



Norme et Contexte : Influence d'une dichotomisation du matériel et de l'évaluation sur la contextualisation de jugements catégoriels

Guylaine Molina, Jean-Marc Fabre

► To cite this version:

Guylaine Molina, Jean-Marc Fabre. Norme et Contexte : Influence d'une dichotomisation du matériel et de l'évaluation sur la contextualisation de jugements catégoriels. *L'Année psychologique*, 2000, 100 (1), pp.37-69. hal-01421091v1

HAL Id: hal-01421091

<https://amu.hal.science/hal-01421091v1>

Submitted on 27 Dec 2016 (v1), last revised 15 Jun 2017 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

*Centre de Recherche en Psychologie cognitive
UMR 6561 CNRS
Université de Provence¹*

**NORME ET CONTEXTE :
INFLUENCE D'UNE DICHOTOMISATION
DU MATÉRIEL ET DE L'ÉVALUATION
SUR LA CONTEXTUALISATION
DE JUGEMENTS CATÉGORIELS**

par Guylaine MOLINA et Jean-Marc FABRE

SUMMARY : Norm and context : Dichotomization of stimuli and response scales influences context effects in absolute judgments.

Range effects, which are contextual effects on perceptual judgments, can be subject to a limitation when a dichotomous pre-established norm is available. These limitations, initially observed by Marsh and Parducci (1978), consist of a neutral-point anchoring and of a symmetrization of the subjective range. Four experiments show that these limitations 1 / do not cancel the usual range effects, 2 / can be observed with various stimuli, and 3 / can be explained by various processes linked to response scales or to stimuli. This diversity can guide the analysis of similar effects that occur in everyday situations.

Key words : judgmental norms, context, dichotomy, subjective range.

INTRODUCTION

Le rôle du contexte dans les jugements psychophysiques et sociaux n'est plus à démontrer. Après les travaux de Helson (1947), Jonhson (1949), Parducci, Calfee, Marshall et Davidson (1960), Harvey et Campbell (1963), Manis (1967), de nombreuses

1. CREPCO, 29, avenue Robert-Schuman, 13621 Aix-en-Provence Cedex 1.
E-mail : gmolina@newsup.univ-mrs.fr

recherches ont été effectuées qui ont mis en évidence la généralité des effets de mise en référence de stimulus unidimensionnels. Cette mise en référence, qui leur assure une position relative sur un ensemble de possibles, guide l'affectation des différentes catégories de jugement. Plusieurs approches théoriques ont été élaborées pour rendre compte de cette influence contextuelle. Dans le cadre de cette étude, on se réfère prioritairement au modèle étendue-fréquence de Parducci (1965, 1974, 1983). Dans la plupart des études les plus récentes qui ont utilisé ce modèle (par ex., Mellers, 1986 ; Wedell et Parducci, 1988 ; Wedell, 1990 ; Haubensak, 1992*a* et *b* ; Fabre, 1993 ; Parducci, 1995 ; Parducci et Fabre, 1995 ; Wedell, 1995), l'aspect actif du contexte est représenté par le paramètre fréquentiel des séries de stimulus. Les manipulations expérimentales font intervenir différentes formes de distributions sur une étendue de variation donnée. Ces distributions, symétriques ou dissymétriques, sont obtenues par le choix de densités de stimulation différentes selon les zones de l'étendue. On obtient cette variation des densités soit en répétant certains stimulus plus souvent que d'autres, soit en choisissant dans certaines zones des espacements inter-stimulus plus réduits que dans d'autres zones. Indépendamment de cette variation des densités, le modèle étendue-fréquence permet aussi de prévoir et d'étudier l'effet de contextualisation produit par l'étendue de stimulation. La marge de variation des stimulus peut ainsi être manipulée de façon à faire varier sa largeur ou sa concentration sur l'échelle. Les prédictions liées à ces deux types de manipulations sont fondées sur deux principes : le principe de fréquence et le principe d'étendue.

