

TROISIÈME PARTIE

**Les comparaisons
entre les groupes**

Vous déambulez dans une grande ville américaine et vous croisez successivement un Blanc, un Noir, un Hispanophone et un Asiatique. Que pouvez-vous dire à propos de leurs QI respectifs? Le lendemain, vous apprenez d'un spécialiste de l'intelligence que le QI moyen des Blancs vivant aux États-Unis est de 100, celui des Noirs, de 85, celui des Asiatiques, de 110 et celui des Hispanophones, de 93. Que pouvez-vous dire ensuite à propos du QI des quatre citoyens rencontrés la veille? Tout simplement rien parce que vous ne disposez d'aucune mesure.

Parmi les controverses entourant l'intelligence, peu ont déchaîné autant de passions que celles relatives aux comparaisons entre les groupes ethniques (chapitre 10) et entre les sexes (chapitre 9). Certains vont même jusqu'à remettre en question la pertinence de recherches sur le sujet, prétextant que des chercheurs seraient trop heureux d'exploiter d'éventuels résultats pour conforter leurs positions racistes, sexistes ou politiques. L'exemple classique, sans cesse évoqué, concerne les comparaisons au plan du QI entre les groupes ethniques pour justifier la discrimination raciale.

Si l'exploitation pernicieuse des données de la recherche reste toujours possible, les progrès de la connaissance servent le plus souvent d'antidote aux idées erronées, aux positions doctrinaires et aux raisonnements fallacieux. Au strict plan éthique, plusieurs raisons d'ordre légal, politique et social justifient les recherches sur l'intelligence. L'un des objectifs de nombreuses sociétés n'est-il pas de fournir à chaque enfant l'éducation qui lui assure l'avenir le plus prometteur? Cet objectif est guidé par le principe de l'égalité pour tous, indépendamment du sexe, de l'origine ethnique ou de la provenance sociale. Or, dans la mesure où de telles recherches permettent de déceler des facteurs défavorables à l'un ou l'autre groupe d'enfants, on pourra d'autant mieux élaborer des programmes destinés à corriger cette situation. Il restera bien sûr à évaluer l'efficacité de tels programmes.

À propos des comparaisons sur le plan cognitif, l'une des erreurs les plus courantes consiste à confondre le score moyen d'un groupe et le score d'un individu appartenant à ce groupe. L'exemple suivant devrait aider à bien faire comprendre cette distinction. Un professeur informe ses élèves que la moyenne du groupe au dernier examen est de 73 %. La probabilité que tous les élèves aient 73 %, même si elle n'est pas théoriquement impossible, est dans les faits quasi nulle. Par contre, que les scores varient entre 40 % et 100 % est tout à fait probable. Appliqué aux différences entre les sexes, cet exemple signifie que même si, comme nous le verrons plus loin, des travaux montrent que les hommes présentent en moyenne de meilleures habiletés visuospatiales que les femmes, ce constat n'autorise d'aucune façon à conclure que tous les hommes ont de meilleures habiletés visuospatiales que les femmes. Bref, on ne peut strictement rien dire à propos de ce genre d'habiletés chez telle femme ou tel homme croisé dans la rue à moins de disposer d'une démarche d'évaluation formelle.

Appliquer une moyenne de groupe à l'un de ses membres constitue non seulement une erreur de raisonnement, mais aussi une faute méthodologique, voire un accroc au fonctionnement démocratique. En somme, on ne peut rien dire sur les capacités de qui que ce soit sans les avoir préalablement mesurées. Qui plus est, comme le montre la figure 3.3, la variance attribuable aux différences individuelles de QI excède, et de loin, la variance liée à l'ensemble des membres d'un groupe. Autrement dit, il existe un large recouvrement dans les résultats entre deux groupes, attribuable précisément aux différences plus importantes entre les individus à l'intérieur d'un groupe donné qu'entre les groupes. Bref, si jamais les différences entre les groupes disparaissaient, les différences entre les individus persisteraient.

L'erreur de jugement évoquée ne relève pas seulement du raisonnement. La nature humaine étant ce qu'elle est, classer procure toutes sortes d'avantages. Au plan émotif, c'est rassurant; au plan cognitif, c'est structurant et efficace; au plan social, le classement peut contribuer au sentiment d'appartenance et au sentiment d'identité.

