Mesures psychologiques[1,2]

J.P. Guilford

Traduction par *Anne Matoul*

1. Introduction

Les progrès et la maturité d'une science s'évaluent souvent à l'aune de son affinité avec l'usage des mathématiques[3]. Les « méthodes psychométriques » sont en fait des procédures de mesures psychologiques. Par « mesures », on entend la description de données en termes numériques, qui, à leur tour, permettent d'exploiter les nombreux avantages fournis par les opérations sur des nombres et par le raisonnement mathématique.

Les mathématiques ne constituent pas en elles-mêmes une science empirique ; elles ne recueillent aucune donnée factuelle de l'observation de la nature. Par contre, elles représentent un langage universel que toute science ou technologie peut utiliser à son meilleur avantage et pour sa plus grande facilité. Leur vocabulaire est illimité, tout en étant d'une rigoureuse exactitude. Les règles de leurs opérations – leur « syntaxe » – sont d'une précision logique sans pareille.

Certaines des conclusions les plus évidentes amenées par les mesures ne sont guère difficiles à voir ou à admettre. Les mesures fournissent en effet des descriptions précises, objectives, aisément communicables et manipulables par le raisonnement. La précision est à la mesure de la minutie observée et de l'instrument utilisé par l'observateur. L'objectivité constitue l'un des objectifs majeurs de la science. Conformément à une définition simple et opérationnelle, le terme « objectivité » désigne un consensus interpersonnel. Là où plusieurs personnes parviennent à un accord concernant des observations et leurs conclusions, les descriptions de la nature sont probablement mieux à même d'éviter des biais individuels. Elles constituent alors un corpus de connaissances considérées comme « vraies ». Par ailleurs, ces descriptions revêtent une forme qui les rend aisément communicables aux autres. La science est une institution sociale. Le fait que les observations et les conclusions émises par quelques individus puissent être transmises à d'autres dans toute leur intégralité et leur intelligibilité constitue l'une de ses valeurs-clé. Dans cette optique, les descriptions quantitatives, ou « mesures », ne font pas exception à la règle, loin s'en faut.

---

[1] Ce chapitre est une traduction du premier chapitre intitulé « Psychological Measurement » de l'ouvrage de J.P. Guilford (1954). Psychometric Methods. New York : MacGraw-Hill.
Les encarts et les notes précédées des lettres MD sont de Marc Demeuse, les notes précédées de la mention de NDT sont de la traductrice. La bibliographie et les références bibliographiques ont été harmonisées et actualisées.

[2] Pour aller plus loin, dans le domaine de la littérature francophone, on peut consulter le chapitre que Jean-Claude Falmagne consacre à *Mesurage, modèles mathématiques et psychophysique* (pp. 79-139) dans le *Traité de psychologie expérimentale* de M. Richelle, J. Requin et M. Robert (1994) paru aux Presses universitaires de France. Il dépasse cependant de loin le niveau de ce chapitre introductif.

[3] MD : Comme le signale Levy (1975, p. 46), on trouve chez Adolphe Quételet, dans son ouvrage intitulé *Théorie des probabilités* (p. 102), une assertion analogue : *Tout ce qui peut être exprimé numériquement est de son ressort [du calcul des probabilités] : plus les sciences se perfectionnent, plus elles tendent à rentrer dans son domaine, qui est une espèce de centre vers lequel elles viennent converger. On peut même juger du degré de perfection auquel une science est parvenue par la facilité plus ou moins grande avec laquelle elle se laisse aborder par le calcul et ainsi même se justifie le proverbe ancien, placé en tête de cet ouvrage : « mundum numeri regunt ».*
Dans le domaine de la psychologie, au *Whatever exists at all, exists in some amount* de E.L. Thorndike (1918, p. 16), répond le *Anything that exists in amount can be measured* de W.A. McCall (1922, p. 15).

Ce chapitre est en grande partie consacré à la philosophie qui sous-tend la prise de mesures en psychologie. En physique, les mesures constituent une étape tellement évidente qu'elles n'ont guère donné lieu à une grande réflexion sur leur nature réelle. Il se trouve quelques physiciens pour prétendre que ce que l'on appelle mesures en psychologie ne l'est pas du tout. Il est exact que le terme *mesures* est quelques fois défini de façon telle qu'il ne s'applique pas à la plupart des opérations réalisées en psychologie et désignées comme telles. La définition d'un terme aussi abstrait que *mesures* relève toutefois de l'arbitraire, et les psychologues pourraient lui attribuer celle qui englobe les opérations qu'ils réalisent dans ce concept, ou inventer un nouveau terme pour désigner ce qu'ils effectuent en manipulant des nombres et des opérations arithmétiques. Fort heureusement, la définition du terme *mesures* peut être suffisamment étendue pour englober les opérations désignées par l'expression « mesures psychologiques ».

Il est exact que les difficultés qui ont jalonné le développement des mesures en psychologie ont été très importantes. Mais, comme le dit Spearman, « le chemin de la science est pavé de réussites là où on ne prévoyait que l'irréalisable. A vrai dire, le traitement mathématique est peut-être l'aspect où la psychologie a réalisé ses avancées les plus solides et les plus surprenantes » (Spearman, 1937, p. 89). Si on ne peut réaliser en psychologie des opérations de mesures conformes au modèle établi en physique, il faut alors chercher à développer ses propres techniques de mesure, adaptées spécifiquement à sa problématique, et c'est ce qui se passe. Il existe une grande variété de ces techniques, nous le verrons dans les chapitres suivants. La psychologie est toutefois obligée de justifier l'application des nombres et, globalement, des mathématiques, à ses données. Il s'agit là d'un problème de logique que nous allons analyser dans le présent chapitre, mais auparavant, examinons brièvement l'histoire des mesures en psychologie, afin d'avoir une meilleure perspective du problème.

2. Les mesures du mental

Les premières étapes de l'histoire des mesures du mental sont marquées par le développement de deux courants relativement indépendants. D'un côté, la tradition psychophysique, précurseur de la psychologie expérimentale proprement dite. De l'autre, la tradition des tests mentaux, dont le centre d'intérêt se focalise sur les différences individuelles. Le premier courant s'est fondé sur la physiologie expérimentale et les méthodes quantitatives qui se sont multipliées conjointement à l'essor des sciences naturelles. La psychophysique s'est inspirée de l'ingéniosité de ceux qui avaient déjà appliqué des méthodes quantitatives et statistiques semblables en astronomie et en physique. La tradition des tests mentaux, inspirée par la biologie en pleine évolution du XIX^e^ siècle et par son intérêt pour les traits héréditaires, s'inspirait directement de ces mathématiciens qui s'adonnaient plus ou moins sérieusement aux problèmes de probabilité et aux méthodes statistiques, qui n'en étaient alors qu'à leurs débuts. Ces deux courants se sont donc largement inspirés, directement ou indirectement, d'une source commune, les probabilités mathématiques. Voyons donc le tableau synoptique du développement des méthodes statistiques, du point de vue qui les apparente aux traditions de la psychophysique et des tests mentaux.

2.1. Les origines des méthodes statistiques

On peut affirmer qu'avant l'an 1600, aucune conception mathématique des probabilités n'était reconnue. Les joueurs avaient élaboré beaucoup de spéculations concernant les jeux de hasard, lorsqu'il devenait temps pour eux de considérer leurs pertes et leurs gains, et surtout leurs pertes. Ils avaient même tenté d'intéresser les mathématiciens à leurs problèmes, sans guère de succès. Les mathématiciens étaient bien trop occupés par les domaines alors récemment découverts de la géométrie analytique et de ses calculs pour se soucier de problèmes de jeux de hasard. Toutefois, le XVII[e] siècle vit apparaître les débuts d'un intérêt réel pour les probabilités. **Bernoulli** (1645-1705) publia le premier ouvrage célèbre entièrement consacré à ce sujet. On peut attribuer à **De Moivre** (1667-1754) la découverte de la courbe de répartition normale aux environs de l'an 1733. C'est à partir de ce moment qu'émergea l'intérêt des astronomes et des mathématiciens pour ce domaine. En 1812, **Laplace** (1749-1827) terminait la rédaction de ce qui est considéré comme l'ouvrage le plus important consacré exclusivement à ce sujet. Il y présentait la démonstration de la **méthode des moindres carrés**[4]. Ce fut **Gauss** (1777-1855) qui fit la preuve du grand intérêt pratique de la courbe normale, en démontrant comment elle s'applique à la répartition des mesures et aux erreurs survenues lors d'observations scientifiques. Ce fut Gauss encore qui conçut les modes fondamentaux des calculs de moyennes, des probabilités d'erreurs, etc. Aujourd'hui, il n'est pas rare de voir la *courbe normale* appelée *courbe gaussienne.*

C'est à **Quételet** (1796-1874), astronome personnel du souverain belge, que revient le mérite d'avoir appliqué le premier la courbe normale et les méthodes statistiques élémentaires à des données biologiques et sociales. Quételet devint le grand initiateur de la méthode statistique en Europe. Il encouragea l'enregistrement de données météorologiques, et celui d'événements sociaux comme les naissances, les décès, les mariages, les maladies et les crimes. Il démontra que la loi normale de la répartition s'applique à divers types de mesures anthropométriques lorsque l'on utilise des populations non sélectionnées. Il était tellement troublé par la répartition normale des populations qu'il a la réputation d'avoir prétendu que la nature visait un homme moyen idéal, *l'homme moyen*[5] , mais qu'elle aurait manqué son objectif et aurait donc créé des déviations de part et d'autre de la moyenne (Boring, 1950, p. 477).

Encart 1 – *Adolphe Quételet.*

Lambert Adolphe Jacques Quételet est né à Gand le 22 février 1796. Il est nommé professeur au collège d'Audenarde en 1813, après avoir terminé brillamment ses études au lycée de Gand où il deviendra professeur de mathématique en 1815. Il est reçu docteur en mathématiques en 1819 et nommé professeur à l'Athénée de Bruxelles la même année. Il devient membre de l'Académie en 1820. On peut mettre à son actif la création de l'observatoire de Bruxelles (1832) et des travaux importants dans le domaine des probabilités et de la statistique. Il préside notamment la commission centrale de statistique, créée à son initiative en 1841. A ce titre, c'est lui qui met en œuvre le premier recensement général dans la jeune Belgique, en 1846. Il préside également le premier Congrès international de statistique à Bruxelles (1853), comme ceux de Vienne, Londres, Berlin, Florence, La Haye et Saint-Pétersbourg, entre 1857 et 1872.

Il s'intéresse tant à la météorologie qu'à la « physique sociale » (c'est notamment le titre de l'un de ses plus célèbres textes). Cette physique sociale sera, par la suite, mieux connue sous le nom de sociologie et popularisée par Auguste Comte.

L'un des sujets de prédilection de Quételet est l'étude de « l'homme moyen ». Le site de l'Institut national de statistique[6] précise comment cette idée émerge chez Quételet : *En relevant les mensurations de conscrits français et en analysant celles de 5.000 soldats écossais reprises en 1817 dans la revue Edinburgh Medical Journal, Adolphe Quételet applique les lois des probabilités aux données biométriques de l'homme, comme le poids, la taille, le périmètre thoracique, devenant ainsi un des fondateurs de l'anthropométrie et de la biostatistique. Basant son analyse sur les régularités temporelles et les distributions de forme gaussienne, il constate que ces données oscillent autour de valeurs moyennes et que celles-ci tendent à être constantes. Il met ainsi en avant une distribution de celles-ci selon un modèle, la "courbe des possibilités", qui sera nommé en 1894 par Pearson "loi normale".*

Suite à sa correspondance avec d'illustres mathématiciens comme Pierre-Simon Laplace, Joseph Fourrier ou Denis Poisson (qu'il avait connus à l'Observatoire de Paris), Quételet devient le premier à utiliser la courbe normale autrement que pour la répartition d'erreurs. Il étendra ensuite ces notions à l'ensemble des caractéristiques physiques en créant la notion "d'homme moyen". Élaborant ainsi une mécanique sociale qu'il présentera en 1835, dans son ouvrage intitulé Sur l'homme et le développement de ses facultés ; Essai d'une physique sociale, comme l'étude des lois qui régissent l'homme du point de vue physique, intellectuel et moral. Dans cet ouvrage, qui sera traduit en anglais dès 1842, la notion "d'homme moyen" est perçue comme la valeur centrale autour de laquelle les mesures d'une caractéristique humaine se groupent suivant une

4 *La méthode des moindres carrés consiste à minimiser la somme des carrés des écarts entre les valeurs observées de la variable dépendante du modèle et les valeurs estimées par le modèle.* (Dodge, 1993, p. 239).