Selon le principe de fréquence, le sujet est supposé égaliser le nombre de stimulus qu'il affecte à chacune des catégories de réponse, le résultat de cette affectation étant analogue à un étalonnage en classes d'effectifs égaux. Les intervalles de classes, selon ce principe, sont inégaux et dépendent des densités de présentation. Selon le principe d'étendue, le sujet élabore une représentation, appelée étendue subjective, des deux marges de stimulation, et la segmente en autant d'intervalles qu'il dispose de catégories de réponse. Le jugement attribué à un stimulus est alors déterminé par sa position vis-à-vis des extrêmes subjectifs.

Comme la « valeur relative » prise par un stimulus peut différer selon les deux principes, le « calcul » d'un compromis permet au sujet d'affecter à chaque stimulus, à chaque essai, une

seule position relative et donc une seule réponse. Les procédures de calcul incluses dans le modèle permettent de prédire et d'analyser ces modes de relativisation à un ensemble. Les prédictions du modèle peuvent être représentées graphiquement. Par exemple avec deux séries de forme asymétrique, gauche ou droite, construites sur une même étendue, la figure 1 expose les prédictions, dans le cas hypothétique où l'échelle des stimulus en abscisse a été ajustée de façon à être fonction affine de leurs valeurs d'échelle perceptives.

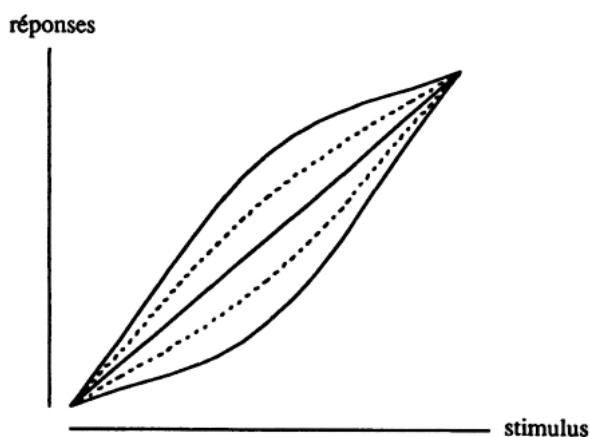


Fig. 1. — Le modèle étendue-fréquence. La fonction linéaire représente les valeurs déterminées par le principe d'étendue (après ajustement de l'échelle). Les courbes pleines représentent les valeurs déterminées par le principe de fréquence dans un contexte asymétrique gauche (courbe la plus haute) et dans un contexte asymétrique droite (courbe la plus basse). Les courbes hachurées représentent les jugements prédits par le modèle étendue-fréquence, lorsque le poids accordé aux deux principes est le même.

Range-frequency theory. The values of stimuli determined by the range principle (after rescaling) are represented by a linear function. Solid curves are the values of stimuli determined by the frequency principle in the positively skewed context (upper curve) and negatively skewed context (lower curve). Dashed curves are predictions of range-frequency theory, with the same weight accorded to the two principles.

L'influence du contexte peut subir des restrictions. L'un des cas les plus intéressants est celui de restrictions à l'effet d'étendue causées par l'intervention d'une norme stable pré-établie, reconnue comme telle par tous les sujets, et qui opère par le biais d'une dichotomisation de l'ensemble des possibles. Marsh et Parducci (1978) ont publié une étude de cette question, étude qui depuis cette date n'a pas fait l'objet de réplication. Cette lacune s'explique peut-être par le fait que ces auteurs avaient conclu qu'il s'agissait d'un effet propre à certaines caté-

gories de jugements, non généralisables, à traiter comme une exception à la loi de relativisation des jugements.

