comités et des consultants dans le but d'expliquer les disparités entre leurs évaluations respectives. Avant que le premier de ces témoins amorce son témoignage, l'employeur a présenté au Tribunal un aveu dans lequel il reconnaît ce qui suit :

[TRADUCTION]

4. L'employeur fait l'aveu et la clarification qui suivent afin de réduire les questions en litige et d'éviter qu'on doive inutilement prendre du temps d'audience pour produire des preuves.
5. L'employeur admet que les disparités entre les consultants et les comités dans les réévaluations Wisner 222 et Willis 300 peuvent s'expliquer par des facteurs autres que la partialité fondée sur le sexe au sein du Comité mixte syndical-patronal.

6. Pour clarifier les questions, l'employeur n'invoquera pas les raisons des disparités comme preuve de partialité fondée sur le sexe dans le processus ou de partialité dans les résultats.
7. Par conséquent, l'employeur soutient que les preuves qui analysent les raisons des disparités n'aident pas le Tribunal à évaluer :
- a) la fiabilité du processus; ou
  - b) la fiabilité des résultats.

(pièce R-154)

762. M. Willis a eu l'occasion de commenter les deux analyses présentées par les témoins de l'Alliance. A son avis, ni l'une ni l'autre ne sont utiles pour déceler la partialité fondée sur le sexe dans une étude d'envergure ou pour explorer les disparités entre les consultants et les comités. D'après son expérience, l'évaluation individuelle des différences en fonction des justifications raisonnées ne saurait révéler l'existence de partialité fondée sur le sexe. Le Tribunal accepte l'avis de M. Willis. Notre décision ne sera pas basée sur les justifications raisonnées des différences individuelles entre les évaluateurs des comités et les

203

consultants sur une question donnée, mais reposera plutôt sur l'examen de l'ensemble de la preuve qui se rapporte aux évaluations des comités et des consultants.

763. M. Willis voulait des questionnaires qui étaient complets et qui mettaient l'accent sur des renseignements concrets. Un questionnaire incomplet amène l'évaluateur à faire des suppositions qui augmentent les disparités éventuelles. Les disparités relevées dans cette étude étaient en général plus nombreuses que ce que M. Willis est habitué de voir. En revanche, M. Willis n'avait jamais participé à une étude de l'envergure de l'étude sur la parité salariale menée par le Comité mixte et il n'était pas en mesure de superviser les 522 réévaluations effectuées, dont certaines ont été exécutées durant l'étude et d'autres une fois que celle-ci a été terminée.

764. Nous examinerons maintenant les questions (i), (iii) et (iv) de M. Willis. A de nombreuses occasions lorsqu'il a témoigné, M. Willis a parlé d'un niveau de tolérance applicable aux différences entre les évaluations des comités et celles des consultants. Les variations en pourcentage qu'il utilise sont simplement une fonction de son expérience et de ce qu'il considère comme acceptable. En se basant sur la qualité de

l'information dont disposait le Comité directeur, il s'attendrait à trouver une variation fortuite de l'ordre de 10 à 12 pour cent, positive ou négative, dans les évaluations. Étant donné que la qualité des informations dont disposaient les comités d'évaluation multiples n'était pas, à son avis, aussi élevée que celle dont bénéficiait le Comité directeur, il s'attendrait à trouver une variation fortuite d'entre 15 et 20 pour cent dans leur cas. Les évaluateurs ont davantage tendance à faire des suppositions lorsqu'ils s'appuient sur une information de moins bonne qualité.

765. M. Willis a expliqué que la variation fortuite se produit lorsqu'on porte des jugements de valeur sur le sens des faits présentés dans le questionnaire. Il considère que dans une vaste étude comme celle sur la parité salariale où le nombre de postes à évaluer est considérable, on peut accepter une plus grande disparité du fait qu'on dispose d'informations relativement faibles sur les emplois. La situation ne devient inquiétante que si, avec le temps, la variation cesse d'être fortuite et devient systématique, a-t-il ajouté. Il définit la variation systématique comme une valeur ou des valeurs qui [TRADUCTION] «sont de façon constante plus élevées ou plus faibles qu'une évaluation objective de certains types d'emploi». M. Willis considère les expressions «variation systématique» et «partialité fondée sur le sexe» comme synonymes.

766. M. Shillington a témoigné sur la distinction entre systématique («pattern», en anglais) et fortuit dans une étude d'envergure et sur la difficulté de qualifier quelque chose de fortuit. Voici ce qu'il dit à ce propos (volume 86, p. 10540, ligne 9, jusqu'à la p. 10541, ligne 13) :

[TRADUCTION]

Q. Comment savoir qu'on est en présence de quelque chose de fortuit par opposition à quelque chose qui ne l'est pas, quelque chose qui révèle une tendance ou qui est systématique?

204

R. Parfois vous êtes à l'aise d'employer un terme sans le définir, et le terme fortuit (ou ses synonymes «aléatoire» et «au hasard») est l'un de ces termes que chacun utilise sans se poser de question. Je pense que tout le monde sait de quoi vous parlez, mais dès que vous cherchez à le définir, les choses se compliquent.

Si vous montrez à quelqu'un une suite de nombres, très souvent les gens l'examineront et vous pouvez dire «Est-ce le fruit du hasard

ou non?» Il est très difficile de montrer qu'une suite quelconque est le fruit du hasard. Il est souvent plus facile de montrer que ce n'est pas le cas.

Je vais écrire une séquence. Supposons qu'à quatre essais de pile ou face j'obtiens face, pile, face, pile. Vous pouvez examiner ce résultat et dire qu'il est possible qu'on obtienne ces résultats avec une pièce non truquée. Vous avez obtenu pile 50 % des fois et face 50 % des fois.

Mais si vous continuez d'obtenir pile, face, pile, face, pile, face, vous commencez à vous dire que cela n'est plus fortuit. Oui, vous avez obtenu pile la moitié des fois et face la moitié des fois, mais cela est beaucoup trop systématique.

Il est très, très difficile de définir ce qui est fortuit. Il est beaucoup plus facile de dire : «Cela n'est pas fortuit, n'est pas dû au hasard. Il semble y avoir une tendance, quelque chose de systématique ici.»

767. M. Shillington affirme plus loin (volume 86, à la p. 10543, lignes 1-8) :

[TRADUCTION]

Alors il est facile de montrer que ce n'est pas fortuit, qu'il y a une séquence. Mais il est presque impossible de prouver que c'est fortuit.