5 NDT : en français dans le texte.

6 http://www.statbel.fgov.be/info/quetelet_fr.asp

courbe normale. L'homme moyen d'une population est, selon Quételet, un individu dont les caractéristiques physiologiques sont chacune égale à la moyenne des caractéristiques physiologiques des autres individus de la population.

Mais, comme le précise Dupréel (1942 , p. 44), qui consacre un ouvrage à son illustre concitoyen, *il arrive fréquemment que l'on rencontre chez certaines intelligences supérieures, chez les esprits les plus solidement équilibrés et très amplement informés, un engouement obstiné pour quelque idée manifestement boiteuse ou inadmissible. En réalité, cet accident est universel ; il n'est peut-être personne qui ne traîne dans son bagage intellectuel quelque opinion dont la faiblesse éclate aux yeux des autres, et qui ne s'explique que par un ensemble de circonstances tout particulier. On ne remarque cette bizarrerie que lorsqu'elle se rencontre dans un esprit de haute qualité, à cause du contraste et du caractère paradoxal de la chose. De cette bizarrerie on trouverait difficilement un plus beau cas que l'attachement de Quételet pour sa théorie de l'homme moyen.*

« La nécessité de faire abstraction des individus pour ne m'occuper que de ce qui se rapporte aux masses m'a conduit, comme je l'ai dit, à admettre l'homme moyen, qui, par rapport au système social, peut être considéré comme l'analogue du centre de gravité dans les corps. Cette distinction, que j'ai établie dans mon premier Mémoire sur la croissance de l'homme et que j'ai développée ensuite dans mes recherches. » Sur le penchant au crime, « ne serait que d'un intérêt médiocre, si je n'avais reconnu que l'homme moyen jouit aussi de propriétés si remarquables par rapport au système social, qu'elles ne semblent ouvrir un vaste champ à un nouvel ordre de recherches » (Physique sociale, tome II, p. 7)

Quételet meurt en 1874 à Bruxelles.

2.2. La tradition des tests mentaux

Le maillon essentiel entre Quételet et la psychologie est **Sir Francis Galton** (1822-1911). Celui-ci, fervent admirateur de l'œuvre de Quételet et poussé par l'ambition de résoudre les problèmes de l'hérédité humaine, entreprit de mesurer les êtres humains sur une grande échelle. Son laboratoire anthropométrique, qu'il installa en 1882 à South Kensington, était équipé pour effectuer divers tests moteurs et sensoriels simples. Estimant que la courbe normale et ses applications les plus simples ne convenaient pas à ses recherches, il inventa plusieurs outils statistiques, dont la méthode de corrélation (avec l'aide de **Karl Pearson**), l'utilisation de notes standardisées, la médiane, et d'autres méthodes d'échelonnage psychologique, comme la méthode des échelles d'appréciation (« rating-scale ») (Galton, 1908). D'autres acteurs de l'histoire des tests mentaux sont davantage célèbres. **Karl Pearson** et **R.A. Fischer** sont ceux dont la contribution a été la plus importante dans le domaine des statistiques. Pour ce qui est de la conception de tests, on retiendra surtout les noms de **Cattell**, **Binet**, **Terman**, **Otis**, **Thorndike**, **Spearman** et **Thurstone**.

2.3. La tradition psychophysique

Les fondements de la psychophysique avaient été jetés bien avant la venue de Galton et de ses tests psychologiques simples. La notion de différences cruciales relevées lors d'expériences sensorielles était déjà largement répandue depuis des siècles. Le principe d'un seuil absolu, ou d'une limite inférieure de la sensation, avait déjà été évoqué par **Herbart**[7] (1776-1841), bien des années avant l'introduction par **Fechner** d'une science psychophysique. **Weber** (1795-1878), quant à lui, avait déjà émis son concept des seuils différentiels de discrimination (*jnd – just noticeable differences*- différences juste perceptibles) et sa loi selon laquelle la croissance différentielle d'une sensation est proportionnelle à l'intensité du stimulus.

Fechner considérait la **loi de Weber** comme le fondement de sa fameuse relation psychophysique, selon laquelle l'intensité du processus sensoriel est proportionnelle au logarithme du stimulus. Lorsque germa en lui cette idée de la loi logarithmique, Fechner initia un programme de recherche sur une échelle réellement remarquable. Il définit la

[7] A noter que Herbart a très astucieusement appliqué quelques concepts de la logique mathématique aux processus mentaux. Il a même proposé certaines équations rationnelles dont l'objectif était d'exprimer les forces d'attraction et de rejet existant entre les « idées ».

psychophysique comme « une science exacte des relations fonctionnelles de la dépendance entre le corps et l'esprit » (Fechner, 1889, p. 8). Cette conception était suffisamment étendue pour inclure, non seulement les mesures des magnitudes sensorielles, mais aussi la quantification des perceptions, des émotions, des actions, de l'attention, en fait de tout processus psychologique pouvant être corrélé aux stimuli. Cette approche avant-gardiste de Fechner se limita essentiellement à l'étude des sensations, même s'il accomplit quelques incursions dans le domaine de la perception esthétique, ce qui fit de lui le fondateur de l'esthétique expérimentale.

Le plus important à nos yeux réside ici dans le fait que, afin de pouvoir réaliser les mesures nécessaires dans le cadre de cette nouvelle science « exacte » qu'était la psychophysique, Fechner fut amené à adapter certaines méthodes déjà en vigueur et à en inventer d'autres. C'est à lui que l'on doit d'avoir posé les fondements de toutes les méthodes psychophysiques traditionnelles, comme la ***méthode de l'erreur moyenne***, la ***méthode des changements minimaux***, et la ***méthode constante*** (Cf. Section relative à la psychophysique).

Parmi les opposants les plus importants de la psychophysique selon Fechner, à avoir aussi apporté de larges contributions théoriques et méthodologiques, citons **Delboeuf** (encart 2), **Wundt** et **G. E. Müller**. Après ceux-là sont venus **Urban**, **Culler** et **Thurstone**, qui ont aussi contribué à la théorie et aux méthodes de psychophysique. Nous verrons plus loin que la ***loi du jugement comparatif*** de Thurstone a considérablement accru le champ des opérations psychophysiques et la compréhension de sa problématique.

Encart 2 – *Joseph Delboeuf.*

Joseph Rémi Léopold Delboeuf naît à Liège le 1er octobre 1831 dans un famille modeste d'artisan. Il fréquente l'école assez tardivement (9 ans) et perd son père au milieu de ses études secondaires. Il parvient à poursuivre ses études grâce à des travaux de répétiteur. Il décroche son doctorat en philosophie et lettres en 1855 et son doctorat en sciences physiques et mathématiques, trois ans plus tard, à l'Université de Liège. Il poursuit sa formation à l'Université de Bonn et publie en 1860 une critique des postulats de la géométrie euclidienne. Il enseigne à l'Université de Gand de 1863 à 1866, où ses recherches sont principalement orientées vers la psychologie expérimentale et l'étude des illusions d'optique (dont une porte encore son nom). Il y travaille notamment avec un autre belge célèbre, Joseph Plateau, à qui ont doit notamment la découverte de la persistance rétinienne. Ses recherches consistent autant en expérimentations qu'en travaux théoriques, par exemple à propos de la loi de Fechner.

Ses travaux dans le domaine de la perception seront cependant publiés plus tardivement (à partir de 1873), notamment parce que Delboeuf revient à l'Université de Liège en 1866 pour y enseigner la philologie (Nicolas, 1995 ; Nicolas, Murray & Farahmand, 1997). Tout en publiant dans le domaine de la psychophysique, Delboeuf va s'intéresser à l'hypnose et publier de nombreux textes à ce sujet à partir de 1885. Il y consacre plusieurs voyages, à Paris, chez Charcot (à la Salpetrière), et à Nancy, chez Bernheim (Duyckaerts, 1992 ; Gauld, 1997). Il interviendra notamment à la Chambre des Représentants de manière à conserver aux non-médecins le droit de pratiquer l'hypnose. Il garde toujours, dans ce domaine très particulier, une approche expérimentale et critique, et considère par exemple l'hypnose à travers les apports que celle-ci peut produire dans le domaine des phénomènes mnésiques.

Il meurt à Bonn le 13 août 1896.

2.4. Tests mentaux et psychologie expérimentale

Ainsi que nous le verrons plus loin, le lien entre psychophysique et tests mentaux est très direct et très étroit. Il est curieux de constater que, même si la psychophysique fut dès le départ étroitement associée à la psychologie expérimentale, et même si le chercheur en psychologie expérimentale a largement utilisé les méthodes de tests psychophysiques en laboratoire, ce n'est pas à lui que revient le mérite d'avoir jeté un pont entre les deux notions. La raison en est que le chercheur en psychologie expérimentale a mis beaucoup de temps à comprendre qu'il utilisait des tests mentaux comme instruments de mesure. Il a longtemps associé tests mentaux et différences individuelles, sans voir que ces tests mesurent également des « différences occasionnelles » chez le même individu.

Bien que bon nombre de mesures effectuées en psychologie expérimentale l'aient été en termes de variables physiques – propriétés physiques du stimulus et de la réponse, dans le système familier des centimètres/grammes/secondes –, à commencer surtout chez Ebbinghaus, les mesures prises par les psychologues expérimentaux faisaient la part belle aux scores obtenus aux tests. Les nombreuses expériences menées dans les domaines de l'apprentissage, la mémoire, la motivation et la pensée fournissent généralement des données sous forme de nombres de bonnes réponses, de nombres d'erreurs, de nombres d'items sur une liste, de nombres de seuils dépassés, etc. Même si la psychologie expérimentale a exploité les tout derniers tests statistiques notoires en élaborant ses conclusions à partir de leurs données, ses chercheurs ont été lents à tirer profit du raisonnement mathématique et statistique très approfondi dispensé par ces tests en l'utilisant comme fondement à des mesures de tests mentaux. Des mesures en termes de temps, de distance ou d'énergie peuvent s'avérer très raffinées du point de vue de la technique expérimentale. On peut toutefois s'interroger sur le bien-fondé de telles échelles pour représenter des phénomènes psychologiques. La variable de la « force de l'habitude », par exemple, est une variable psychologique tout à fait admise, quel que soit le nom ou la définition qu'on lui attribue. Ses différents indices, même ceux exprimés sous forme d'unités physiques, donnent lieu à diverses formes de courbes d'apprentissage. Celles-ci ne peuvent être toutes reliées à la force de l'habitude car elles ne sont pas linéairement reliées entre elles.

2.5. Les échelles psychologiques

Les méthodes d'échelonnement psychologique, comme la ***méthode de comparaison par paires***[8], la ***méthode d'échelonnement***, les ***échelles de cotation***, la ***méthode des intervalles apparemment égaux*** et leurs variantes, ont permis de déboucher sur une base commune aux tests psychophysiques et mentaux. L'objectif majeur de ces méthodes consiste à évaluer des stimulus sur des échelles linéaires, par exemple l'échelle des valeurs affectives, des valeurs liées à la croyances ou au pouvoir de persuasion ; ou encore l'échelle des valeurs liées à la qualité d'écriture, de dessin ou de composition ; ou celle des valeurs liées aux traits de personnalité comme le leadership, la délicatesse ou la sociabilité ; etc. Dans tous ces cas, il n'y a pas d'évaluation physique des stimuli, le traitement psychophysique intégral est donc hors de question. Un grand nombre de ces méthodes d'échelles trouvent toutefois leurs origines en psychophysique et, ces dernières années[9], on doit surtout à Thurstone de les avoir rationalisées sur base de la théorie psychophysique.