Dans l'article de Marsh et Parducci (*ibid.*), deux expériences sont relatées. L'une porte sur des jugements moraux, l'autre sur des jugements de gains et de pertes à un jeu de hasard. Les stimulus sont des descriptions de comportements dans le premier cas, des sommes positives ou négatives dans le second. Dans les deux cas, on dispose d'une norme habituelle préétablie susceptible de dichotomiser à la fois les stimulus et les jugements qui leur sont attribués, en distinguant le bien et le mal, le gain et la perte. La dichotomisation effectuée par la norme est plus aisée dans le cas des valeurs numériques (gains et pertes) que dans celui des descriptions de comportements, pour au moins deux raisons. La première est évidente, elle tient à la présence dans seulement l'un des cas d'un stimulus qui symbolise le neutre (le nombre zéro). La deuxième est plus fondamentalement liée à la nature même de ce neutre : dans le cas du jeu le zéro est le maintien d'un état stable, alors qu'il est pratiquement impossible de définir un « neutre » qui séparerait le bien du mal autrement qu'en ne lui décernant ni louange ni sanction (mais on voit bien que cette définition n'est pas opérationnelle puisqu'elle peut s'appliquer à tout objet ne relevant pas d'un jugement moral). C'est vraisemblablement pour ces raisons que les résultats de Marsh et Parducci sont plus nets dans le cas de l'expérience sur le jeu que dans celui des jugements moraux.

Dans l'étude sur la satisfaction associée aux issues d'un jeu, quatre séries d'étendue différente étaient présentées à des groupes disjoints. Deux séries étaient symétriques autour de zéro : une série étroite et une série large étirée bilatéralement. Deux séries dissymétriques, négative et positive, présentaient une extension en dessous ou au-dessus de zéro. Les distributions étaient construites de façon à ce que chacune des deux séries dissymétriques partage une borne avec la série étroite : il s'agissait de la borne inférieure pour la série positive, et de la borne supérieure pour la série négative, l'autre borne étant partagée avec la série large. Le principe d'étendue prévoyant l'ajustement de l'échelle de réponse aux marges de la série, l'ancrage devait alors être le même pour les séries positive et étroite sur la partie basse, et pour les séries négative et étroite sur la partie haute. Cet ajustement de l'échelle aurait impliqué une surévaluation systématique des stimulus présentés dans la série négative comparative-

ment à la série positive. Les résultats n'ont pas été conformes à ces prédictions : ils ont montré qu'une dissymétrie de l'étendue ne modifiait pas l'évaluation du zéro, et suscitait des réajustements particuliers sur l'échelle de réponse. En comparaison des jugements obtenus avec la série étroite, les sujets se sont déclarés moins insatisfaits par les pertes présentées dans la série positive, et moins satisfaits par les gains présentés dans la série négative. Les schémas, présentés en figure 2, permettent de comparer les habituels effets d'étendue et leur restriction observée par Marsh et Parducci (*ibid.*).

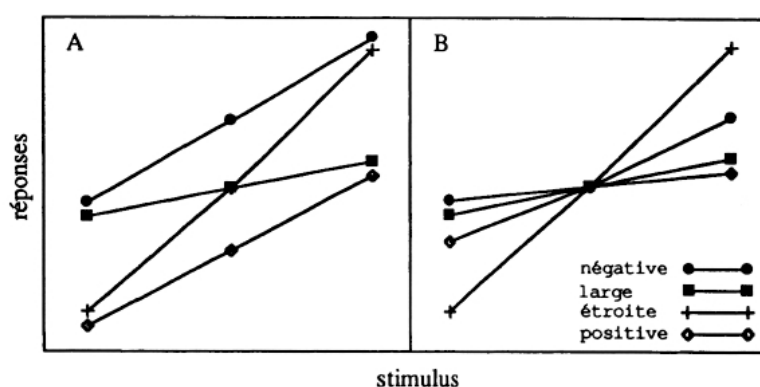


Fig. 2. — Représentation schématique des données de Marsh et Parducci (1978). Le graphique A représente les habituels effets d'étendue. Le graphique B représente un ancrage au point neutre et une équilibration de l'étendue subjective : la valeur supérieure de la série négative occupe une position subjective intermédiaire entre les valeurs correspondantes des séries étroite et large. Corrélativement, il en est de même pour la valeur inférieure de la série positive.