On rencontre une autre erreur de raisonnement dans la population et, parfois, chez des chercheurs en sciences humaines et sociales. Cette erreur consiste à s'élever contre l'idée que les comportements complexes puissent être déterminés génétiquement dans l'espèce humaine, alors qu'ils l'admettent volontiers pour les autres espèces. Par exemple, on reconnaîtra volontiers que les *border collies* – ces chiens apprennent plus vite que les autres chiens – sont plus intelligents que les lévriers afghans en raison de différences d'ordre génétique. Refuser que des différences intellectuelles entre des groupes humains puissent être en partie génétiquement déterminées, c'est oublier que, à l'instar des chiens, l'environnement en général et l'apprentissage en particulier viennent moduler les différences génétiques (chapitre 3 – «L'analyse génétique en psychologie»).

Les comparaisons entre les groupes en matière intellectuelle soulèvent en outre un problème sémantique. Par exemple, comme *tous* les hommes ne diffèrent pas de *toutes* les femmes, ni *tous* les Noirs de *tous* les Blancs, comment dès lors décrire simultanément les ressemblances et les différences entre les groupes étudiés? Pour éviter d'attribuer à tous les membres d'un groupe les caractéristiques d'un pourcentage des membres de ce groupe, il vaut mieux s'en remettre à des fréquences relatives. Ainsi, il serait plus exact de dire «il y a plus de garçons que de filles qui sont dyslexiques» que de dire «les filles lisent mieux que les garçons». En ce qui concerne les comparaisons entre les sexes, les précautions sémantiques sont d'autant plus nécessaires que le facteur comparatif est biologique et donc incontournable, ce qui pourrait conduire à la conclusion que les différences observées sont immuables.

Enfin, les étudiants se disent souvent surpris des résultats de recherche contradictoires. Malgré les efforts d'objectivité dont font preuve la majorité des chercheurs, les recensions traditionnelles des écrits ne permettent pas toujours de trancher de façon satisfaisante, compte tenu de l'espace laissé

aux choix arbitraires et aux impressions personnelles. Pour contrer cet éventuel subjectivisme, une procédure quantitative a été mise au point vers la fin des années 1970, la méta-analyse (encadré III.1)

Encadré III.1 – La méta-analyse

La méta-analyse permet de tenir compte d'un grand nombre d'études publiées ou non. Ramenés à une base commune, les résultats de ces études peuvent alors être traités comme s'ils provenaient d'un large échantillon. Les études utilisées dans une méta-analyse doivent minimalement comprendre des statistiques descriptives (N , X , S^2) ou inférentielles (t , X^2 , F , r). Les études de cas, les études cliniques ou toute étude dépourvue de données quantitatives sont exclues, bien qu'elles puissent éventuellement servir à la discussion. Cette procédure dégage une moyenne des résultats obtenus pour la comparaison de deux groupes ou encore en établissant une corrélation moyenne appelée estimé pondéré de l'ampleur ou de la taille de l'effet (en anglais, *effect size*). La mesure la plus commune de l'ampleur de l'effet lors de l'estimation de différences de groupes, d , est égale à la moyenne des scores obtenus par les sujets d'un groupe, moins la moyenne des scores obtenus par les sujets de l'autre groupe, divisée par la somme des écarts types des recherches concernées. Le résultat alors obtenu mesure l'écart entre les deux groupes et s'exprime en unités d'écart type. Un d de zéro indique une absence totale de différence entre les deux groupes, un d négatif ou positif indique la présence d'une différence entre les deux groupes comparés. Habituellement un d de 0,20 est considéré faible, un d de 0,50, moyen et un d de 0,80 et plus, important. Quelle que soit l'ampleur de la différence, il existe quand même un large recouvrement entre les deux groupes : un d faible, moyen ou important implique respectivement un recouvrement de 85%, 67% et 53% entre les deux courbes (figures III.1, III.2 et III.3).

FIGURE III.1

Distributions normales théoriques lorsque la différence entre deux groupes correspond à un $d = 0,20$

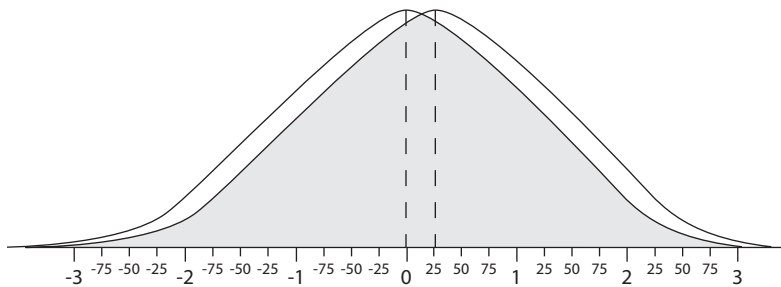


FIGURE III.2

Distributions normales théoriques lorsque la différence entre deux groupes correspond à un $d = 0,50$