Les développeurs de tests mentaux apprécient les méthodes d'échelonnement pour leur utilité dans le cadre de la problématique pédagogique, et tout particulièrement pour leur mode systématique et objectif grâce auquel on peut obtenir une évaluation précise des individus. Ces méthodes ont fourni aux développeurs de tests des critères qui leur ont beaucoup servi pour éprouver la validité de leurs tests ; elles leur ont également tenu lieu de mode d'expérimentation pour évaluer des traits de personnalité pour lesquels il n'existait pas encore de tests spécifiquement établis. On peut donc relier les méthodes d'échelles à la fois à la psychophysique et aux tests mentaux. La psychophysique leur a apporté des bases rationnelles et mathématiques et a favorisé leur usage en psychologie expérimentale ; quant aux tests mentaux, ils ont tiré bon nombre d'informations empiriques de leur application à l'enseignement et aux problèmes issus des différences individuelles.

[8] Cette méthode a souvent été appelée méthode des *comparaisons pairées*. Or, ce ne sont pas les comparaisons qui vont par paires, mais bien les stimuli.

[9] NDT : Ne pas oublier que l'ouvrage a été publié en 1954 !

3. Théorie générale de la mesure

Le consensus actuel semble emboîter le pas à l'opinion de **Campbell** (Campbell et *al.*, 1940), qui définit la mesure comme *l'attribution de valeurs numériques selon certaines règles à des objets ou des événements*[10]. En tout cas, cette définition est bien souvent citée. L'auteur la considère comme satisfaisante, si ce n'est en ce qui concerne le terme de « valeurs numériques », auquel d'autres auteurs préfèrent le terme « nombres ». La distinction entre les deux n'est toutefois pas toujours claire. En gros, les valeurs numériques sont de simples symboles. Pour d'aucuns, elles ne sont que des griffonnages sur des feuilles de papier. Bien sûr, elles sont bien plus que cela, puisque chacune d'elle revêt une forme distincte de toutes les autres. Elles sont donc souvent utilisées pour désigner des objets ou des groupes, sans pour autant donner une forte implication à la signification numérique. Ainsi, on parle souvent des « groupe 1 », « groupe 2 », etc., sans raison particulière, si ce n'est celle de distinguer les groupes. Cependant, les valeurs numériques utilisées en ce sens suggèrent au moins une propriété des nombres, celle de la singularité, de l'identité. Ce qui justifie ici que nous parlions de nombres. Nous verrons plus loin que l'utilisation de nombres pour désigner des catégories est la forme la plus élémentaire de la mesure.

3.1. La nature des mathématiques

L'étroite relation entre mesure et mathématiques a été évoquée plus haut. Nous ne pouvons comprendre la nature de la mesure si nous ignorons les propriétés des mathématiques. Beaucoup d'étudiants suivent des cours de mathématiques sans comprendre vraiment ce que sont les mathématiques. Enseignants en mathématiques et auteurs de manuels semblent souvent mésestimer la nature sous-jacente de leur sujet. Peut-être est-ce parce qu'ils en sont beaucoup trop proches. Il a fallu les philosophes, comme Bertrand Russel, pour comprendre ce que sont réellement les mathématiques. Ils sont parvenus à la conclusion que les mathématiques constituent un langage extrêmement logique, sinon une branche de la logique.

3.1.1. Postulats et théorèmes

Toute branche des mathématiques repose d'abord sur une série de postulats. Un postulat est un énoncé considéré comme vrai sans qu'il soit nécessaire de le prouver d'aucune façon. Un postulat énonce une hypothèse que l'on fait au sujet d'une relation entre des objets. Par exemple, nous pouvons affirmer en postulat que $a + b = b + a$. Cela veut simplement dire que si nous associons deux objets, a et b, l'ordre dans lequel l'association a lieu n'affecte pas le résultat. Nous pourrions aussi bien affirmer le contraire[11], que l'ordre *importe* en fait, mais alors, nous tirerions de ce second postulat des conclusions différentes de celles issues du premier. Un postulat est utile pour les conclusions ou déductions que nous pouvons en tirer,

---

[10] MD : Michell (1999) analyse et critique longuement cette conception qu'il considère à la fois comme hégémonique et intenable en psychologie. Il rappelle (p. 185) qu'avant 1951, année de la « révolution » imposée par Stevens (1951), les auteurs comme Fechner (1860, p. 38) indiquent que « généralement la mesure d'une quantité consiste à définir combien une quantité unitaire de même nature est contenue en elle ». Cette définition correspond bien à celle d'autres auteurs classiques comme Titchener, et au recours classique à un étalon, mais s'écarte de la simplification généralisante proposée par Stevens, Campbell et leurs successeurs jusqu'aujourd'hui.

[11] MD : cette idée est sans doute assez difficile à admettre car, pour nous, l'opération mathématique prend ordinairement le pas sur la réflexion logique. Imaginons cependant que nous disposons d'un objet fragile (par exemple, un vase en cristal) et un objet lourd (disons, un piano). Le fait de poser l'objet fragile, puis, dessus, l'objet lourd, ou l'inverse, ne conduira pas au même résultat, si nous désirons mesurer la hauteur de l'empilement des deux objets. Dans ce modèle physique, le signe « + » signifie « mettre a sur b ».

ou tirer de son association avec d'autres postulats. Pour élaborer un système au départ d'une série de postulats, il est impératif qu'aucun d'eux ne soit en contradiction avec un autre. Il faut une cohérence interne. Par souci d'économie, aucun postulat ne doit être itératif ou empiéter sur un autre, bien que, même si c'était le cas, cela n'invaliderait pas le système. Le nombre de postulats dépendra du nombre nécessaire pour étayer le système.

Par déduction logique, on peut inférer d'autres énoncés, appelés *théorèmes*. Si le raisonnement est étroitement ajusté, cohérent avec les postulats, les théorèmes déduits seront exacts puisque les postulats sont exacts. L'exactitude évoquée ici est de type logique, non de type empirique. Des postulats aux théorèmes, nous sommes en plein dans le monde des idées. Il n'y a pas lieu de rechercher de preuve expérimentale à ces déductions. Pareille démarche n'aurait aucun sens. La seule nécessité de preuve qui soit appropriée relève entièrement du monde de la logique.

3.1.2. Modèles mathématiques

En d'autres mots, les postulats pas plus que les théorèmes mathématiques ne rendent compte de quoi que ce soit du monde dans lequel nous vivons, le monde de l'observation. L'ancienne idée des Grecs, selon laquelle le monde fonctionne sur une base mathématique, est inexacte. *Les mathématiques sont une invention de l'homme, non une découverte*. Il est tout aussi inexact de prétendre que la courbe de répartition gaussienne, ou normale, est une courbe biologique ou psychologique. Elle n'est ni l'un ni l'autre. C'est une courbe mathématique, purement et simplement. Le fait qu'elle puisse servir à décrire des répartitions obtenues lors d'observations faites en biologie et en psychologie est une coïncidence. Ce qui n'enlève rien à la grande commodité ou à la précision conférées par l'usage de cette courbe en tant que modèle pour décrire des événements de nature biologique et psychologique. Voici, en effet, un très bon exemple de la *fonction générale des mathématiques* : fournir des modèles pratiques et avantageux pour décrire la nature. La nature n'est jamais *exactement* décrite par quelque modèle mathématique que ce soit. Toutes ces descriptions ne sont que des approximations, dont certaines sont meilleures que d'autres.

3.1.3. Isomorphisme

On ne peut donc dire que la nature obéit à des lois mathématiques. Mais si cette affirmation est exacte, comment se fait-il que nous utilisions des modèles mathématiques pour décrire la nature ? Comment peut-on assigner des valeurs numériques ou des nombres à des objets ou des événements ? Comment peut-on mesurer ce qui n'existe pas sous une forme numérique ? La réponse est que la structure de la nature telle que nous la connaissons comporte des propriétés qui sont suffisamment parallèles à la structure des systèmes logiques en mathématiques. Il existe entre les deux ce qu'il convient d'appeler un *isomorphisme*, c'est-à-dire une équivalence de forme. A certains points, cette équivalence est parfaite jusque dans le moindre détail, et à d'autres points, elle est très grossière. L'*application* de quelque modèle mathématique que ce soit à quelque aspect de la nature que ce soit *peut* être testé empiriquement. Par exemple, si nous appliquons la courbe de répartition normale à la description d'une série donnée de mesures, nous pouvons tester la « qualité de l'ajustement » en effectuant un test de χ^2. Si la valeur du χ^2 est faible, nous acceptons le modèle de la courbe normale comme applicable à ces données. Si la valeur du χ^2 est tellement grande qu'on ne peut l'obtenir qu'avec une très petite probabilité, nous rejetons le modèle. Si nous considérons que l'ajustement est acceptable, nous pouvons alors exploiter les propriétés mathématiques de la courbe normale pour tirer des conclusions sur les données et pour faire des prévisions qui reposent sur ces propriétés. Nous pouvons agir ainsi avec la certitude que nos conclusions et nos prévisions ne comporteront que d'infimes erreurs, et qu'elles pourront donc servir nos objectifs.

3.2. Nature des nombres

3.2.1. Introduction

Puisque les nombres constituent le fondement des mesures et puisque le système numérique est une part importante du corpus général des mathématiques, nous allons maintenant nous intéresser aux propriétés des nombres. Et nous verrons que celles-ci correspondent à l'image des mathématiques que nous avons brossée plus haut.

Il n'y a pas de définition des nombres qui englobe toutes les sortes de nombres. Il existe bien une définition, généralement attribuée à Bertrand Russell, qui fonctionne bien lorsqu'elle s'applique aux **nombres rationnels**. Selon elle, un nombre est une « catégorie de toutes les catégories ». Il s'agit bien là d'une abstraction de haut niveau. Sa signification peut être représentée sous forme d'une illustration. Certaines catégories d'objets (peu importe de quels objets il s'agit) appartiennent à la même catégorie parce qu'elles ont en commun la propriété d'avoir deux objets en elles. Deux poissons, deux hommes, deux crayons, deux idées, tout cela appartient à la même catégorie pour une seule raison : l'élément commun de la dualité. Toutes les paires, les trios, les quartets, etc., constituent des catégories simplement en raison de leur « numérosité » identique. L'opération expérimentale qui nous permet de déterminer si deux catégories appartiennent à la même catégorie est d'assembler deux par deux les membres des deux catégories et voir si cela fonctionne. Si oui, ils ont une numérosité semblable et on peut leur attribuer le même nombre. Si l'une des catégories comporte un ou plusieurs membres qui ne correspond(ent) pas à l'un ou l'autre de l'autre catégorie, c'est qu'ils n'ont pas la même numérosité et qu'on ne peut donc leur attribuer le même nombre. La séquence du système numérique des nombres entiers est établie par l'organisation de catégories comportant juste un membre de moins lorsque l'on compare chacune d'elle avec celle qui la suit directement au sein du système.

Encart 3 – *Les nombres.*

Qu'est-ce qu'un nombre ? La question n'est pas simple ! Suivons ici le canevas de l'exposé de S. Baruk (1995, p. 758 et suivantes).

L'idée courante de nombre, portée par la langue commune, n'a pas grand-chose à voir avec les problèmes qu'a posé et que pose toujours la question en mathématiques. Elle n'est pas simple pour autant : « Nous 'rencontrons' des cailloux et des arbres », dit le philosophe ; « mais trois cailloux, deux arbres ? Jamais. Pour les voir, il y faut déjà quelque opération ».

Tout autant que le point, la droite ou le plan en géométrie, les nombres sont des idéalités. Trois, deux, sont des nombres au sens où on l'entend en mathématiques. Mais, quand dans la vie courante on énonce un nombre (au sens mathématique), il est pratiquement toujours suivi de ce qu'il compte ou énumère : « la distance de Paris à Rouen est de 130 kilomètres, j'ai roulé à une moyenne de 90 kilomètres à l'heure, au premier péage il faut payer 7 francs » ; ou alors c'est qu'il le sous-entend : « nous partîmes cinq cents » (sous-entendu : « hommes ») ; il s'agit donc toujours d'un nombre de quelque chose.

Les nombres sont mathématiques ; les nombres-de couramment utilisés dans la vie quotidienne font un usage spécifique des nombres – le plus souvent limité aux décimaux – qui constitue en particulier ce qu'on peut appeler le quantitatif ; les nombres-de d'usage savant utilisent les nombres de façon propre à chaque science.