Nous utilisons les expressions «fortuit» et «au hasard» comme termes polyvalents pour désigner quelque chose que nous ne connaissons pas. Si on joue à pile ou face pendant un certain temps, on dit que le résultat est fortuit parce qu'on ne peut pas bien prédire si la prochaine fois ce sera pile ou face.

768. M. Willis a confirmé à la fin de l'étude qu'il était prêt à accepter une grande disparité dans les évaluations pourvu qu'il n'y ait pas de variation systématique. Il n'aime pas voir de variation systématique du tout. Dans ses premiers témoignages, il a dit que si la variation était inférieure à 2 pour cent il ne rajusterait probablement pas les évaluations. Il a mentionné au volume 61, à la p. 7596, lignes 5-11 :

[TRADUCTION]

R. En dernière analyse lorsque l'étude est terminée, dans bien des cas nous devons passer aux recommandations et à la mise en oeuvre.

A ce stade-là, je pourrais décider que certains correctifs s'imposent. Mais évidemment, si [la variation] est inférieure à 2 pour cent, la disparité

205

salariale est tellement minime que j'imagine qu'il faudrait que je l'accepte.

769. En règle générale, même si on dispose des meilleures informations possible sur les emplois, il est normal, selon M. Willis, qu'il y ait un écart de plus ou moins 10 pour cent entre les comités et les consultants. Tout écart supérieur à 10 pour cent doit nous alerter à la possibilité qu'il puisse y avoir un problème quelconque concernant les évaluations. Dans une vaste étude comme celle sur la parité salariale, M. Willis demande l'aide d'un statisticien pour déterminer si la variation est systématique.

770. La nature de cette démarche, que M. Willis considère davantage comme un art qu'une science, fait qu'il est difficile de quantifier l'évaluation des emplois statistiquement ou mathématiquement. Le Tribunal s'est longtemps attardé à la présentation de données statistiques. Finalement, les experts en statistique qui ont témoigné, MM. Shillington et Sunter, ont déclaré que l'analyse statistique ne peut déceler la présence de partialité fondée sur le sexe.

771. Les conclusions de M. Sunter résultent de la vérification d'hypothèses. Dans ses analyses explicatives, il s'appuie sur des critères de probabilité et des modèles mathématiques pour expliquer la variation dans les données. Ses conclusions ne reposent pas entièrement sur le raisonnement scientifique et les applications mathématiques; elles s'appuient en partie sur des hypothèses concernant la «nature du monde». M. Sunter a signalé à différents moments dans son témoignage quand son intuition l'avait aidé à tirer ses conclusions. Les exemples suivants que nous reproduisons ne sont pas exhaustifs. Aux lignes 8 à 17, de la p. 13221 du volume 110, il fait remarquer ce qui suit :

[TRADUCTION]

Quand j'ai dit que le truc initial était très inattendu, c'était parce que j'estimais que si le consultant a toujours raison et le comité toujours tort, alors mon intuition de statisticien me dit que cela devrait donner lieu à une plus grande variance des scores des comités ainsi qu'à une covariance négative et à une corrélation négative entre la différence et les scores des comités, ce qui est

exactement la relation qui est reproduite par le modèle 2 que vous voyez.

772. Plus loin, aux lignes 10 à 20, p. 14387 du volume 119, il dit :

[TRADUCTION]

Il y a une association plus forte, positive entre DIFF et CONS qu'entre DIFF et COMM. Maintenant, je vous dirai que d'après mon intuition de statisticien -- je n'ai pas à justifier cela, c'est juste qu'on finit par développer une intuition, et mon intuition de statisticien est surprise par ceci, si c'est réellement le consultant qui a tort... pardon, si c'est le comité qui a tort. Je m'attendrais à ce que les associations soient quelque peu différentes, mais c'est mon intuition seulement.

206

773. De la p. 15046, ligne 19, jusqu'à la p. 15047, ligne 2, du volume 123, M. Sunter fait l'observation suivante :

[TRADUCTION]

Je pense qu'il a demandé s'il s'agissait d'outils pertinents dans le contexte de ce que M. Shillington faisait dans le sous-comité de fiabilité inter-évaluateurs, et j'ai dit «oui». Vous savez, il était dans une situation différente, préoccupé par des choses différentes, et je suppose qu'il a utilisé les deux tests à la suite d'une évaluation plus ou moins intuitive, ce qu'il avait parfaitement le droit de faire dans les circonstances...

774. Et une fois de plus, dans le volume 217, à la p. 28225, lignes 9 à 23, M. Sunter fait remarquer ce qui suit :

[TRADUCTION]

Généralement, en théorie de la prise de décision, on associe pertes et gains à diverses décisions, et la façon dont vous prenez une décision est un facteur à considérer -- si vous vouliez le faire techniquement, il vous faudrait entrer dans tous ces trucs, et j'essaie d'éviter tout ça en vous disant : «Je n'ai pas de fonction de perte à offrir ici. J'ignore comment vous devriez prendre cette décision.» Si vous insistiez pour que j'élaboré une fonction de prise de décision, je suppose que je pourrais le faire.

Voilà pourquoi je ne prends pas position sur cette question. Faites le rajustement ou ne faites pas le rajustement -- cela dépend du genre de processus décisionnel intuitif qui est le vôtre, mais je ne suis pas près de prendre cette décision pour vous.

775. Les deux statisticiens reconnaissent que l'analyse statistique peut appuyer la preuve même si en soi elle n'est pas forcément concluante. M. Shillington discute de la signification et de la non-signification des résultats en termes de preuve faible ou solide. A son avis, un résultat significatif n'est pas concluant en soi. Il peut, toutefois, amener le statisticien à faire une inférence qui met en question l'hypothèse. Selon M. Sunter, l'analyse statistique peut venir étayer quelque chose qui semble déjà plausible. Seule, l'analyse peut très rarement fournir des explications plausibles. Dans cette fonction limitée, nous estimons que les analyses statistiques sont des outils appropriés et utiles. Par conséquent, nous concluons que les statistiques sont accessoires à la fonction première des évaluateurs -- qui est de porter un jugement de valeur -- ainsi que du Tribunal, qui est de déterminer la fiabilité des résultats.