Schematic representation of range-frequency analysis of Marsh and Parducci (1978). Panel A shows typical adjustment to stimuli range. Panel B shows range-frequency compromise for neutral-point anchoring, under the assumption that subjective range is symmetrical around neutral point : the highest endpoint of the negative context is in an intermediate subjective position between the corresponding values of narrow and wide contexts. The same holds for the lowest endpoint of the positive context.

Ainsi, tout se passe comme si la valeur séparatrice dite neutre jouait le rôle d'une norme stable permettant de distinguer les éléments positifs et négatifs sur une dimension orientée. L'hypothèse d'une suppression de l'effet de contexte étant exclue, puisque la signification des termes d'une échelle de réponse n'est établie que par l'ajustement de cette échelle à la série de stimulus, Marsh et Parducci (*ibid.*) réunissent dans une même analyse la propriété de relativité, la stabilité du neutre et l'indifférence aux

marges effectives de l'étendue de stimulation. Rappelons que le principe d'étendue confère aux échelles de réponse des propriétés d'échelles d'intervalles. Si les sujets ajustaient l'échelle aux marges effectives de stimulation tout en la centrant sur une valeur non médiane, les valeurs des intervalles ne pourraient plus être les mêmes au-dessus et en dessous de la norme. En revanche, ils peuvent prendre le parti de sous-utiliser certaines catégories extrêmes de réponse, comme si elles correspondaient à des stimulus possibles mais non présents. C'est apparemment ce qui s'est passé dans ces expériences. L'analyse est conforme au principe d'étendue, mais introduit la spécification suivante : dans l'élaboration des marges de l'étendue subjective, qui peut reposer sur des processus d'apprentissage ou d'évocation, le rôle central attribué au neutre facilite l'évocation de stimulus absents et l'équilibration des marges de part et d'autre du neutre.

PROBLÉMATIQUE

Les expériences présentées ici ont pour objectif : 1 / de vérifier l'hypothèse du maintien d'un effet de relativisation même dans le cas d'une altération de l'ajustement à une série présente ; 2 / de généraliser cette altération à d'autres types de stimulus ; et 3 / d'en analyser le mécanisme de façon plus fine. D'une part, nous nous efforcerons d'évaluer le rôle de l'habillage : nous nous demanderons si d'autres habillages sont susceptibles de fournir des résultats du même type, et si ces résultats peuvent être obtenus en l'absence de tout habillage, avec une tâche d'évaluation de séries numériques. D'autre part, nous nous intéresserons au rôle spécifique de la nature algébrique des stimulus, qui associe à des signes positifs et négatifs des valeurs absolues aisément repérables, ce qui nous conduira à opposer cette structure à des nombres entièrement positifs. Enfin, nous étudierons le rôle joué par l'échelle de réponse, et ceci en tenant compte de deux caractéristiques différentes : la dénomination de l'échelle, liée au libellé des catégories, et la disponibilité d'une catégorie médiane facilitant le processus d'ancrage. Ces expériences nous conduiront à dissocier les effets de stabilité de la norme et d'équilibration de l'étendue, en montrant qu'ils relèvent de caractéristiques différentes et qu'ils peuvent donc être obtenus séparément. Selon

nous, le résultat de Marsh et Parducci (*ibid.*), qui conjugue stabilité du neutre et extension unilatérale du contexte, n'est pas une exception, mais un résultat général, susceptible de se reproduire dès que certaines propriétés de dichotomisation se retrouvent au niveau des stimulus et/ou des catégories de réponse.

Le tableau I propose une vue d'ensemble des expériences réalisées avec un matériel numérique qui sera, dans un premier temps, structuré en nombres relatifs. Dans les expériences 1A et 1B, les possibilités de dissociation des effets de stabilité et d'équilibration sont exploitées, à travers la mise en place de situations favorisant ou au contraire gênant l'équilibration des marges. L'expérience 2 met à l'épreuve l'ancrage de la norme sur l'échelle en l'absence de catégorie de réponse médiane. L'objectif de l'expérience 3 est de reproduire, en l'absence d'habillage explicite, la mise en place d'un critère stable de dichotomisation. L'expérience 4 achève la décomposition des effets, et met l'accent sur l'influence déterminante de la structure du matériel dans l'effet d'équilibration.