Résultant de siècles d'interrogations et de recherches sur la spécificité des nombres, un ordre d'exposition devenu « classique » et fondé sur les « extensions » successives de leurs « aptitudes » face aux opérations permet d'aborder successivement :

N ensemble des nombres entiers naturels [par exemple, 0 ; 1 ; 2 ; 3]

Z ensemble des nombres entiers relatifs [par exemple,-3 ; -2 ; 0 ; 1 ; 2 ; 3]

D ensemble des nombres décimaux (relatifs) [par exemple,-3 ; -2 ; -1,36 ; 0 ; 1 ; 2 ; 2,658 ; 3]

Q ensemble des nombres rationnels (relatifs) [par exemple, $-3 ; -2 ; -1{,}36 ; \frac{-1}{4} ; 0 ; 1 ; \frac{1}{3} ; \frac{6}{9} ; 2 ; 2{,}658 ; 3 ; \frac{32}{10}$ *]*

R ensemble des nombres réels (relatifs) [ensemble des nombres rationnels et des nombres irrationnels[12]]

A ces ensembles qui s'incluent progressivement ($N \subset Z \subset D \subset Q \subset R$), on peut en ajouter encore d'autres, comme l'ensemble des nombres complexes.

Se demander ce qu'est un nombre, c'est à la fois s'interroger sur les relations qu'entretiennent ou que n'entretiennent pas entre elles l'histoire des mathématiques et l'histoire des sociétés, que l'on pourrait presque rapporter aux relations qu'entretiennent entre eux nombre et nombres-de. En effet, aussi loin qu'il est possible de remonter dans l'histoire des civilisations, le nombre apparaît comme une notion fondamentale et caractéristique de l'esprit humain ; ceci dans plusieurs domaines correspondant à des activités distinctes, que l'on pourrait schématiser ainsi :

a. Une activité liée aux nécessités de la vie pratique : la nécessité et le désir de comptabiliser les biens, qui sont, semble-t-il, à l'origine même de l'écriture – les premières traces écrites qui apparaissent dans l'histoire de l'humanité sont celles mentionnant les revenus de temples et palais sumériens, plus de 3000 ans avant Jésus-Christ -, celles aussi de codifier la propriété et de régler les échanges font apparaître les nombres à partir de quantités matérielles ou de mesures de grandeurs : troupeaux dont on compte les bêtes, terrains dont on mesure les dimensions, récoltes dont on évalue le volume ou le poids, échanges de marchandises dont on calcule les valeurs respectives (avant l'apparition de la monnaie, ces valeurs se calculaient, par exemple, en têtes de bétail ou en ration d'orge), etc.

b. Une activité scientifique et technique : une activité technique, également liée aux nécessités de la vie pratique, et une activité scientifique, fondée sur la ***quantification*** *des phénomènes et sur la recherche des lois les exprimant et les prévoyant, mettent en jeu des nombres qui, ici aussi, sont des nombres-de.*

c. Une activité spéculative, spécifiquement mathématique : les mathématiques en tant que telles sont une activité spéculative qui étudie les nombres pour eux-mêmes, et ce, quelle que soit leur origine, fût-elle matérielle. Si, incontestablement, les ***entiers*** *et les* ***fractions*** *se sont présentés à l'esprit humain d'abord sous forme de nombre-de, très vite ils ont été étudiés pour eux-mêmes. Pour reprendre l'exemple célèbre de* π*, en supposant que son calcul s'impose en raison de nécessités matérielles ou techniques, « quatre décimales font connaître* π *avec une précision suffisante pour les besoins pratiques. Avec 16 décimales, on obtient à l'épaisseur d'un cheveux près la longueur d'une circonférence ayant pour rayon la distance moyenne de la Terre au Soleil. Remplacer le Soleil par la nébuleuse la plus lointaine et le cheveu par le plus petit corpuscule connu des physiciens, pour atteindre cette précision fantastique, vous n'avez guère besoin que de 40 décimales ». Or on connaît aujourd'hui un milliard de décimale de* π*..*

Le lecteur intéressé par l'histoire des nombres(-de) et de leur écriture (les chiffres) peut utilement consulter les ouvrages d'Ifrah (1994) et de Warusfel (1961).

Cet exposé peut paraître quelque peu prosaïque à certains lecteurs, mais il s'agit là d'un sujet de grande importance si l'on veut assurer des fondements logiques. Il s'agit là de l'un des traits particuliers de l'histoire de la pensée humaine que la nécessité d'avoir de grands esprits pour saisir ce qui apparaît d'abord comme des idées très simples.

3.2.2. Développement du système numérique

Le système numérique ou, plus exactement, les systèmes numériques, ont connu un développement particulièrement intéressant. Depuis les débuts, avec le système des entiers naturels, jusqu'au système actuel des nombres complexes, les nouvelles propriétés des nombres sont apparues toujours croissant, à l'instar des opérations qu'on peut leur appliquer.

Le *système naturel* des nombres comprend les nombres entiers positifs. Ce système a sans aucun doute été conçu pour répondre aux besoins de comptage d'objets discrets. Cet objectif ne nécessite en effet que des nombres entiers positifs. On a aussi découvert que ce système permettait d'additionner et de multiplier. L'une ou l'autre de ces opérations aboutit à un nombre qui lui aussi est un entier positif ; en d'autres termes, le résultat appartient lui aussi au même système. La soustraction, quant à elle, était limitée dans son application au sein de ce système. Elle fonctionnait, sauf si on s'efforçait de soustraire un nombre de lui-même ou d'un nombre plus petit. Dans ces cas, il n'y avait pas de nombres qui correspondent dans le

[12] MD : Nombres irrationnels : nombres décimaux illimités non périodiques, comme $\sqrt{2}$ ou π.

système. Cette limitation permit l'extension du système au zéro et aux nombres négatifs, soit deux nouveaux concepts très importants. L'opération de multiplication était encore plus limitée. Elle fonctionnait seulement tant que les deux nombres impliqués étaient dans un rapport simple, soit lorsque l'un était un multiple exact de l'autre. Pour prendre en compte les autres cas, on inventa les nombres fractionnaires.

Le système numérique qui comprend les nombres positifs, négatifs et fractionnaires est appelé le *système rationnel*. Au sein de ce système, n'importe quel nombre peut être exprimé comme le rapport entre deux nombres entiers du système. On peut y effectuer les quatre opérations fondamentales, à l'exception de la division par zéro. Le système rationnel fournit également tout le matériel nécessaire pour toutes sortes de mesures.

Toutefois, le traitement des données métriques requiert souvent des opérations mathématiques qui sont impossibles dans le système rationnel. Les racines carrées de beaucoup de nombres, par exemple, ne peuvent être exprimées sous forme de nombres rationnels. La racine carrée de 2 est en-dehors du système numérique rationnel. Il a fallu inventer le concept des nombres irrationnels pour prendre en compte des résultats de ce type. Pour des raisons d'ordre pratique, on propose comme approximation de la racine carrée de n'importe quel nombre un nombre arrondi qui est inclus dans le système rationnel. Cela permet de rencontrer la majorité des objectifs de calculs que l'on vise en pratique.

3.3. Application des nombres aux mesures

En vertu du principe d'isomorphisme évoqué ci-dessus, on peut valablement se servir des nombres pour effectuer des mesures (et les attribuer à des objets ou des événements), du moment que les propriétés des nombres sont parallèles à celles des objets ou des événements auxquels ils sont attribués. Ce qui nous amène à effectuer un examen plus approfondi des propriétés des nombres et des relations entre eux.

3.3.1. Quelques propriétés des nombres utiles à la prise de mesures

Brièvement (car nous aurons l'occasion d'y revenir plus en détails dans les chapitres suivants), les propriétés des nombres les plus importantes dans le cadre de la prise de mesures sont au nombre de trois : identité, propriété d'ordre et additivité. A l'exception des cas d'égalité, les nombres peuvent être localisés dans un ordre incontestable sur une échelle linéaire. Par « additivité », on entend le fait qu'une addition produit des résultats qui sont intrinsèquement cohérents. Nous développerons ce point particulier plus loin. Au stade actuel, il est surtout important d'examiner ce qu'implique le concept d'additivité, à savoir que les quatre opérations fondamentales peuvent être appliquées, puisque la soustraction, la multiplication et la division peuvent être chacune considérée comme un cas particulier d'addition. Si l'addition peut être appliquée à des nombres rationnels, alors les trois autres opérations le peuvent également. La soustraction est une addition de deux nombres dont l'un est négatif. La multiplication est une série d'additions successives du même nombre. A l'inverse, la division est une série de soustractions successives, ce qui, en vertu de l'énoncé sur les soustractions cité un peu plus haut, correspond à une série d'additions de nombres négatifs. Ainsi, la nécessité de la propriété d'additivité vaut pour toutes les opérations numériques fondamentales.

Quand les propriétés d'ordre et d'additivité interviennent-elles dans les phénomènes naturels ? A moins que l'une des deux, ou les deux, soi(en)t présente(s), l'attribution de nombres à un domaine spécifique d'observation ne nous apportera pas grand-chose. Les nombres ainsi attribués ne seront guère porteurs de sens, et on ne pourra donc pas en faire grand-chose de vraiment utile ou significatif. Que faut-il faire et que pouvons-nous faire en termes d'expérimentations pour réussir à prouver que ces deux propriétés des nombres s'appliquent à un domaine particulier ? On peut juste supposer qu'elles s'appliquent, sur base d'une connaissance superficielle du domaine, mais nous n'en serons pas vraiment sûrs, sauf si nous disposons de quelques informations qui renforcent cette supposition. Certains parmi ceux qui se sont penchés sur le problème insistent sur le fait qu'il faut pouvoir démontrer le parallélisme entre les propriétés numériques et le domaine observé, et ce en menant de véritables expérimentations.

3.3.2. Démonstration expérimentale de la propriété d'ordre

Comment démontrer les propriétés d'ordre et d'additivité ? Il est assez facile de trouver un exemple de démonstration de la propriété d'ordre. En géologie, une échelle d'ordre de dureté des minerais est établie grâce à des expériences de frottement. Le minerai qui provoque des rayures sur l'autre est le plus dur des deux. On peut ainsi établir un ordre de dureté. En psychologie animale, une poule qui donne régulièrement des coups de bec à une autre lorsqu'elles sont en conflit pour de la nourriture, est considérée comme la poule dominante. Grâce au test des coups de bec, on peut établir un ordre de coups de bec (ou de domination). Par le biais d'observations directes, on peut ordonner des tonalités sonores en fonction de la hauteur de tonalité sur base du jugement « est plus haut que » ; les couleurs rouges peuvent être ordonnées en fonction de l'intensité du rouge sur base du jugement « est plus rouge que » ; des films peuvent être ordonnés en fonction de leur drôlerie sur base du jugement « est plus drôle que ». Nous ne nous attarderons pas ici sur le cas d'inversions d'ordre occasionnelles pour des paires d'objets juxtaposés. De telles inversions sont généralement considérées comme le fruit « d'erreurs d'observation » ou « d'erreurs de mesure » et ne justifient pas d'invalider le principe général de la propriété d'ordre ou de l'opération de mesure. Il arrive néanmoins qu'il soit nécessaire de prouver par des méthodes corrélatives ou autres qu'il y a bien une certaine cohérence entre les rangs. C'est le problème de la fiabilité (Guilford, 1954, chapitre 13).

3.3.3. Démonstration expérimentale de l'additivité

La propriété d'additivité est assez difficilement démontrable de façon expérimentale, même dans le cas de la physique. Voyons comment l'on peut procéder. Un exemple de la propriété de longueur est probablement l'un des plus évidents à démontrer. Considérons d'abord de quelle façon sont démontrées les propriétés d'ordre dans ce domaine. Si nous prenons deux objets linéaires (fils rigides, segments ou planches) et que nous les disposons côte à côte, l'une des extrémités d'un des deux objets étant sur le même plan que l'extrémité correspondante de l'autre objet, nous pourrons dire, en comparant les deux autres extrémités, si les objets sont d'égale longueur ou si l'un est plus long que l'autre, et dans ce cas, lequel des deux. A noter que le jugement concernant cet ordre est effectué sur base d'une observation visuelle. Si la mise en ordre de longueurs est possible sur base d'un jugement humain, il en va de même pour la mise en ordre des perceptions et des émotions évoquées ci-dessus, hauteur de tonalité, intensité de rouge et drôlerie. Toutes reposent sur la réaction d'un observateur.