776. Lors de sa dernière comparution devant le Tribunal, M. Sunter a reconnu qu'il y avait des limitations à l'applicabilité des statistiques au règlement des questions dont le Tribunal est saisi. On trouve ses commentaires à cet effet à la p. 28301, lignes 13-22, du volume 217 :

207

MEMBRE FETTERLY : Je suppose que ce à quoi je veux en venir est ceci : les statistiques ne nous disent pas nécessairement tout. Je pense que vous pouvez être d'accord sur ce point, n'est-ce pas?

LE TÉMOIN : Oui, je souscris à cette observation générale.

MEMBRE FETTERLY : Alors nous pourrions devoir considérer d'autres facteurs qui ne relèvent peut-être pas de votre spécialité?

777. Les tests de M. Sunter aident à relever les différences statistiquement significatives entre les résultats de l'étude Wisner 222, de la Willis 300 et de la base de données combinées (522) par comparaison aux évaluations des comités. Selon M. Sunter, les différences ne présentent aucune tendance particulière. Il a trouvé des différences significatives entre les consultants et les comités dans les deux études -- mais davantage dans la Wisner 222 que dans la Willis 300 -- relativement aux questionnaires provenant des emplois à prédominance masculine. Les résultats de ses tests ont relevé des différences principalement à l'égard

des postes de niveau supérieur à majorité masculine et un peu moins à l'égard des postes de niveau supérieur à prédominance féminine. Dans l'ensemble, les questionnaires relatifs aux postes à prédominance féminine avaient une distribution des cotes numériques plus faible que les questionnaires portant sur les emplois à majorité masculine. Nous sommes conscients du fait que les différences constatées par rapport aux questionnaires à prédominance féminine n'étaient pas significatives statistiquement.

778. M. Shillington a avancé une opinion concernant les analyses que M. Sunter a effectuées des «autres causes» -- autres que celles fondées sur le sexe -- pouvant expliquer les différences entre les cotes données par les consultants et celles des comités. Une de ces analyses a consisté à comparer les différences pour voir si elles étaient associées à la distribution relative des questionnaires dans les fourchettes de cotes supérieure et inférieure. Contrairement à M. Sunter, M. Shillington était d'avis qu'il serait très difficile de séparer ces deux questions d'analyse de données dans le but de déterminer si les différences pouvaient s'expliquer par un facteur autre que le sexe. Sur ce point, M. Shillington dit ce qui suit (volume 131, p. 16045, ligne 21, jusqu'à la p. 16046, ligne 21) :

[TRADUCTION]

R. Oui, et l'analyse qui est derrière cela.

Les analyses de régression ont été effectuées dans un certain sens pour voir s'il y avait une relation entre les différences entre les consultants et le comité du point de vue du sexe. Il est aussi possible que les différences qui ont pu exister entre les cotes des consultants et celles des comités n'aient pas été directement reliées au sexe, mais peut-être à des valeurs numériques élevées par opposition à des valeurs faibles.

208

Ce facteur de confusion tient au fait qu'il y a une forte tendance dans les données qui fait que les questionnaires des hommes ont tous des cotes ou valeurs élevées par rapport à ceux des femmes et que les questionnaires des femmes ont plutôt tendance à provenir de la fourchette inférieure de cotation, ce qui veut dire qu'on ne peut séparer ces deux questions d'analyse de données, ou qu'il est difficile de les séparer.

LA PRÉSIDENTE : Que voulez-vous dire?

LE TÉMOIN : Vous ne pouvez séparer la question de savoir si la tendance est reliée au sexe ou si elle est reliée au fait que les cotes étaient élevées ou faibles.

779. Sur le même sujet, M. Shillington tient les propos suivants dans le même volume (p. 16048, ligne 16, jusqu'à la p. 16049, ligne 11) :

[TRADUCTION]

Dans les circonstances, si on revient à l'analyse des cotes Willis et du rajustement éventuel des résultats, nous avons une situation qui -- dans la mesure où il y a une tendance ici, si quelqu'un arrivait et disait que cela n'était peut-être pas dû au sexe, au fait d'être un homme ou une femme, mais plutôt à la professionnalisation ou au fait que certains questionnaires aient eu des cotes beaucoup plus fortes que d'autres, vous auriez de la difficulté à extraire ces deux hypothèses distinctes de l'analyse parce que vous avez une situation dans laquelle les hommes ont surtout eu des cotes élevées, tandis que les femmes ont surtout eu des cotes faibles. Il y a donc confusion entre le fait d'être un homme [ou une femme] et celui d'avoir eu des cotes élevées ou faibles.

Ce phénomène est reflété dans la distribution. C'est pourquoi il s'agit d'une question de distribution. La distribution des cotes Willis pour les hommes avait tendance à être sensiblement plus élevée que la distribution des cotes Willis pour les femmes. C'est un facteur confusionnel. Vous devrez donc faire preuve de prudence dans vos interprétations.

780. Tout compte fait, M. Shillington estime que l'on doit utiliser ces analyses avec prudence, et nous nous reportons à la réponse qu'il a donnée dans le volume 131, p. 16049, ligne 20, jusqu'à la p. 16052, ligne 7 :

[TRADUCTION]

LE TÉMOIN : C'est plus une question d'interprétation et... je ne pense pas qu'on puisse trop -- je ne suis pas M. Sunter, mais je crois que nous devons faire attention lorsque nous utilisons ces analyses, à cause des différences dans les distributions.

LA PRÉSIDENTE : Par exemple, lorsque l'on compare des lignes de régression, nous examinons habituellement les différences -- en tout cas nous avons examiné l'écart salarial à l'aide des lignes de régression,

par exemple, pour calculer la distance entre elles. Alors vous les comparez pour voir quelle est la distance.

LE TÉMOIN : Oui.

LA PRÉSIDENTE : C'est à cela que je pense lorsqu'on me dit que vous ne pouvez comparer ces deux lignes de régression. Alors quand M. Sunter affirme qu'on ne peut pas comparer ces deux lignes de régression, je réponds «pourquoi les comparer?» Voilà pourquoi je suis un peu mêlé.

Quand vous dites que vous ne pouvez pas les interpréter, voulez-vous dire que parce que dans la ligne de régression des hommes on a une distribution de valeurs élevées et de valeurs faibles, mais une tendance à trouver surtout des valeurs élevées, tandis que dans celle des femmes on trouve aussi des deux mais une tendance à trouver surtout des cotes faibles, que lorsque vous interpréter ces lignes, dis-je, vous ne pouvez pas affirmer avec certitude si cela reflète une partialité fondée sur le sexe, par exemple?