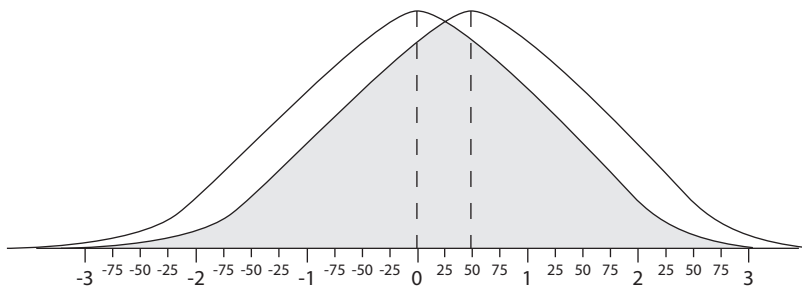
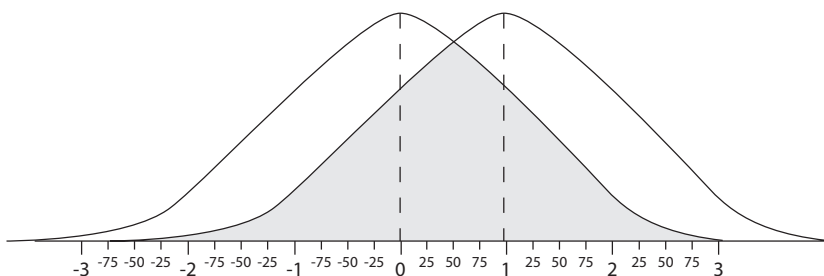


FIGURE III.3

Distributions normales théoriques lorsque la différence entre deux groupes correspond à un $d = 1,00$



Sans nier les écueils de cette procédure [1], on peut dire qu'elle permet de dresser le bilan des connaissances dans un champ d'études donné et peut déboucher sur de nouvelles perspectives de recherche [6]. Certains [2, 9] minimisent par exemple l'importance des différences entre les hommes et les femmes relativement à certaines habiletés spécifiques en invoquant l'indéniable recouvrement entre les performances des deux groupes. Ils oublient alors que dans d'autres domaines, moins empreints d'idéologie, le même type de résultats consacre le phénomène étudié comme un fait établi. Ils oublient en outre que, même si les différences observées n'ont pratiquement aucune conséquence sur le plan individuel, elles peuvent quand même être significatives au plan social vu l'ampleur des échantillons concernés [des informations supplémentaires sont présentées aux références 3, 4, 5, 7, 8].

Cette troisième partie comprend deux chapitres; le chapitre 9 est consacré aux comparaisons entre les sexes et le chapitre 10 traite des comparaisons entre les groupes ethniques, dont l'homogénéité n'est pas aussi évidente que dans le cas des sexes. Dans la mesure du possible, nous présenterons des résultats de méta-analyses plutôt que des résultats découlant de recensions traditionnelles des écrits.

Références

1. Eysenck, H.J. (1995). Meta-analysis squared: Does it make sense? *American Psychologist*, 50, 110-111.
2. Favreau, O.E. (1993). Do the Ns justify the means? Null hypothesis testing applied to sex and other differences. *Canadian Psychology*, 34, 64-78.
3. Fortin, F., Lévesque, J., et Vitaro, F. (2007). La méta-analyse au service de la prévention et de l'intervention. *Revue de psychoéducation*, 36 (1), 167-193.
4. Hall, J.A., et Rosenthal, R. (1991). Testing for moderator variables in meta-analysis: Issues and methods. *Communication Monographs*, 58, 437-448.
5. Hedges, L.V., et Pigott, T.D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, 9, 426-445.
6. Hunter, J.E., et Schmidt, F.L. (1990). *Methods of meta-analysis*. Newbury Park, CA: Sage.
7. Mullen, B. (1989). *Advanced basic meta-analysis*. Hillsdale, NJ: Lawrence Erlbaum.
8. Mullen, B., et Rosenthal, R. (1985). *Basic meta-analysis: Procedures and programs*. Hillsdale, NJ: Lawrence Erlbaum.
9. Sherman, J. (1978). *Sex-related cognitive differences*. Springfield, IL: Charles C. Thomas.