TABLEAU I. — *Présentation des conditions expérimentales*

Overview of experimental conditions

Expérience	Habillage	Présentation	Échelle
1A	notes, scores	successive	linéaire
1B	notes	simultanée	linéaire
2	notes	simultanée	linéaire à 6 catégories
3	aucun	simultanée	dichotomisée, linéaire
4	aucun	successive	linéaire

EXPÉRIENCE 1 : STABILITÉ AVEC ET SANS ÉQUILIBRATION

L'objectif principal de l'expérience 1 est de proposer une réplication des effets de stabilité et d'équilibration obtenus par Marsh et Parducci (*ibid.*), et d'élargir ainsi le champ d'application de l'effet de normes dichotomisantes préétablies. L'habillage retenu est celui de notes obtenues par des étudiants à une épreuve : il véhicule un schéma dichotomique réussite-échec, qui définit le critère de réussite comme point de repère normatif

pour l'évaluation. L'hypothèse est que l'habillage va provoquer l'application d'une règle de décision visant à différencier les deux catégories d'événements, réussite et échec, et que cette règle se traduira par une stabilité du critère de réussite.

Une autre caractéristique de cet habillage est liée au type de représentation de la notation qu'il est susceptible d'activer chez le sujet. Nous supposons l'existence d'un modèle « gaussien » de la notation centré sur la valeur du critère de réussite, tel que des notes extrêmes à chacun des deux pôles de l'échelle accompagnent toujours les notes moyennes. Cette représentation pourra conduire, en cas de dissymétrie de l'étendue, au rétablissement du modèle habituel, et donc à la mise en place d'un effet d'équilibration. Le rétablissement des caractéristiques du modèle gaussien est envisagé en relation avec l'illusion bien connue de la compensation, qui fonctionne quelle que soit la taille des échantillons (Tversky et Kahneman, 1971, 1982). Toutefois, nous faisons l'hypothèse que ce type de représentation ne pourra jouer que si le sujet ne connaît pas à l'avance l'ensemble des valeurs possibles, ce qui est le cas avec une présentation successive des stimulus. En revanche, si la présentation est simultanée, le sujet pourra constater, avant la phase d'évaluation, que la notation est biaisée, et l'effet d'équilibration sera susceptible de disparaître.

EXPÉRIENCE 1A : LA PRÉSENTATION SUCCESSIVE

La méthode de présentation successive constitue *a priori* la condition la plus favorable à une équilibration, dans la mesure où le sujet ne dispose, au moment où débute l'évaluation, d'aucune information sur l'étendue possible des notes. En conséquence, ce mode de présentation devra nous permettre de tester l'efficacité de l'habillage retenu pour obtenir les deux effets, stabilité et équilibration. Deux habillages seront utilisés : des notes à un examen et des scores à un questionnaire de type vrai-faux donnant lieu à une correction automatisée, dans laquelle la part de subjectivité du correcteur est neutralisée. S'il se manifeste, le modèle gaussien peut s'appliquer à une représentation de l'activité d'évaluation, mais il peut aussi viser une représentation des objets de cette évaluation (« compensation » objectivée), et donc concerner tout autant l'épreuve standardisée que la notation traditionnelle.

MÉTHODE

HABILLAGE

Selon la condition expérimentale, les valeurs numériques étaient identifiées soit comme des notes obtenues par des étudiants à un examen en vue de l'obtention d'un diplôme, soit comme des scores déterminés par la différence entre le nombre de bonnes réponses et le nombre d'erreurs à un questionnaire demandant des réponses de type vrai-faux. Dans tous les cas, il était fait référence à une différenciation réussite-échec, et le critère de partition était zéro.