Dans le cas de segments de droite, nous pouvons démontrer l'addition d'une façon qui ne peut être reproduite pour la hauteur de tonalité, l'intensité du rouge ou la drôlerie. On peut placer ces segments de droite bout à bout, pour obtenir une longueur totale. Si nous avons une échelle de distances calibrée, nous pouvons démontrer que la combinaison $a + b$ est égale à la longueur prévue c. Ce type d'empilement expérimental est impossible au niveau des propriétés psychologiques. S'il nous faut démontrer l'addition de propriétés psychologiques, il faudra le faire autrement.

On peut également démontrer empiriquement l'addition en physique, dans le domaine du poids et celui des résistances électriques. On peut superposer des poids sur une échelle. Des résistances peuvent être placées « bout à bout » d'une façon évidente. Mais en dehors de ces exemples simples, même en physique, une addition empirique reste à ce jour difficile, voire impossible à démontrer. Sur les échelles habituelles de température, il n'y a pas d'opération permettant de faire la somme de deux niveaux de température. Même dans le domaine des longueurs ou des distances, on rencontre de sérieuses restrictions à la démonstration de l'addition par expérimentation. Personne n'a encore jamais placé bout à bout des années lumière, pas plus que personne n'a encore démontré l'addition de distances atomiques. On peut même douter que les distances linéaires de proportions astronomiques (et peut-être de proportions atomiques) puissent être en réelle continuité avec les types de distances linéaires que l'on rencontre tous les jours. Ainsi, dans toutes les sciences, l'hypothèse de l'applicabilité des propriétés numériques repose sur des preuves empiriques extrêmement restreintes. Il faut donc laisser à chaque science l'opportunité de produire son propre type de preuve empirique de la mesurabilité. Nous verrons qu'il existe en psychologie des façons spécifiques de démontrer que les phénomènes psychologiques comportent les propriétés qui justifient nécessairement la prise de mesures.

3.3.4. Mesures dans le cas de propriétés numériques restreintes

Un autre point important à relever est qu'il n'est pas nécessaire que des phénomènes présentent *toutes* les propriétés numériques, comme l'additivité, pour que nous puissions effectuer des mesures utiles. La propriété d'ordre suffit à beaucoup d'objectifs. Toutefois, lorsque l'additivité n'est pas présente, les nombres que nous attribuons ont une signification restreinte et on ne peut leur appliquer toutes les opérations numériques. Comme nous le verrons plus loin, les mesures telles que l'on peut les relever aujourd'hui peuvent présenter divers degrés de complétude. Or, elles sont utiles et porteuses de sens dans la mesure où elles sont complètes. Il est donc important que nous puissions identifier leurs lacunes, afin d'éviter de vouloir leur attribuer plus de sens qu'elles n'en ont.

L'***échelle de centiles*** généralement utilisée en psychologie est un bon exemple de mesure limitée. Un centile est un rang parmi 100 valeurs de rang différentes. Un centile de P_{80} indique donc le 80^{e} rang à partir du bas de l'échelle. La différence $P_{80} - P_{60}$ indique qu'il y a 20 % de sujets observés entre ces deux limites. Ici, nous faisons bien une soustraction porteuse de sens. La différence $P_{60} - P_{40}$ offre la même interprétation. On peut considérer ces deux différences comme égales de ce point de vue. Toutefois, d'un autre point de vue, elles ne sont pas égales. Si nous considérons une échelle de compétence, où nous voulons que les différences indiquent des variations de sommes de compétences, la différence $P_{80} - P_{60}$ est plus grande que la différence $P_{60} - P_{40}$. Celle-ci se situe au milieu des rang de compétence. Le même nombre d'individus (20 %) présentent une plus large compétence au-dessus du centre de la distribution, où les sujets observés se raréfient.

3.4. Quelques postulats fondamentaux pour les mesures

Avant d'examiner plus avant les quatre niveaux généraux de mesures tels qu'on les distingue généralement aujourd'hui, il est nécessaire de spécifier précisément les propriétés numériques qui doivent être rencontrées. Les neuf postulats proposés ci-dessous sont ceux essentiellement émis par Campbell et revus par la suite avec quelques variantes (Comrey, 1950 ; Helson, 1951 ; Reese, 1943 ; Stevens, 1951). Les trois premiers concernant l'identité, les deux suivants l'établissement d'un ordre, et les quatre derniers, l'additivité.

1. Soit $a = b$ ou $a \neq b$
2. Si $a = b$, alors $b = a$
3. Si $a = b$ et $b = c$, alors $a = c$
4. Si $a > b$, alors $b \not> a$
5. Si $a > b$ et $b > c$, alors $a > c$
6. Si $a = p$ et $b > 0$, alors $a + b > p$
7. $a + b = b + a$
8. Si $a = p$ et $b = q$, alors $a + b = p + q$
9. $(a + b) + c = a + (b + c)$

Quelques commentaires vont servir à donner plus de sens à ces postulats. Le premier postulat établit l'identité d'un nombre. Les nombres sont identiques ou ils sont différents. Le deuxième postulat affirme que la relation d'égalité est symétrique. Une affirmation d'égalité peut être inversée et rester exacte. Le troisième postulat exprime sous forme d'équation le dicton familier : « Des objets égaux à un même objet sont égaux entre eux. » Le 4e postulat signifie que la relation > est *asymétrique*. On ne peut inverser les affirmations $a > b$ ou $a < b$, en inversant a et b, et aboutir à un énoncé exact. Le postulat 5 indique que la relation d'ordre est transitive. Un exemple expérimental d'une série intransitive de relations pourrait consister à observer que l'équipe *A* a battu l'équipe *B*, l'équipe *B* a battu l'équipe *C*, mais l'équipe *C* a battu l'équipe *A*. Ici, l'ordre est plutôt circulaire que linéaire. D'autres irrégularités peuvent également annuler la transitivité d'une série d'objets[13]. Le postulat 6 indique la possibilité d'opérer une addition ; il implique également que l'addition du nombre 0 laisse un nombre inchangé. Le postulat 7 signifie que l'ordre dans lequel des objets sont additionnés ne modifie en rien le résultat. Le postulat 8 signifie que des objets identiques

[13] MD : Charpak et Broch (2002, pp. 100-101), en rappelant le ***Paradoxe de Condorcet,*** en fournissent une bonne illustration dans le domaine électoral : *Un sondage est fait auprès des électeurs afin de savoir comment ils classent les candidats* [A, B et C] *par ordre de préférence décroissante. Mille personnes sont interrogées et l'ensemble des résultats obtenus lors de ce sondage se trouve dans le tableau ci-dessous.*

Nombre d'électeurs sur 1000	***385***	***370***	***205***	***25***	***10***	***5***
Choix 1	*A*	*B*	*C*	*A*	*C*	*B*
Choix 2	*B*	*C*	*A*	*C*	*B*	*A*
Choix 3	*C*	*A*	*B*	*B*	*A*	*C*

Ainsi, dans la première colonne, nous pouvons voir que 385 personnes préfèrent A à B et B à C.
Nous avons 385 + 205 + 25 soit 615 personnes sur 1000, soit 61,5%, qui, d'une manière ou d'une autre, préfèrent A à B. Avec un tel résultat, aucun problème donc pour l'élection si A est confronté à B.
De la même manière, nous voyons que 385 + 370 + 5 soit 760 personnes sur 1000, c'est-à-dire 76% des personnes interrogées préfèrent B à C . Le plébiscite est tel que B n'a pas de souci à se faire s'il est confronté à C, il l'écrasera à plate couture.
Résumons la situation : A est préféré à B par 61,5% des personnes (A>B), et B est préféré à C par 76% des personnes (B>C). Il semblerait que l'on puisse logiquement en déduire que A sera préféré à C (A>C).
Regardons de plus près. A est préféré à C par 385 + 25 + 5 soit 415 personnes, alors que C est préféré à A par 370 + 205 + 10 soit 585 personnes ! Donc, contrairement à ce que l'on pensait pouvoir logiquement déduire de nos premiers constats, C est préféré à A par 58,5% des personnes sondées.
Ce petit paradoxe, le paradoxe de Condorcet, qui peut apparaître dans des situations de choix sur la base de trois critères, est surprenant parce que l'on suppose qu'une relation exprimée en termes de préférence est toujours transitive, ce qui n'est pas le cas. Une relation de préférence n'est pas une relation d'ordre au sens mathématique.

peuvent être remplacés l'un par l'autre dans une addition. Enfin, le postulat 9 signifie que l'ordre de combinaisons et d'associations ne fait aucune différence dans une addition.

3.5. Les quatre niveaux généraux de mesures

Habituellement, on distingue quatre niveaux de mesures, qui ont été clairement décrits par Stevens (1946, 1951). Du niveau le plus bas au plus élevé, il y a l'échelle de mesure nominale, ordinale, d'intervalles et de rapports. Ces échelles se distinguent entre elles selon plusieurs critères. En fonction de notre définition des mesures – assignation de numéros à des objets ou des phénomènes selon des règles -, les règles déterminant la façon dont les numéros sont assignés constituent les critères essentiels servant à définir les échelles. Celles des niveaux supérieurs nécessitent des règles plus strictes ; la plupart des postulats fondamentaux sont d'application. Il y a aussi des différences entre les quantités d'opérations mathématiques et statistiques qui peuvent être effectuées sur les nombres selon les divers niveaux de mesures. Plus l'échelle se situe à un niveau élevé, plus il est possible d'exploiter les nombres obtenus par la prise de mesures. Examinons à présent les règles et les opérations possibles pour chaque niveau d'échelle.

3.5.1. Les échelles nominales

Dans le cas de mesures sur des échelles nominales, un nombre[14] est simplement utilisé comme une étiquette pour désigner une classe ou une catégorie[15]. Les objets d'une classe sont considérés comme égaux ou équivalents sous certains aspects. Nous les désignons collectivement comme le « groupe 1 » , le « groupe 2 », etc. Les nombres attribués seraient substitués l'un à l'autre qu'ils serviraient aussi bien nos objectifs. Les seules règles qui régissent l'attribution des nombres sont que tous les objets d'une même classe reçoivent le même numéro et qu'un même numéro ne soit pas attribué à deux classes distinctes.

Parmi les postulats fondamentaux, seuls les trois premiers, relatifs à l'égalité, sont d'application. L'égalité entre deux objets, du point de vue où ils sont considérés, constitue le fondement sur lequel on se base pour les placer dans la même catégorie. Cela nécessite quelques éclaircissements concernant le terme *égalité*. Selon les postulats fondamentaux, égalité signifie identité. L'opération consistant à évaluer deux objets comme égaux ne constitue qu'une approche de cette condition idéale. Il arrive que des objets soient placés dans la même catégorie car ils sont indifférenciables selon les méthodes d'observation classiques. Il arrive aussi que des différences soient observables, mais ces différences peuvent toutefois être tolérées de manière à éviter d'avoir affaire à un nombre trop élevé de catégories. La finesse de discrimination dépend alors de nos capacités d'observation ainsi que de nos degrés d'exigence à l'égard d'une certaine précision ou de tolérance à l'égard d'une certaine approximation. Pour des raisons pratiques, il peut arriver que l'on accepte, voire que l'on préfère, des catégories relativement larges[16].

Il est intéressant de constater que la classification constitue assez logiquement la forme la plus élémentaire des mesures. Lorsque l'on peut ordonner des classes sur une échelle linéaire, ce qui est la cas pour la classification quantitative, nous avons fait un pas de plus vers une

[14] L'auteur emploie ici le terme *nombre* car les numéros utilisés, même de cette façon rudimentaire, impliquent un sens lié au nombre, en clair celui de l'identité.

[15] Cet emploi du terme *échelle* peut paraître surprenant. Dans un contexte numérique et mathématique, une échelle comporte des attributs de série et de graduation. Dans le dictionnaire, la définition du terme *échelle* fait état de « ce qui détermine des alternatives », et cela semble toutefois correspondre à l'emploi de ce terme. Les propriétés de séries et de graduations interviennent aux niveaux supérieurs des mesures.