Est-ce cela que vous voulez dire?

LE TÉMOIN : Oui. Je crois qu'il s'agit davantage de déterminer si les tendances que vous observez sont clairement reliées au sexe ou si elles tiennent au phénomène des cotes élevées par opposition aux cotes faibles parce que dans les données, finalement, les deux se produisent en même temps. Les hommes obtiennent surtout des cotes élevées et les femmes, des cotes faibles.

LA PRÉSIDENTE : Alors il n'est pas question de les comparer en vue de calculer un écart salarial, n'est-ce pas?

LE TÉMOIN : Je pense que c'est là une question différente que nous aborderons plus tard, je pense.

LA PRÉSIDENTE : OK. Mais si vous regardez ces résultats dans le but de décrire ce que révèle leur distribution, ce que vous pouvez dire finalement, c'est que les hommes ont tendance à avoir des cotes élevées et les femmes des cotes faibles, mais que vous ne pouvez rien dire de plus avec certitude à cause de cet effet confusionnel. C'est bien cela que...

LE TÉMOIN : C'est exact. Il faut faire très attention en interprétant les résultats parce que vous ne devez jamais oublier que quelqu'un d'autre pourrait arriver avec une explication différente des données et affirmer que ce phénomène n'a rien à voir avec le sexe, que c'est l'effet des cotes élevées/faibles, que vous avez recueilli les données d'une telle façon que la plupart des cotes élevées sont celles d'hommes et que la plupart des cotes faibles proviennent des femmes. Vous vous trouvez donc devant deux explications également valables des mêmes données.

L'attitude raisonnable à adopter dans les circonstances, c'est d'interpréter les résultats avec prudence.

210

781. M. Sunter a effectué une autre analyse en vue de la présenter en réplique. Dans cette analyse, qu'il a qualifiée d'analyse «effet valeur», il a voulu chercher à expliquer davantage la différence de traitement des questionnaires ayant reçu une cote élevée par opposition à ceux qui ont été cotés faiblement. Les deux statisticiens ont des vues opposées sur la question de savoir s'il est possible de séparer, dans l'interprétation des résultats, l'effet des valeurs de l'effet d'une partialité éventuelle fondée sur le sexe. Nous retenons la mise en garde de M. Shillington sur les risques que comporte toute tentative de départager ces effets dans les circonstances. Cependant, l'analyse de M. Sunter est utile en ce sens qu'elle démontre que les différences entre les consultants et les comités surviennent dans la fourchette supérieure de l'échelle de cotation. Puisque nous avons conclu que, dans l'optique de la question que nous avons à trancher, les statistiques apportent un éclairage accessoire plutôt que des réponses définitives, nous ne sommes pas convaincus de la nécessité, voire de la validité, de faire intervenir les autres conclusions que M. Sunter a tirées de son analyse «effet valeur». Par ailleurs, les travaux antérieurs de M. Sunter sur la recherche des différences significatives demeurent utiles pour comprendre où se manifestent les différences entre les comités et les consultants.

782. Nous examinerons maintenant la question (iv) de M. Willis. L'étude Wisner 222 a été réalisée pendant que l'étude sur la parité salariale était en cours. A ce moment-là, M. Willis n'a pas fait d'analyse en profondeur afin de déterminer les raisons des différences entre les réévaluations de la Wisner 222 et les évaluations des comités, comme il l'avait fait pour les réévaluations antérieures que les consultants avaient effectuées des postes-repères du Comité directeur. M. Willis aurait préféré passer immédiatement à la deuxième partie de son plan, qui consistait à exécuter une étude plus vaste. Il estimait que cette autre

étude était souhaitable parce que la Wisner 222 n'avait pas été concluante sur la question de la partialité fondée sur le sexe. A l'époque, M. Wisner lui-même a reconnu qu'il pouvait y avoir d'autres explications plausibles de ce qu'il a défini comme une tendance marquée par des différences d'évaluation dans la Wisner 222 (pièce PSAC-4). M. Wisner affirme en fait que les postes dans les classifications à prédominance masculine, qui comportent des tâches et responsabilités plus complexes, peuvent avoir été la cause de ce phénomène. Il affirme :

[TRADUCTION]

[...] Étant donné que cela est vrai, la tendance marquée par les différences d'évaluation que nous avons relevée aurait pu se produire si les comités avaient eu tendance à sous-estimer les postes plus complexes, par rapport à la discipline du Comité directeur telle que le consultant l'interprétait.

(pièce PSAC-4, p. 8)

783. Il y a lieu de signaler que durant les travaux du Comité directeur M. Willis a remarqué que ce dernier appliquait une discipline qu'il a qualifiée de «conservatrice». Cela est évident dans la tendance du Comité directeur à ne pas attribuer de cote au-delà d'un certain niveau.

211

Le plan d'évaluation Willis prévoyait un niveau de complexité variant de A à G pour ce qui est de la connaissance fonctionnelle de l'emploi. Selon M. Willis, le niveau élevé G est un niveau assorti de [TRADUCTION] «[...] l'exigence de posséder une expertise ou de maîtriser une sphère d'activité professionnelle» (volume 35, à la p. 4448).

784. Pendant les travaux du Comité directeur, M. Willis a jugé qu'il y avait quatre ou cinq questionnaires qu'on aurait dû coter au niveau G. Lors d'une réunion spéciale avec les évaluateurs du Comité directeur, il a encouragé ceux-ci à attribuer des cotes supérieures à F. Dans son témoignage, M. Willis a fait les observations suivantes à propos du phénomène qu'il a observé (volume 35, p. 4448, ligne 19, jusqu'à la p. 4450, ligne 24) :

[TRADUCTION]

R. Parmi ceux que le Comité directeur avait évalués. Vers la fin, d'ailleurs, j'ai eu une rencontre spéciale avec eux pour voir si nous pouvions franchir le cap du F supérieur pour passer au niveau

G. C'était un phénomène intéressant. Tous reconnaissaient le problème, mais ils ne semblaient pas pouvoir parvenir à sélectionner des postes qu'ils étaient prêts à coter au-dessus de ce niveau F.