CONTEXTE

Quatre contextes, qui différaient en fonction de l'étendue de stimulation, ont été créés : deux contextes symétriques autour du seuil de réussite (le zéro) et deux contextes dissymétriques. Les séries symétriques, l'une étroite et l'autre large, présentaient une contraction ou une extension. L'étendue étroite variait de -30 à $+30$, les stimulus contextuels occupant une position médiane sur l'échelle. L'étendue large était étirée bilatéralement, et variait de -50 à $+50$. Les séries dissymétriques par rapport à zéro, une série négative et une série positive, comportaient une extension unilatérale en dessous ou au-dessus de zéro, et se caractérisaient par un nombre différent de stimulus négatifs et positifs. La série négative variait de -50 à $+30$, et la série positive de -30 à $+50$.

STIMULUS

Chacune des quatre distributions utilisées était composée de 43 stimulus, dont 13, étirés de -30 à $+30$ et espacés d'un écart de 5 points, étaient communs à tous les contextes. Dans le choix des stimulus contextuels, c'est-à-dire spécifiques d'un contexte, une stricte répétition en miroir des valeurs absolues des nombres a été exclue. Ces stimulus n'ont donc pas été répartis de façon symétrique autour de zéro, ceci afin d'éviter un traitement exclusif des valeurs absolues.

PROCÉDURE

La passation de l'expérience était individuelle, et en début de session chaque sujet recevait les instructions spécifiques de l'habillage qui lui était affecté. La consigne informait les sujets de leur tâche, qui était d'évaluer la qualité de résultats obtenus par des étudiants à un examen ou bien à une

épreuve standardisée, selon la condition expérimentale. Quel que soit l'habillage, l'échec correspondait à des notes négatives, et la réussite à des notes positives.

L'ordre de présentation des stimulus a été déterminé de façon aléatoire à l'intérieur de deux contraintes. D'une part, les stimulus communs occupaient la même position à l'intérieur de chaque série. D'autre part, les premiers stimulus présentés étaient contextuels, puis une alternance régulière de stimulus communs et contextuels a été réalisée pour le reste de la série. Le signe des nombres, négatif bien sûr mais aussi positif, était indiqué afin que les sujets repèrent clairement la structure dichotomique en nombres relatifs de ce matériel.

Les stimulus étaient présentés sur un écran d'ordinateur de façon séquentielle, c'est-à-dire que les sujets devaient évaluer chaque stimulus successivement sans avoir pris au préalable connaissance de l'ensemble de la distribution. Dans cette situation les sujets ne disposaient, au moment où commençait l'évaluation, d'aucune information sur l'ensemble des valeurs possibles, et ne pouvaient élaborer une représentation de la série qu'à mesure que s'effectuaient leurs évaluations.

L'échelle d'évaluation utilisée était composée des cinq catégories verbales suivantes : « très mauvais », « mauvais », « moyen », « bien », « très bien ». Ces catégories étaient disponibles sur un support visuel tout au long de l'expérience.

PLAN EXPÉRIMENTAL ET SUJETS

L'habillage (notes, scores) et l'étendue (négative, large, étroite, positive) sont des facteurs inter-sujets. Les différentes valeurs des stimulus communs constituent un facteur intra-sujets. Cinquante-six étudiants du Centre littéraire de l'Université de Provence ont participé à l'expérience, et ont été équirépartis de façon aléatoire dans huit groupes disjoints.

HYPOTHÈSES

Nos hypothèses sur les éventuels effets d'équilibration et de stabilité du point de référence (le zéro) ont été mises à l'épreuve en comparant les jugements des stimulus communs à tous les contextes, ordonnés de -30 à $+30$. Elles ont porté sur les séries dissymétriques (négative et positive), les séries symétriques (étroite et large) servant de conditions contrôles. Les traitements statistiques ont été effectués après avoir codé numériquement (1, 2, 3, 4, 5) l'échelle d'évaluation.