[16] Pour un exposé concernant les procédures de classification des données, voir Guilford (1950, p. 14 à 17).

mesure complète. Même dans des types de mesures plus élaborées, le principe de classification est toujours d'application.

Même dans la classification de catégories qualitatives, on peut utiliser quelques opérations statistiques. Nous pouvons compter le nombre d'objets dans chaque catégorie, nous obtenons ainsi des *fréquences*. On peut chercher à savoir quelles sont les classes les plus denses et les plus communes. Cette classe-là serait le *mode* de répartition des classes. La classe modale est très utile pour élaborer des prévisions sur les catégories[17]. Si on a deux objets identiques classés de deux façons, sur base de deux aspects ou de deux principes de classification, on peut déterminer l'interdépendance de ces deux aspects en calculant un coefficient de contingence. Les fréquences, les modes et les coefficients de contingence sont des éléments statistiques qui peuvent très bien être utilisés dans le cas de données dans un système de catégories. Nous verrons que ces éléments de statistiques peuvent également servir lors de prises de mesures plus élaborées.

3.5.2. Les échelles ordinales

Dans le cas de mesures sur une échelle ordinale, les nombres attribués exploitent la propriété d'ordre. Les postulats 4 et 5 ci-dessus présentent le fondement logique de l'ordre des rangs (ou *arrangement*). Si *a* et *b* ne sont pas égaux, ils peuvent se différencier selon une certaine dimension qui concerne également d'autres paires d'objets. Cette dimension se traduit généralement en termes de l'un des aspects des objets –hauteur, chaleur, sonorité ou aptitude orale.

Il arrive que le fondement de la classification en catégories soit un composé de deux variables ou plus. Par exemple, on peut chercher à ordonner des individus selon leur « niveau socioéconomique », pour lequel il existe deux ou davantage d'indices de la variable, comme les revenus, l'instruction reçue et la profession. Deux individus peuvent se situer au même rang pour les revenus, et à des rangs totalement distincts pour le niveau d'instruction. Tant que nous nous confinons à l'une des deux variables, le postulat n° 4 ne sera pas violé. Nous ne pouvons toutefois satisfaire ce postulat à l'égard de la variable du composé, à moins que nous n'adoptions des pondérations relatives pour les indices constituant le composé de façon à ce que ce dernier établisse un et un seul arrangement.

Les arrangements opérés sur les variables d'un composé sont des phénomènes courants en psychologie. La plupart des variables psychologiques que nous traitons sont des variables latentes, basiques, ou primaires. On fixe ce qui constitue essentiellement une quantité multidimensionnelle sur une échelle linéaire. Quelques fois la pondération est explicite, comme c'est le cas lorsqu'on utilise plusieurs ***équations de régression***, mais le plus souvent, elle est implicite, comme par exemple avec des évaluations faites par des observateurs humains. Dans ce cas, la pondération s'effectue de manière intuitive, et il n'y a aucun moyen de savoir ce que sont ces pondérations ou si elles sont appliquées de manière uniforme, sauf si on a recours à des opérations *d'échelonnage multidimensionnel* – ***analyse factorielle***[18].

On peut considérer l'opération consistant à « mettre en rang » ou ordonner comme une classification de catégories quantitatives. La distinction entre les catégories est ici basée sur une qualité ou une propriété des objets mis en ordre. Une discrimination complète signifierait que l'on ne place qu'un seul objet dans chaque catégorie, comme dans la *méthode d'arrangement*. Chaque catégorie a alors une fréquence de 1. Dans un sens plus général, toutefois, on peut tolérer certains liens afin de ne pas imposer des discriminations au-delà des

[17] Voir Guilford (1950, p. 365)

[18] Se reporter au chapitre 10 de Guilford (1954) pour un exposé sur l'échelonnage multidimensionnel, et au chapitre 16 du même ouvrage pour une description de l'analyse factorielle.

limites de la précision d'observation. On obtient alors des fréquences supérieures à 1 dans une ou dans toutes les catégories. La méthode qui correspond le mieux à cette description est appelée *méthode des catégories successives*. Cela n'implique pas que les catégories ordonnées par rang soient nécessairement espacées de la même façon sur une échelle, que les intervalles entre elles soient égaux. Nous allons aborder après le sujet des échelles d'intervalles égaux, dans le cadre du niveau supérieur de mesures.

Les éléments statistiques que l'on peut utiliser au niveau des mesures sur échelles nominales peuvent également être appliqués aux mesures sur échelles ordinales – ***fréquences***, ***modes***, ***coefficients de contingence***/de ***corrélation.*** Le principe d'ordre permet d'utiliser d'autres éléments statistiques, comme les ***médianes***, les ***centiles***, les ***coefficients de corrélation*** de rangs. Une médiane sépare les fréquences associées en deux parts égales au-dessus et au-dessous du point médian. Le nombre attaché à une médiane, issu des numéros d'ordre des rangs, indique lui-même simplement la position du rang dans l'intervalle de variation des numéros de rangs utilisés. La médiane peut coïncider avec l'un des numéros de rang utilisés, ou elle peut se situer à mi-chemin entre deux numéros de rangs successifs. Des démarches de discrimination plus fines pour évaluer des médianes de rangs seraient inutiles[19]. Les centiles étant également des numéros de rang, la médiane de toute succession de rangs en centiles est elle-même un numéro de rang parmi cent numéros.

Certains auteurs hésitent beaucoup à dire que le coefficient de corrélation de rangs s'applique convenablement aux valeurs d'échelles ordinales[20]. Cette hésitation nous paraît injustifiée. Par exemple, le coefficient de corrélation de Spearman ρ présuppose une série complète de rangs, avec des numéros successifs, chaque lien étant convenablement évalué dans les séries. La formule pour calculer ρ est une formule de covariance. Elle repose sur l'hypothèse d'intervalles égaux *entre les numéros d'ordre* (ce qui explique pourquoi on utilise une distribution complète des objets et des nombres entiers successifs). Donc, les numéros de rangs sont espacés de façon égale. Cela n'implique pas qu'une échelle idéale d'unités égales comporterait des objets successifs espacés par des intervalles égaux. Le coefficient ainsi obtenu ne fournit d'information directe que sur la concordance entre les deux séries de numéros de rangs. Quand on évalue un coefficient de corrélation de Pearson r à partir d'un coefficient ρ, on est simplement en train de prédire ce que serait ce coefficient de Pearson si, au lieu de connaître les rangs des objets, on avait leurs évaluations sur deux échelles d'unités semblables et qu'on leur appliquait la formule de Pearson.

L'application de la formule du coefficient de corrélation de Pearson aux valeurs des rangs en centiles dans des cas courants fournit probablement une estimation assez juste de la corrélation. En utilisant des centiles, qui s'appliquent parfaitement à un échantillon où les variables X et Y sont corrélées, on se rendrait compte que le coefficient de corrélation r obtenu différerait d'un coefficient r que l'on aurait obtenu par la corrélation de valeurs d'intervalles, dans la même mesure où l'ordre de rang ρ diffère de r.

---

[19] Le rang médian peut servir dans trois situations distinctes. Si on a les premiers nombres entiers consécutifs n attribués à n objets, la médiane correspond au rang du milieu si n est un nombre pair et à une valeur intermédiaire entre deux rangs si n est un nombre impair. S'il y a des fréquences supérieures à 1 dans les catégories, et s'il y a une fréquence supérieure à 1 dans la catégorie où se situe la médiane, l'interpolation n'est porteuse d'aucun sens. L'interpolation repose sur l'hypothèse qu'il y a une continuité dans la catégorie, que la dimension de la catégorie est connue et que les objets de la catégorie sont répartis de façon régulière. Ici, aucune de ces hypothèses n'est étayée. Le cas où l'application de la médiane à des rangs est le plus utile, c'est lorsque le même objet apparaît plusieurs fois à des rangs différents sur une liste d'objets semblables, avec les mêmes catégories. Ici aussi, l'interpolation est exclue. La médiane doit être indiquée comme un rang existant ou comme un rang à mi-chemin entre deux rangs successifs.

[20] Voir notamment Stevens (1946, 1951).

3.5.3. Les échelles d'intervalles

Les échelles d'intervalles sont aussi appelées *échelles d'intervalles jugés égaux*. La propriété importante qui les distingue des types d'échelles plus rudimentaires est que les distances numériquement égales qu'elles comportent représentent des distances empiriquement égales entre les objets, sous un certain aspect. Ainsi, on peut dire par exemple à propos de valeurs quantitatives observées, que deux objets ayant reçu les numéros 5 et 10 sont aussi distants l'un de l'autre sur l'échelle que deux autres objets auxquels on a attribué les numéros 15 et 20. On peut également dire que la distance entre A et B plus la distance entre B et C est égale à la distance entre A et C. Si l'on mettait ces deux affirmations sous forme d'équation, on aurait $R - Q = T - S$ et $AB + BC = AC$. Certaines opérations expérimentales permettent de vérifier ces affirmations.

Bien que l'on puisse parler d'addition d'intervalles, cela ne signifie pas que l'on mette en pratique l'importante propriété d'additivité au sens absolu. La somme de quantités sur l'échelle n'a pas beaucoup de sens. La raison en est que le zéro est situé arbitrairement. La valeur numérique de la propriété définissant l'échelle représentée par le zéro n'est probablement pas la limite inférieure à laquelle cette propriété disparaît. Prenons deux nombres sur une échelle d'intervalles, par exemple 7 et 11. Le premier représente 7 unités à partir du zéro arbitrairement fixé et le second, 11 unités à partir du zéro. Additionnons les deux et nous obtenons 18 unités, ce qui signifierait que l'on aurait une quantité de 18 unités à partir de zéro. Supposons maintenant que ce zéro arbitraire soit fixé 5 unités plus haut que le point zéro initial. Il faudrait alors que les deux valeurs deviennent 12 et 16 si l'on veut maintenir les écarts réels entre elles et le zéro absolu et réellement porteur de sens. En additionnant 12 et 16, on obtient 28, et non plus 18. la variation de la somme concorde avec la variation de la position du point zéro sur l'échelle.