En fait, je leur ai dit : «Choisissons-en un seulement. Je veux que les autres comités voient qu'ils ont un emploi hautement professionnel comportant une véritable expertise. Je ne veux pas qu'ils pensent qu'ils ne peuvent pas coter plus haut que F.

Alors, choisissez le poste le plus fort que vous pouvez. Voyons si nous ne pouvons pas le faire passer au niveau G.» Et ils n'ont tout simplement pas pu le faire.

Cela n'a pas été sans m'inquiéter. Mes craintes ont toutefois été réduites par deux facteurs : (1) Il y avait plusieurs postes au niveau F supérieur. Le total des cotes pour ce niveau est le même que pour le niveau [G] inférieur...

LA PRÉSIDENTE : Excusez-moi, vous disiez qu'il y avait plusieurs postes au niveau F supérieur?

LE TÉMOIN : Oui, le F tendant vers le G.

Si vous vous rappelez, dans le système d'évaluation le G «inférieur» qui tend vers le F a la même cote totale. Alors je ne m'en faisais pas du point de vue de la cote. Mais comme le Comité directeur se trouvait à établir le cadre de référence pour les autres comités, je voulais qu'il puisse coter des postes au niveau G. Cela ne s'est pas produit dans le Comité directeur.

Finalement, je me suis trouvé plus tard à conseiller les comités d'évaluation. Je leur ai expliqué le problème. Je ne me souviens plus combien de postes ils ont finalement cotés au niveau G, mais je crois comprendre qu'ils ont effectivement franchi ce seuil et qu'ils ont attribué la cote G à certains des 4 000 postes évalués.

Q. Alors cette tendance, au bout du compte, vous a-t-elle paru comme quelque chose dont il ne fallait pas s'inquiéter?

R. L'autre facteur atténuant, c'est que même s'ils étaient très conservateurs ici, leur conservatisme a été uniforme. En examinant l'alignement, je pouvais constater que l'alignement interne était toujours approprié. S'ils étaient très

conservateurs par rapport aux niveaux supérieurs de l'échelle, cela n'a pas créé, disons, une inversion dans les rapports d'évaluation.

Il y avait tellement peu de postes -- et je me souviens d'en avoir discuté avec Paul Durber après l'étude. Ils ont examiné les postes qui auraient pu être cotés à un niveau supérieur et il y en avait tellement peu qu'ils n'auraient pas influé de façon sensible sur les résultats.

785. Selon la preuve, au début du processus, plus particulièrement lorsque les consultants ont réévalué les postes du Comité directeur, les consultants évaluaient les postes différemment des comités. Cela s'est produit pour la première fois lorsque la consultante de M. Willis, Mme Drury, a examiné les évaluations du Comité directeur à la demande même de ce dernier. C'est également arrivé plus tard, lorsque M. Wisner a examiné les évaluations du Comité directeur qui avaient été contestées. La discipline de M. Wisner, a-t-on constaté, était légèrement plus libérale que celle du Comité directeur. M. Willis a témoigné à cet effet dans le volume 56, à la p. 6940, lignes 14-24 :

[TRADUCTION]

R. Alors, cela nous ramène à votre commentaire comme quoi M. Wisner était probablement plus libéral.

R. Il était un peu plus libéral, mais cela ne me dérangeait pas. J'avais mes raisons pour ne pas vouloir effectuer les évaluations moi-même ou pour les faire faire par Jan Drury, même si nous avions discuté du Comité, j'étais prêt à accepter le fait que la discipline de M. Wisner puisse être légèrement différente. Mais c'était l'aspect systématique des différences que je surveillais dans les évaluations. Donc, M. Wisner a fait le meilleur choix.

786. M. Willis était prêt à reconnaître que la discipline de Wisner puisse être légèrement différente de celle des comités. Cela ne le préoccupait pas aussi longtemps qu'il n'y avait pas de tendance systématique dans les différences.

787. Certaines évaluations étaient plus faciles que d'autres selon les informations contenues dans le questionnaire. M. Willis a témoigné que les réponses des titulaires des postes à prédominance féminine ont été reçues plus rapidement et contenaient de meilleures informations que ce n'était le cas pour les titulaires des postes occupés majoritairement par des hommes. On lui a demandé si cela pouvait influencer, de façon limitée, sur la

fiabilité des évaluations. Sa réponse figure dans le volume 68, à la p. 8575, lignes 3-13 :

213

[TRADUCTION]

Q. Mais voici où je veux en venir : cela pouvait-il influencer sur la fiabilité des évaluations en ce qui concerne, disons, les groupes professionnels? Autrement dit, obteniez-vous des informations plus fiables des groupes à prédominance féminine et des informations moins fiables des groupes à prédominance masculine?

R. Je ne l'ai pas vérifié, mais je pense que c'est possible, bien sûr, puisque les groupes à prédominance féminine ont tendance à fournir de meilleures informations.

788. M. Willis a par ailleurs fait remarquer que les questionnaires des titulaires occupant des postes techniques et professionnels de haut niveau prenaient plus de temps à rentrer que les questionnaires des commis de bureau ou des gens de métier. A cet égard, M. Willis a expliqué que, «en règle générale», les emplois professionnels et techniques sont plus difficiles à comprendre pour les évaluateurs. Voici ce qu'il dit à ce sujet (volume 69, à la p. 8582, lignes 11-20) :

[TRADUCTION]

LE TÉMOIN : Les questionnaires sur les postes professionnels et techniques sont moins faciles à comprendre que, disons, ceux des gens de métier ou des commis.

M. FRIESEN :

Q. Et cela est en partie dû au fait qu'ils n'étaient pas aussi bien décrits dans les informations.

R. En partie, et en partie aussi parce qu'il est plus difficile de comprendre un emploi plus complexe. [c'est nous qui soulignons]

789. L'opinion de M. Willis est corroborée par le témoignage d'au moins deux évaluateurs. Mme Crich, qui a été membre de la première version du comité 5, a témoigné que son comité avait de la difficulté à évaluer les questionnaires des groupes à prédominance masculine. A son avis, cela a contribué aux problèmes auxquels son comité a dû faire face. Nous avons

également le témoignage du membre Latour, aussi du comité 5, quant à la difficulté que ce comité a éprouvée à évaluer les postes techniques.