HYPOTHÈSE 1 : EFFETS D'ÉTENDUE

Selon le principe d'étendue (Parducci, 1983), les catégories d'évaluation extrêmes sont supposées s'appliquer aux limites de l'étendue. En conséquence, les jugements des stimulus communs doivent couvrir la totalité de l'échelle de réponse avec la série étroite, et occuper une place plus médiane avec la série large en raison de son extension bilatérale. Puisque chacune des deux séries dissymétriques partage une borne avec la série étroite et une borne avec la série large, on peut s'attendre à ce que l'ajustement de l'échelle de réponse à ces séries reproduise l'ajustement à la série étroite ou à la série large selon la borne concernée. Dans cette hypothèse, l'ancrage sur la partie basse de l'échelle devra être le même pour la série positive et la série étroite d'une part, et pour la série négative et la série large d'autre part. De même, sur la partie haute de l'échelle, l'ancrage devra être le même pour la série négative et la série étroite d'une part, et pour la série positive et la série large d'autre part.

HYPOTHÈSE 2 : EFFET D'ÉQUILIBRATION

Un éventuel effet d'équilibration de l'étendue pourra être observé en comparant les évaluations obtenues avec les séries dissymétriques et la série étroite. Dans cette hypothèse, la présentation d'une valeur extrême dans un sens suscitera l'évocation de son opposé symétrique dans le sens inverse. Les stimulus bas seront alors surévalués dans la série positive comparativement à la série étroite, et les stimulus élevés sous-évalués dans la série négative comparativement à la série étroite. La méthode choisie pour tester cet effet est globale : pour chacune des séries, on calculera la moyenne des jugements portés sur les six stimulus négatifs (-30 , -25 , -20 , -15 , -10 , -5), ainsi que celle des jugements portés sur les six stimulus positifs ($+5$, $+10$, $+15$, $+20$, $+25$, $+30$). Il sera alors possible de comparer à la série étroite la série positive pour ce qui est de la partie basse de l'échelle, et la série négative pour ce qui est de la partie haute de l'échelle.

HYPOTHÈSE 3 : EFFET DE STABILITÉ DE LA NORME

La stabilité du critère de réussite à l'épreuve (le zéro) sera testée en comparant les jugements moyens attribués à ce stimulus avec les séries dissymétriques. L'ajustement de l'échelle de réponse aux bornes effectives de l'étendue implique une surévaluation de chacun des stimulus présentés dans une série négative comparativement à une série positive. En revanche, si l'effet du critère de dichotomisation se traduit par un ancrage de cette valeur au centre de l'échelle, son évaluation restera stable quel que soit le sens de la dissymétrie, alors que les effets d'étendue joueront sur le reste de la série. Dans ce cas, il y aura une interaction entre l'ensemble des stimulus communs et le contexte défini par les séries dissymétriques.

L'objectif sera alors de mettre à l'épreuve la stabilité du zéro, et de faire la démonstration d'une absence de différence entre les évaluations moyennes obtenues avec les séries dissymétriques. Pour traiter cette question, nous avons retenu une procédure fiduciaire (Lépine et Rouanet, 1975), qui permettra d'inférer sur l'importance de l'écart à l'hypothèse nulle et de conclure, le cas échéant, à un effet négligeable de la dissymétrie. La recherche d'une conclusion d'effet négligeable passe par le choix d'un critère arbitraire défini *a priori*. Le critère que nous avons retenu est le suivant : un écart entre évaluations moyennes est négligeable s'il n'excède pas 5 % de l'échelle de réponse (0,20). Appliqué à la comparaison des séries dissymétriques, ce critère permet, grâce à la distribution *t* de Student, de définir le taux de confiance unilatéral avec lequel il est possible de dire que la valeur vraie de la différence entre les moyennes ne dépasse pas le critère retenu. Cette recherche unilatérale part du postulat selon lequel tout écart entre les évaluations obtenues avec les séries dissymétriques ira dans le sens habituel des effets de contraste, avec une surévaluation de la série négative comparativement à la série positive.

RÉSULTATS

La figure 3A représente les résultats obtenus, les deux habillages confondus, et montre une stabilité du zéro et une équilibration lorsque les séries sont dissymétriques.