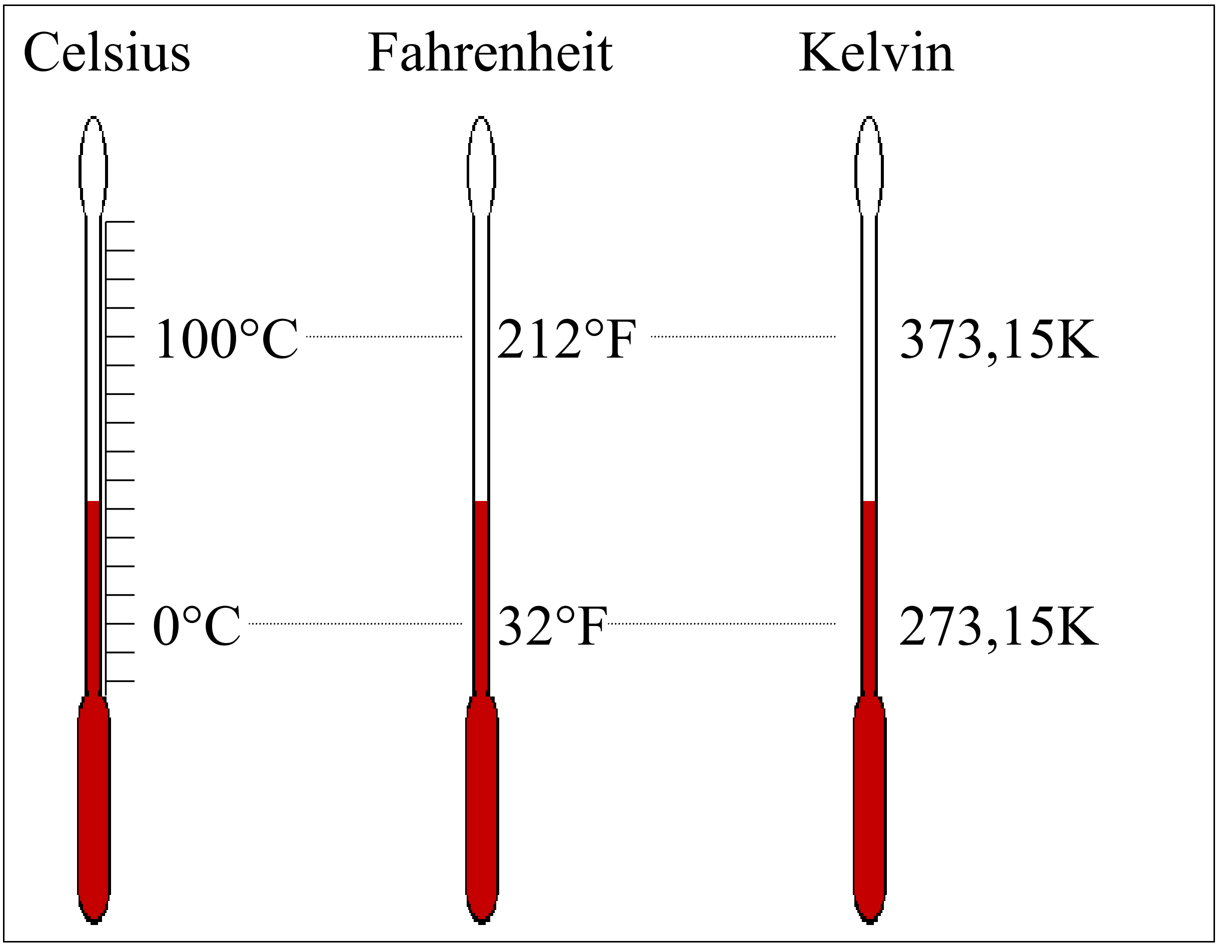
Parmi les exemples d'échelles d'intervalles, on a l'échelle courante des températures, qu'il s'agisse de l'échelle centigrade ou de l'échelle Fahrenheit (encart 4). La calendrier est lui aussi une échelle d'intervalles, le zéro ayant été fixé arbitrairement, par convention. Tout intervalle de temps est une mesure prise à partir d'un point de départ déterminé, tout comme l'altitude d'une montagne est une distance calculée à partir d'un autre point de départ déterminé, qui est le niveau de la mer. Dans les deux cas, il s'agit de distances –ou différences– réelles dans des échelles d'intervalles, mais on peut les considérer comme des quantités absolues. En d'autres termes, globalement, l'additivité étant une propriété des *distances* dans les échelles d'intervalles, on peut traiter ces distances de la même façon que l'on traite les mesures des échelles de niveau supérieur, les *échelles de rapport*, que nous verrons ensuite.

Il nous faut pourtant examiner quelles sont les opérations statistiques pouvant être appliquées aux mesures sur échelles d'intervalles. Nous faisons ici référence aux nombres de l'échelle, et non aux différences entre ceux-ci. Pratiquement toutes les procédures statistiques habituelles peuvent s'appliquer aux valeurs des échelles d'intervalles, en tout cas les procédures importantes comme *l'écart moyen*, *l'écart type*, *le coefficient de corrélation de Pearson r* et d'autres opérations statistiques qui reposent sur ces valeurs. Le seul élément statistique courant qui ne peut s'y appliquer, et qui n'est guère utilisé, est le ***coefficient de variation***. Il se fonde sur la rapport entre l'écart type et la moyenne de distribution. Sa formule est $CV = 100\sigma_x/M_x$. La raison en est que le rapport entre σ_x et M_x dépend de la position arbitraire du zéro. L'écart type est une distance fixe sur l'échelle de mesures ; il sera donc toujours le même, quelque soit l'endroit où se situera le zéro. Par contre, l'écart moyen variera en fonction de la position du zéro. Si l'on ajoute une constante à toutes les mesures, l'écart moyen sera augmenté de la même constante. Si nous déplaçons le point zéro vers le bas, la valeur de M_x va augmenter et celle de CV va chuter. Si nous déplaçons le zéro vers le haut, M_x va diminuer et CV va augmenter.

Comme nous allons examiner chaque méthode de mesure dans les chapitres suivants, nous verrons à quel type d'échelle psychologique l'on parvient dans chaque cas. Au stade actuel de notre développement, on peut dire que de nombreux chercheurs estiment que la plupart des échelles psychologiques, qui sont traitées comme s'il s'agissait d'échelles d'intervalles, sont en fait des échelles ordinales. Pourtant, les psychologues hésitent rarement à appliquer les processus statistiques qui nécessitent une échelle d'intervalles, comme s'ils considéraient qu'ils ont affaire à des mesures d'échelles d'intervalles. Ils n'ont guère conscience de l'hypothèse que cela implique une échelle d'intervalles. Cette hypothèse reste occultée, comme le sont beaucoup d'autres dans le cadre d'investigations psychologiques. Toutefois, il arrive parfois que les psychologues reconnaissent cette hypothèse et s'efforcent de faire quelque chose pour la vérifier.

Il y a plusieurs démarches pour parvenir à des unités égales. Diverses techniques d'observation, méthodes de transformation et autres procédures d'échelonnage permettent d'obtenir des unités égales. Toutefois, même si on ne les applique pas, les données expérimentales se rapprochent souvent suffisamment du critère d'égalité des unités pour que le taux d'erreur soit acceptable et permette d'appliquer les diverses procédures statistiques que ces données nécessitent. C'est l'un de ces cas où l'utilisation d'approximations, même grossières, permet de tirer un maximum d'informations de ses données. Cette démarche se justifie souvent par l'évidence d'une cohérence interne et de la validité des résultats obtenus. Néanmoins, cela ne devrait pas éviter au chercheur d'être vigilant à l'égard d'approximations inacceptables et de résultats et de conclusions provenant essentiellement d'une application inappropriée des procédures statistiques.

Encart 4 – *Les thermomètres Celsius et Fahrenheit ne possèdent pas un zéro vrai, contrairement aux thermomètres Kelvin, mais il existe un système de correspondance entre ces trois types d'appareils.* La température de fusion de la glace (0°C) correspond à 32°F et 273,15 K. L'évaporation de l'eau liquide, à pression normale, correspond, quant à elle, à 0°C, 212°F et 373,15 K. La valeur du 0 absolu correspond à 273,15°C (et donc 0 K ou -459,67°F).



3.5.4. Les échelles de rapports

Les échelles de rapports ont des zéros absolus, où le zéro ne représente ni plus ni moins que l'absence de propriété mesurée par l'échelle. Tous les postulats fondamentaux énoncés plus haut sont d'application. On peut mettre sur un même pied de façon tout à fait significative des rapports entre les nombres de l'échelle. Par exemple, le rapport 12/8 est égal au rapport 3/2, et tous deux indiquent la même relation entre deux quantités réelles. En fait, toutes les opérations fondamentales sont possibles et porteuses de sens, et toutes les opérations statistiques peuvent être appliquées, y compris le *coefficient de variation*. Il se peut qu'il soit impossible de démontrer l'additivité de façon expérimentale avec les quantités empiriques auxquelles on applique des numéros de l'échelle des rapports. Ce n'est pas fondamental. Il existe certains modes d'observation que l'on peut légitimement utiliser pour démontrer que l'échelle de rapports existe. Par exemple, la méthode de fractionnement (voir Guilford, 1954, chapitre 9), qui requiert de trouver des paires de stimuli avec un rapport simple, comme ½. Si l'opération est bien menée, ce qui peut être vérifié en examinant les mesures de cohérence interne, on obtient une échelle de rapports. Les échelles psychophysique dont les unités ont été baptisées « mel » (Stevens, 1936) et « sone » (Stevens et Volkmann, 1940) par Stevens appartiennent à cette catégorie[21]. Il existe d'autres échelles dont l'objectif est de répondre à la nécessité du principe de rapport égal[22].

Les mesures de dénombrement[23] (obtenues en comptant les objets) sont des mesures d'échelles de rapports. Il y a un zéro initial (pas d'objet), et l'on peut évoquer significativement les rapports de fréquences. Les méthodes statistiques f et N, qui sont tellement répandues dans nos calculs, comportent la propriété d'additivité et peuvent généralement être soumises à toutes les opérations numériques. Nous disons « généralement », parce que certaines utilisations des fréquences font sortir ces éléments statistiques du champ de l'échelle de rapports. Par exemple, lorsque le résultat d'un test correspond au nombre d'items auxquels on a correctement répondu, on obtient ce résultat par un comptage. Si on se limite au « nombre de réponses correctes », on obtient des valeurs de rapports. Mais ce nombre est doté d'une fonction ou d'un usage supplémentaire lorsqu'il sert à indiquer la position d'un individu sur une échelle de compétence ou d'une autre caractéristique. En fait, dans ce cas, nous parlons de deux échelles, et non d'une seule. Lorsque elles sont transférées sur une échelle de compétence, de telles valeurs de fréquences perdent leur propriété de rapport. Aucune réponse correcte peut signifier bien davantage qu'une compétence zéro. Cinquante réponses correctes signifient vraisemblablement une compétence deux fois plus importante que 25 réponses correctes. Il se peut même qu'il n'y ait pas d'unités égales, il arrive en fait rarement qu'il y ait des unités exactement égales. Il est toutefois avéré que de tels scores se rapprochent des mesures d'intervalles lorsque les tests sont longs et que les items sont bien répartis au niveau de leur difficulté[24].

3.6. Transformation et invariance des valeurs d'échelles

Il est intéressant et significatif de chercher à savoir dans quelles conditions on peut transformer des nombres sur une échelle en d'autres nombres par le biais de diverses opérations, en conservant toutefois le même degré d'exactitude. Des exemples simples de transformation sont des changements d'unités ou du point zéro. Lorsque la description reste aussi précise avec de tels changements, on dit qu'il y a « invariance ». La notion d'invariance est importante car elle est garante d'un fonctionnement correct. Elle autorise aussi la généralisation des conclusions.

C'est avec les mesures d'échelles nominales que l'on a la plus grande latitude d'opérer des transformations. Puisqu'un nombre est aussi valable que n'importe quel autre nombre pour décrire une classe d'objets, toute opération systématique laisse la structure de l'échelle –les classes et les contenus des classes– inchangée. L'égalité des objets dans les classes et les distinctions entre les classes ne sont en aucune façon perturbées.

Avec les mesures d'échelles ordinales, toute transformation qui préserve l'ordre de rangs laissera l'échelle intacte. On peut multiplier tous les nombres par une constante ou ajouter à chacun une constante, les nombres obtenus conserveront le même ordre de rangs que précédemment. On peut même élever chaque nombre au carré, extraire sa racine carrée (positive) ou chercher son logarithme. Les nombres que l'on obtiendra seront toujours correctement ordonnés. Tous ces types de changements sont appelés transformations *monotones* car les relations fonctionnelles entre les numéros de rangs d'origine et les valeurs

[21] MD : il s'agit de deux échelles relatives, respectivement, aux perceptions de l'intensité sonore et de la hauteur des sons.
[22] Se reporter au chapitre 9 de Guilford (1954).
[23] MD : voir à ce propos le chapitre consacré au dénombrement.
[24] Voir Culler (1926).

obtenues restent sur le même continuum, augmenté ou diminué. Il n'y a pas de limites supérieures ou inférieures à la fonction qui relie les deux types de valeurs.

Dans le cas des mesures d'échelles d'intervalles, les transformations autorisées ne sont pas seulement de type monotone, mais également linéaire. Les fonctions impliquant des racines, des puissances et des logarithmes ne sont, à l'évidence, pas linéaires. La transformation linéaire, exprimée par l'équation $y = a + bx$, aboutit à un changement d'unités et du point zéro. On effectue à la fois une multiplication par une constante et une addition d'une constante. Chacune de ces opérations est effectuée séparément si a ou b est égal à zéro. Les valeurs ainsi transformées se situent également sur une échelle d'intervalles. L'égalité des unités et des intervalles est maintenue, mais la position du point zéro reste arbitraire.

Pour les mesures d'échelles de rapports, il n'y a qu'un seul type de transformations qui permettra aux numéros d'échelle de décrire correctement les données, toujours sur une échelle de rapports. C'est la multiplication par une constante. C'est l'équation de type $y = bx$, où b est la constante qui multiplie ; cette constante peut être supérieure ou inférieure à l'unité, mais elle ne peut équivaloir à zéro.