790. Dans l'ensemble, il ne doit pas être tenu compte du travail du comité sur l'AQ à la lumière des critiques que M. Willis a faites de ses travaux. Toutefois, le témoignage de deux de ses participants (Mme Crich et Yates) mérite qu'on s'y attarde étant donné qu'il illustre la difficulté qu'ont éprouvée les membres de ce comité à évaluer les 25 questionnaires des titulaires masculins provenant de l'étude Wisner 222 et leur incapacité à en arriver à un consensus dans ces cas.

791. Par contraste, les consultants n'ont pas connu la même difficulté que les membres des comités à évaluer les questionnaires les plus complexes. Les consultants bénéficiaient d'une expérience et d'une formation professionnelles en évaluation des emplois, ce qui leur a permis

214

d'évaluer ces postes plus facilement que les évaluateurs des comités. Le fait que ces derniers n'avaient pas cette expertise professionnelle a contribué, à notre avis, à l'inefficacité du processus d'évaluation des postes et aux longues discussions qui ont eu lieu au cours des évaluations.

792. M. Willis a témoigné qu'il estimait très haut la compétence et l'expérience de ses consultants en ce qui a trait à l'évaluation des emplois aux fins de la parité salariale. Il a reconnu que ceux-ci étaient plus libéraux dans leur évaluation des postes de haut niveau. Compte tenu de l'expérience, des antécédents et de la formation des consultants, il estimait par ailleurs qu'ils comprenaient probablement mieux les postes de haut niveau que les membres des comités.

793. Les exemples fournis par l'employeur ont été confirmés au cours du contre-interrogatoire que l'avocat de ce dernier a fait subir aux témoins Sunter et Shillington quant à l'effet sur l'écart salarial qu'avait pu avoir le traitement différent accordé aux questionnaires des femmes et des hommes. Si les questionnaires sont traités différemment (à cause d'une partialité fondée sur le sexe), cela aura un impact direct sur l'écart salarial. Il y a deux situations distinctes qui peuvent accroître cet écart. Si les comités sous-estiment les questionnaires relatifs aux postes à prédominance masculine, l'écart salarial peut augmenter. Il peut également s'élargir lorsque les comités surestiment les questionnaires des postes à majorité féminine. Dans l'un et l'autre cas, l'effet est le même. Exprimé autrement, l'écart salarial sera «surestimé» si l'un ou l'autre de ces événements se produisent.

794. Si l'écart de 2,3 pour cent entre les comités et les consultants est attribuable à la partialité fondée sur le sexe, alors il découle soit du fait que les évaluateurs, consciemment ou non, traitaient les postes à prédominance masculine moins favorablement que les consultants, soit du fait qu'ils surestimaient les questionnaires des postes à prédominance féminine et qu'ils avaient un parti pris contre les questionnaires des postes à prédominance masculine. Les analyses statistiques de M. Sunter n'ont pas relevé de préférence pour les questionnaires des postes à prédominance féminine de la part des comités d'évaluation multiples. Les résultats des tests sur la fiabilité inter-évaluateurs montrent que la majorité des évaluateurs qui produisaient des cotes aberrantes penchaient du côté des hommes; or si on compare les évaluations finales des comités aux évaluations des consultants, on constate que les disparités révèlent un parti pris contre les postes occupés majoritairement par des hommes.

795. En déterminant pourquoi les différences se sont produites, le Tribunal est fondé à examiner certains faits probants. Le fait le plus important, c'est que le Comité directeur a été conservateur dans sa discipline par comparaison avec les consultants. Premièrement, selon M. Willis, la discipline du Comité directeur était plus exacte que celle des consultants, comme en fait foi le rapport que M. Willis a présenté au Comité mixte (pièce R-22) sur la réévaluation des évaluations du Comité directeur qui a découlé des 103 contestations du sous-comité sur la fiabilité inter-évaluateurs et des contestations du Conseil du Trésor. Ce rapport dit notamment ce qui suit :

[TRADUCTION]

215

Nous n'avons pas de réserve importante quant à la façon dont le Comité directeur comprend et applique le plan d'évaluation. Sa façon d'appliquer le plan d'évaluation aux postes (sa «discipline») diffère à certains égards de la façon dont les consultants l'appliqueraient. Toutefois, compte tenu de la manière dont on a décidé de la composition du Comité directeur, la discipline de ce dernier constitue un reflet plus fidèle de la valeur des postes telle qu'on la conçoit en général dans la fonction publique fédérale que ce que le consultant pouvait déterminer d'un point de vue extérieur.

(pièce R-22, p. 8)

796. Deuxièmement, avant l'étude sur la parité salariale du Comité mixte les consultants de M. Willis avaient une discipline établie fondée

sur leur expérience dans d'autres études. Il y a amplement de preuves qui permettent de conclure que la discipline Willis était plus libérale que celle du Comité directeur. Selon Mme Owen, une autre consultante de M. Willis, la discipline Willis a influencé les consultants dans les évaluations qu'ils ont effectuées durant l'étude menée par le Comité mixte. Les consultants étaient des évaluateurs expérimentés et professionnels. Ils connaissaient mieux les emplois de niveau supérieur, tant chez les cadres que chez le personnel technique, expérience qu'ils avaient acquise dans d'autres études sur l'équité salariale. C'était la première fois, durant l'étude du Comité mixte, que les consultants effectuaient une évaluation au sein de la fonction publique fédérale.

797. Troisièmement, dans l'ensemble les comités d'évaluation ont suivi la discipline du Comité directeur. Il est arrivé trois ou quatre fois qu'ils attribuent des cotes au-dessus du niveau F, en l'occurrence au niveau G faible. Selon M. Willis, cela n'a eu aucun effet sur la discipline du Comité directeur.

798. Quatrièmement, les évaluateurs ayant donné des cotes aberrantes n'ont pas exercé d'influence observable sur les évaluations des comités, que ce soit au sein du Comité directeur ou dans les comités multiples d'évaluation. La preuve statistique corrobore les propres conclusions de M. Willis selon lesquelles ces évaluateurs n'ont eu aucun effet discernable sur les évaluations des autres membres des comités.

799. Enfin, tant M. Willis que les évaluateurs ont témoigné que les postes de haut niveau étaient difficiles à évaluer. La distribution des questionnaires entre les groupes professionnels à prédominance masculine et féminine n'était pas la même sur le plan des valeurs ou des cotes. Les questionnaires les plus difficiles concernaient les postes de haut niveau à prédominance masculine, où l'on a relevé le plus grand écart entre les évaluations des comités et celles des consultants.

## VIII. CONCLUSION

800. A la lumière de ces faits, ainsi que d'autres facteurs mentionnés précédemment par le Tribunal, il est raisonnable de conclure de la discipline conservatrice établie par le Comité directeur, de l'inexpérience

des évaluateurs et de la difficulté à évaluer les postes de haut niveau, compte tenu par ailleurs de la nature très subjective de l'évaluation, que la disparité entre les consultants et les comités s'explique par les facteurs que nous avons mentionnés. Nous concluons que cela a donné lieu à

un phénomène qui s'est traduit par la réticence de la part du Comité directeur à attribuer des cotes élevées aux questionnaires concernant les postes de niveau supérieur. Le fait que les informations sur les postes laissent à désirer et la difficulté à comprendre ces informations ont aussi contribué à ce phénomène. En appliquant la norme de preuve raisonnable exigée par l'article 11 de la Loi, il est raisonnable de conclure que la différence entre les comités et les consultants n'est pas attribuable à la partialité fondée sur le sexe. Nous concluons, de fait, que les disparités ont résulté d'une incapacité ou de la réticence, ou des deux, de la part des évaluateurs à évaluer les postes de niveau supérieur à prédominance masculine selon la discipline des consultants.

801. L'attitude conservatrice des évaluateurs du Comité directeur est à l'origine de ce phénomène, qui s'est répandu et qui a persisté tout au long des travaux des comités multiples. Ce conservatisme a eu l'effet le plus marquant sur les postes à prédominance masculine de niveau supérieur.

802. Dans son témoignage, M. Willis a été incapable de souscrire sans réserve aux résultats de l'étude sur la parité salariale menée par le Comité mixte. Il était d'avis, cependant, qu'il n'y avait pas lieu de les rejeter carrément. Selon lui, le Tribunal pouvait les accepter tels quels ou après leur avoir apporté certains rajustements. Certaines questions demeurent toutefois en suspens, vu les réserves de M. Willis, quant à l'efficacité avec laquelle le processus a fonctionné.

803. Jusqu'ici l'audience s'est échelonnée sur 232 jours. Le Tribunal a pu entendre un large éventail de témoignages d'experts et de profanes, y compris des preuves anecdotiques. En nous prononçant sur la question de la fiabilité, nous n'oublions pas la large part d'entente qu'il y a eu entre les consultants et les comités d'évaluation au sujet des réévaluations. La norme de preuve en l'espèce est la norme du raisonnable. Nous constatons qu'en général les comités et les consultants ont pu s'entendre sur les cotes d'évaluation, exception faite des postes professionnels et techniques plus complexes distribués dans la gamme supérieure des emplois à prédominance masculine. Le phénomène qui avait pris source au sein du Comité directeur s'est propagé aux comités d'évaluation multiples et a été entretenu par d'autres facteurs, lesquels ont contribué à la disparité entre les consultants et les comités.

804. Nous concluons de fait que la preuve établit que les résultats des évaluations sont suffisamment faibles, selon toute norme raisonnable, pour permettre de calculer l'existence ou non de disparité salariale entre les hommes et les femmes employés dans un même établissement et qui accomplissent un travail de valeur égale au sens de l'art. 11 de la Loi et de l'Ordonnance. L'employeur n'a pas réussi à produire de preuves qui

inciteraient le Tribunal à conclure autrement ou à modifier sa décision.

Fait à Vancouver, Colombie-Britannique, le 19e jour de janvier 1996.

217

Donna Gillis, présidente

Norman Fetterly, membre

Joanne Cowan-McGuigan, membre

ANNEXE A  
MANDAT DES COMITÉS

1. Sous-comité sur le plan d'évaluation commun

a) Mandat

Le mandat officiel de ce sous-comité consistait à déterminer les plans d'évaluation à examiner et à faire des recommandations à l'ensemble du Comité mixte syndical-patronal sur la mise en oeuvre de la parité salariale.

2. Comité mixte syndical-patronal sur la mise en oeuvre de la parité salariale

a) Mandat

La tâche du Comité mixte syndical-patronal sur la mise en oeuvre de la parité salariale consistait à établir des paramètres définis pour la mise en oeuvre de la parité salariale, conformément aux dispositions de l'article 11 de la Loi canadienne sur les droits de la personne, et à rédiger un plan détaillé de mise en oeuvre pour le segment de la fonction publique dont le Conseil du Trésor est l'employeur.

### 3. Sous-comité sur la stratégie de communication

#### a) Mandat

Le mandat de ce sous-comité consistait à analyser les options en matière de communication et à recommander la mise en oeuvre des plus efficaces.

A-1

### 4. Sous-comité sur la formation

#### a) Mandat

Le mandat de ce sous-comité consistait à dresser l'ébauche d'une trousse de formation pour les coordonnateurs et à faire des recommandations à ce sujet. Ce sous-comité est devenu par la suite le Sous-comité de l'administration.

### 5. Sous-comité de l'examen du plan d'évaluation Willis

#### a) Mandat

Le principal objectif de ce sous-comité consistait à présenter au Comité mixte des recommandations concernant :

- (i) la modification ou la clarification des définitions et des facteurs relatifs aux quatre tableaux du plan d'évaluation;
- (ii) le choix du tableau 1 ou du tableau 2 pour l'évaluation des conditions de travail.

### 6. Sous-comité sur le questionnaire Willis

#### a) Mandat

Le mandat de ce sous-comité consistait à mettre la dernière main à la présentation et au contenu du questionnaire Willis (y compris la mise au point d'exemples). On a demandé au sous-comité de passer en revue le questionnaire et de veiller à ce qu'il soit approprié pour recueillir les données nécessaires.

A-2

## 7. Sous-comité de l'administration

### a) Mandat

Le mandat de ce sous-comité consistait à examiner toutes les questions relatives à l'administration de l'Étude sur la parité salariale et à présenter des recommandations ou à prendre des décisions à ce sujet, sauf pour les responsabilités confiées au Secrétariat de l'étude sur la parité salariale. De façon plus particulière, ce sous-comité a :

- (i) conçu, mis en oeuvre et contrôlé toutes les mesures administratives demandées par le Comité mixte;
- (ii) fourni au SEPS les conseils nécessaires en ce qui a trait aux questions administratives;
- (iii) recommandé au Comité mixte des mesures à prendre;
- (iv) assuré le bon déroulement de l'étude, selon le cadre établi par le Comité mixte, grâce à l'établissement de priorités, à la délégation des tâches, à la résolution des problèmes et à l'évaluation des progrès réalisés;
- (v) coordonné la formation nécessaire pour les coordonnateurs, les évaluateurs, les examinateurs et les secrétaires.