3.7. Les échelles de nombres indices

Les chercheurs en psychologie et autres ont souvent affaire à des nombres que l'on appelle des valeurs indices. Celles-ci ne sont pas comme la plupart des mesures. Un exemple familier en est le QI, et un autre, le coefficient de corrélation. La division d'un âge mental par un âge chronologique, tous deux exprimés en années ou en mois, fournit un nombre qui ne se situe pas sur l'échelle des années [d'âge]. L'occurrence d'une unité commune au numérateur et au dénominateur annihile l'unité originale de l'échelle. Le résultat est un pur rapport. Ceci ne justifie pas l'hypothèse selon laquelle l'échelle de QI se situe au niveau de l'échelle de rapports ou même qu'elle a des unités égales. Sans élément probant supplémentaire, il vaut mieux la considérer comme une échelle ordinale. Si la distribution des QI d'une population donnée était de type gaussien et si la distribution obtenue dans un échantillon aléatoire simple était normale, nous aurions la preuve qu'il s'agit d'une échelle d'intervalles. Mais si l'on ne peut établir le fait que la distribution de la population concernée est normale, on ne peut se fier complètement à ce critère de preuve. Affirmer que la répartition de l'échantillon établit le profil de la population élude la question. Ce que l'on fait généralement, c'est prétendre présomptueusement que la distribution de la population est normale ; ainsi, si la distribution obtenue d'un échantillon représentatif est normale, on peut avoir une certaine confiance dans l'égalité des unités, quelque soit l'échelle utilisée.

Encart 5 – *Tableau récapitulatif (d'après Bernier et Pietrulewicz, 1997, pp. 30-31).*

Type d'échelle	Postulats, propriétés, limites	Opérations empiriques de base	Conditions d'invariance	Opérations algébriques permises	Opérations statistiques permises	Exemples
Nominal	Postulats d'égalité : Soit $a = b$ ou $a \neq b$ Si $a = b$, alors $b = a$ Si $a = b$ et $b = c$, alors $a = c$ Attribut codifiable se prêtant à des catégories exclusives. Utilisation de symboles numériques (scores) à des fins qualitatives. N'indique pas le degré de présence des attributs.	Détermination des égalités Identification Classement Dénombrement	X'=f(X) où f(X) indique toute substitution d'un symbole à un autre.	$=, \neq$	Calcul des fréquences. Identification du mode. Calcul du coefficient de contingence.	Numéros des joueurs de football.
Ordinal	Postulats hiérarchiques (relation d'ordre) : si $a > b$, alors $b \not> a$ si $a > b$ et $b > c$, alors $a > c$ L'attribut mesuré varie en intensité de + à -. Présence d'un continuum. Les symboles numériques (scores) indiquent l'ordre des individus.	Détermination du plus grand et du plus petit. Mise en ordre.	X'=f(X) où f(X) indique toute fonction monotone croissante.	$=, \neq, <, >$	Celles des échelles nominales. Calcul de la médiane. Calcul des centiles. Calcul du coefficient de corrélation de rang (ρ de Spearman). Calcul du χ carré[25].	Dureté des minéraux. Qualités de laine, de cuir et d'autres matériaux. Ordre des arrivées dans une course. Ordre des candidats à un tournois, un concours.

[25] Le χ carré peut aussi être calculé dans le cas particulier des échelles nominales qui ne comportent que deux modalités (dichotomie).

Encart 5 – *Tableau récapitulatif (d'après Bernier et Pietrulewicz, 1997, pp. 30-31).* (suite)

Type d'échelle	Postulats, propriétés, limites	Opérations empiriques de base	Conditions d'invariance	Opérations algébriques permises	Opérations statistiques permises	Exemples
Intervalles égaux	Postulats des échelles nominales et des échelles ordinales. Postulat d'addition. Les unités sont égales (report d'un étalon constant). Les symboles numériques indiquent l'ordre des individus, les distances séparant ces derniers en fonction d'un attribut, mais n'indiquent pas l'amplitude absolue de l'attribut.	Détermination de l'égalité des intervalles. Mise en ordre. Comparaison des positions.	X'=a+bX (fonction linéaire)	=, ≠, <, > + et – (appliqués aux intervalles et aux scores). x et / (appliqués aux intervalles).	Celles des échelles nominales et des échelles ordinales. Calcul de la moyenne. Calcul de l'écart-type. Calcul du coefficient de corrélation de Bravais-Pearson.	Températures exprimées en °F et °C. Energie. Dates du calendrier. Scores standard à un test de rendement.
De proportions	Propriété parfaite d'addition. Postulats des échelles nominales, ordinales et d'intervalles égaux. Présence d'un zéro absolu. Les symboles numériques fournissent les mêmes informations que celles de l'échelle à intervalles égaux et indiquent l'amplitude absolue de l'attribut.	Détermination de l'égalité des proportions. Mise en ordre. Comparaison des intervalles. Comparaison de l'amplitude absolue des attributs.	X'=aX	=, ≠, <, > +, -, x , / (appliqués aux intervalles et aux scores).	Celles des échelles nominales, ordinales et d'intervalles égaux. Calcul de la moyenne géométrique[26]. Calcul du coefficient de variation[27].	Longueur, poids, densité, résistance. Température exprimée en °Kelvin.

[26] MD : La moyenne géométrique consiste à prendre la racine n^{ième} du produit d'un ensemble de n termes. Par exemple, si on souhaite calculer la moyenne géométrique des trois valeurs suivantes : 3, 6 et 15, on calculera $\sqrt[3]{5x6x15}$. Ce type de moyenne est utile dans certains cas particuliers.

[27] MD : Le coefficient de variation est le rapport entre l'écart-type d'une distribution et sa moyenne.

Le coefficient de corrélation est un autre nombre indice. Ici aussi, les valeurs du numérateur et du dénominateur s'expriment sous la forme des unités de mesures originales, qui s'annulent. L'échelle de Pearson *r* a un point zéro porteur de sens, mais pas d'unités égales. Si on élève *r* au carré pour obtenir le coefficient de détermination, on a un nombre indice de la proportion de variance totale de l'une des deux variables, représentée par la variance de l'autre. Les théories d'analyse de fiabilité et d'analyse factorielle, se fondent généralement sur l'hypothèse que les proportions de variance se situent dans une échelle de rapports. Ainsi, r se situe principalement dans une échelle ordinale, tandis que r^2 se situe sur une échelle de rapports.

Pour les autres nombres indices, il faudra examiner les propriétés de chacun d'eux pour tirer une conclusion quant au type d'échelle de mesures qu'ils représentent. De cette conclusion dépendront le type d'opérations possibles avec le nombre indice concerné et l'interprétation à lui attribuer. Si, par nature, un nombre indice appartient à un type inférieur d'échelle, certaine transformation pourrait lui attribuer les conditions requises pour se situer sur une échelle de type supérieur, comme c'est le cas du coefficient de corrélation.

Bibliographie

Bartlett, R.J. (1939). Measurement in psychology. *Rep. Brit. Advance. Sci.* 1-20.

Bell, E.T. (1945). *The Development of Mathematics.* 2d edition. New York : McGraw-Hill.

Boring, E.G. (1950). *A History of Experimental Psychology.* 2e ed. New York : Appleton-Century-Crofts.

Boring, E.G. (1920). The logic of the normal law of error in mental measurement. *Amer. J. Psychol., 31,* 1-33.

Bridgman, P.W. (1928). *The Logic of Modern Physics.* New York : Macmillan.

Campbell, N.R. (1938). Symposium : Measurement and its importance for philosophy. *Aristotelian Soc. Suppl., 17.*

Campbell *et al.* (1940). Final report. *Advancement of Sci., 2,* 331-349.

Carnap, R. (1939). Foundations of logic and mathematices. *International Encyclopedia of Unified Sciences, 1(3).*

Charpak, G., Broch, H. (2002). *Devenez sorciers, devenez savants.* Paris : Odile Jacob.

Comrey, A.L. (1950). An operational approach to some problems in psychological measurement. *Psychol. Rev., 57,* 217-228.

Culler, E. (1926). Studies in psychometric theory. *J. exp. Psychol., 9,* 271-279.

Dantzig, T. (1939). *Number : The Language of Science.* New York : Macmillan.

Dupreel, E. (1942). *Adolphe Quételet. Pages choisies et commentées.* Bruxelles : Office de publicité.

Duyckaerts, F. (1992). *Joseph Delboeuf philosophe et hypnotiseur.* Paris : Les empêcheurs de penser en rond.

Dodge, Y. (1993). *Statistique. Dictionnaire encyclopédique.* Paris : Dunod.

Fechner, G.T. (1860). *Elemente der Psychophysik.* Leipzig : Breitkopf and Härtel. Traduction anglaise : H.E. Adler, 1966, *Elements of Psychophysics, vol. 1,* D.H. Howes and E.G. Boring (ed), New York : Rinehart and Winston.

Fechner, G.T. (1889). *Elemente der Psychophysik* (reprint). Leipzig : Breitkopf und Härtel.

Galton, F. (1908). *Memories of My Life.* London : Methuen.

Gauld, A. (1997). Joseph Delboeuf (1831-1896) : A foorerunner of modern ideas on hypnosis. *Contemporary Hypnosis, 14(4),* 216-226.

Guilford, J.P. (1950). *Fundamental Statistics in Psychology and Education.* 2d ed. New York : McGraw-Hill.

Helson, H. (1951). *Theoretical Foundations of Psychology.* New York : Van Nostrand.

Hogben, L. (1940). *Mathematics for the Millions.* New York : Norton.

Ifrah, G. (1994). *Histoire universelle des chiffres : l'intelligence des hommes racontée par les nombres et le calcul. (2 tomes).* Paris : Robert Laffont.

Levy, P.M.G. (1975). Quételet, le pète et le statisticien. In MINISTERE DES AFFAIRES ECONOMIQUES. *Adolphe Quételet 1796-1874. Recueil des travaux et contributions présentés en 1974 en hommage à son rôle de statisticien.* Bruxelles : Ministère des affaires économiques.

McCall, W.A. (1922). *Measurement. A revision of How to Measure in Education.* New York : The Macmillan Company.

Michell, J. (1999). *Measurement in Psychology. A Critical History of a Methodological Concept.* Cambridge : Cambridge University Press.

Müller, G.E. (1878). *Zur Grundlegung der Psychophysik.* Berlin : Grieben.

Nicolas, S. (1995). Joseph Delboeuf on visual illusions : A historical sketch. *American Journal of Psychology, 108(4),* 563-574.

Nicolas, S., Murray, D.J., Farahmand, B. (1997). The psychophysics of J-R-L Delboeuf (1831-2896). *Perception, 26,* 1297-1315.

Reese, T.W. (1943). The application of the theory of physical measurement to the measurement of psychological magnitudes, with three experimental examples. *Psychol. Monogr., 55(251).*

Russell, B. (1920). *Introduction to Mathematical Philosophy.* 2d ed. New York : Macmillan.

Russell, B. (1937). *The Principles of Mathematics.* New York : Norton.

Scates, D.E. (1937). How science measures. *J. exp. Educ., 5,* 296-312.

Smith, B.O. (1938). *Logical Aspects of Educational Measurement.* New York : Colombia University Press.

Spearman, C. (1937). *Psychology Down the Ages.* New York : Macmillan.

Stevens, S.S. (1936). A scale for the measurement of a psychological magnitude : loundness. *Psychol. Rev., 43,* 405-416.

Stevens, S.S. (1946). On the theory of scales of measurement. *Science, 103,* 670-680.

Stevens, S.S. (1951). *Handbook of Experimental Psychology.* New York : Wiley.

Stevens, S.S., Volkmann, J. (1940). The relation of pitch to frequency : a revised scale. *Amer. J. Psychol., 53,* 329-353.

Torndike, E.L. (1918). *The Seventeenth Year Book of the National Society for Study of Education.* Bloomington : Public School Publishing.

Titchener, E.B. (1905). *Experimental Psychology.* Vol. II, Pt. II. New York : Macmillan.

Walker, H.M. (1929). *Studies in the History of Statistical Method.* Baltimore : Willams and Wilkins.

Warusfel, A. (1961). *Les nombres et leurs mystères.* Paris : Editions du Seuil.

Weizenhoffer, A.M. (1951). Mathematical structures and psychological measurement. *Psychometrika, 16,* 387-406.

Weyl, H. (1949). *Philosophy of Mathematics and Natural Science.* Princeton, NJ : Princeton University Press.