A-3

## 8. Comité directeur de l'évaluation

### a) Mandat

L'objectif premier du Comité directeur de l'évaluation (CDE) consistait à évaluer un échantillon représentatif de postes, en vue de fournir un cadre d'évaluation aux cinq comités chargés de celle-ci (ce nombre est passé à neuf par la suite), de façon qu'à la fin de l'évaluation des postes dans le cadre de l'étude, on puisse établir un lien juste et équitable entre les 4 400 postes évalués. Le mandat du Comité directeur de l'évaluation consistait à :

- (i) établir des cotes repères pour environ 600 postes, par suite d'une évaluation initiale d'un nombre représentatif de postes échantillonnés, ainsi qu'un cadre de référence pour guider les comités subalternes dans le processus d'évaluation;
- (ii) fournir des conseils et de l'aide aux comités subalternes pour les cas d'évaluation particulièrement difficiles;
- (iii) mettre en oeuvre un système de contrôle pour assurer l'uniformité et l'impartialité des évaluations des comités subalternes;
- (iv) résoudre les cas controversés, en dernier recours, après qu'un comité d'évaluation ait fait tous les efforts nécessaires pour convenir d'une cote, sans succès.

A-4

## 9. Mini-comité mixte syndical-patronal

### a) Mandat

Le mandat du Mini-comité mixte syndical-patronal consistait à résoudre les problèmes de procédures découlant de l'étude. Au départ, le Comité

mixte a passé beaucoup de temps à examiner ces problèmes, mais il a finalement été décidé d'établir un comité restreint pour les résoudre.

#### 10. Secrétariat de l'étude sur la parité salariale

##### a) Mandat

Le Secrétariat de l'étude sur la parité salariale était un secrétariat mixte syndical-patronal. Il était situé à l'immeuble Jackson et fournissait tout le soutien administratif nécessaire au volet d'évaluation de l'étude. Le chef était responsable de la coordination de toutes les activités de soutien et de la transmission efficace des instructions du Comité mixte et du Sous-comité de l'administration.

#### 11. Sous-comité de fiabilité inter-évaluateurs et de méthodologie

##### a) Mandat

Le mandat de ce sous-comité consistait à :

- (i) déterminer la méthode et les recherches nécessaires pour vérifier le coefficient d'objectivité des comités d'évaluation et à faire des recommandations à ce sujet;
- (ii) évaluer la méthode de recherche, dans son application à l'ensemble de l'étude du Comité mixte, et à faire des recommandations à ce sujet.

A-5

#### 12. Cinq comités d'évaluation multiples

##### a) Mandat

Le mandat des cinq comités d'évaluation multiples consistait à:

- (i) évaluer environ 750 postes chacun;
- (ii) tenir le Comité directeur de l'évaluation au courant de leurs délibérations, des résultats obtenus et des questions soulevées, par l'entremise de leurs présidents.

Les cinq comités d'évaluation multiples ont été réorganisés en neuf comités le 14 avril 1989.

### 13. Sous-comité de fiabilité inter-comités

#### a) Mandat

Le mandat de ce sous-comité consistait à :

- (i) examiner les résultats des tests administrés aux comités d'évaluation, en fonction des critères de base fournis par les experts-conseils;
- (ii) examiner les notes de base fournies par les experts-conseils;
- (iii) déterminer les écarts significatifs entre les cotes établies par consensus par les comités et les cotes repères et critères de base;
- (iv) formuler au besoin des recommandations pour la formation et le recyclage devant être assurés par les experts-conseils, ainsi que des mesures à soumettre à l'examen du Comité mixte;

A-6

- (v) déterminer les problèmes touchant les procédures et les méthodes ainsi que les possibilités d'amélioration, y compris la révision de la formulation des justifications.

### 14. Mini-comité directeur de l'évaluation

#### a) Mandat

Le Mini-comité directeur de l'évaluation était chargé de passer en revue les contestations du Comité mixte concernant les évaluations effectuées par le CDE. Le Comité mixte a demandé à Johanne Labine, de l'AFPC, et à Michel Cloutier, du Conseil du Trésor, siégeant tous les deux au CDE, de passer en revue les conditions de travail des 100 postes repères, du point de vue du travail par poste, des heures supplémentaires et des modalités de travail, d'évaluer le nombre de points à modifier, le cas échéant, et de corriger au besoin les justifications.

Il a été décidé, en dernière analyse, que le CDE ne tiendrait pas d'autre réunion et qu'un mini-comité, à savoir un noyau ou un petit nombre d'évaluateurs du CDE, reprendrait cet exercice. Deux membres du

CDE ont été choisis pour représenter ce mini-comité, à savoir Michel Cloutier et Johanne Labine. On avait l'intention que M. Willis rencontre ces deux personnes en vue de résoudre avec elles les différends.

#### 15. Sous-comité sur l'indemnité totale

##### a) Mandat

L'ébauche du mandat de ce sous-comité, au 21 septembre 1989, consistait à :

- (i) déterminer les éléments de la rémunération au gouvernement fédéral qui correspondent à la définition de salaire du paragraphe 11(6) de la Loi canadienne sur les droits de la personne;

A-7

- (ii) compiler les données nécessaires pour fixer le salaire des postes évalués;
- (iii) concevoir une méthode pour établir le coût de la rémunération totale, en vue de corriger les disparités salariales décelées.

#### 16. Comité sur l'analyse de la qualité

##### a) Mandat

Paul Durber de la Commission a créé le Comité sur l'analyse de la qualité pour examiner les 25 emplois à prédominance masculine dont on pensait qu'ils avaient été sous-évalués dans le rapport Wisner, en mai 1990. Le Comité devait déterminer si le fait que ces postes soient à prédominance masculine avait joué un rôle pour l'attribution des cotes, ou encore si les écarts entre l'évaluation de M. Wisner et celle des comités étaient simplement dus à une perception différente du travail.