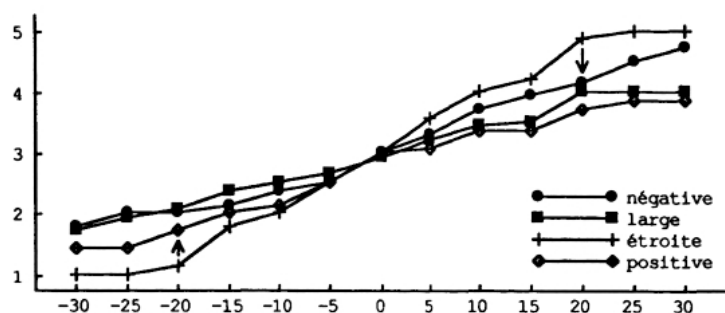


Fig. 3A. — Expérience 1A. Présentation successive. Interaction entre l'étendue (négative, large, étroite, positive) et les stimulus : jugements moyens des stimuli communs aux quatre séries, les deux habillages confondus. L'équilibration est indiquée par les flèches.

Experiment 1A. Successive presentation. Interaction between Range (negative, wide, narrow, positive) and Stimuli : mean ratings of stimuli common to the four sets, the two presentations combined. Arrows indicate the symmetrizing of the endpoints.

La série négative ($M = 3,08$) est globalement surévaluée comparativement à la série positive ($M = 2,73$) ($F(1,48) = 28,48$, $p < .0001$), et cet effet de la dissymétrie est plus marqué aux extrémités de l'étendue qu'au niveau du zéro ($F(12,576) = 3,15$, $p = .0002$). Une analyse spécifique des évaluations du stimulus zéro ne montre aucun effet contextuel significatif. La différence entre les jugements moyens obtenus avec les séries dissymétriques ($M = 3,00$) est nulle, et une inférence fiduciaire permet de conclure, avec une garantie de 99,7 %, que l'éventuel effet de la dissymétrie est négligeable.

L'effet d'équilibration est vérifié au niveau statistique, puisque les séries dissymétriques donnent lieu à des évaluations significativement moins extrêmes que la série étroite, aussi bien sur la partie qui porte l'extension que sur la partie opposée. Une analyse globale des jugements des six stimulus négatifs montre que ceux-ci sont surévalués lorsqu'ils sont présentés dans la série positive ($M = 1,87$) comparativement à la série étroite ($M = 1,57$) ($F(1,48) = 9,24$, $p = .004$). De même, les six stimulus positifs sont globalement sous-évalués dans la série négative ($M = 4,05$) en comparaison de la série étroite ($M = 4,44$) ($F(1,48) = 10,27$, $p = .002$).

Le type d'habillage ne modifie aucun des résultats observés. Ce facteur ne produit pas de différence significative dans l'équilibration de l'étendue, ni dans la stabilité du zéro. Seule une tendance non significative à un effet principal de la nature des stimulus apparaît, tel que les notes ($M = 2,98$) sont globa-

lement surévaluées par rapport aux scores ($M = 2,89$) ($F(1,48) = 3,03$, $p = .088$). Ce second habillage, peut-être parce qu'il rend possibles des réponses correctes dues au hasard, a suscité une moindre indulgence des sujets dans leurs évaluations.

DISCUSSION

Cette expérience démontre l'efficacité de l'habillage retenu pour établir à la fois un effet de stabilité et un effet d'équilibration, et nous permet donc d'affirmer que la stabilité de normes préétablies n'est pas un résultat aussi rare que Marsh et Parducci (*ibid.*) l'avaient suggéré. La confusion entre la dichotomisation de surface du matériel (nombres positifs et négatifs) et celle de l'évaluation (réussites et échecs) entraîne l'ancrage du critère de réussite à la catégorie de réponse médiane, et suscite une représentation équilibrée de l'étendue de part et d'autre du critère. La conjonction des deux effets ne contredit pas le principe d'étendue, puisque la stabilité de la norme est compatible avec sa position subjective centrale dans la série.

Comme c'était déjà le cas dans la recherche de Marsh et Parducci, la coexistence d'un effet d'équilibration et d'un effet de contexte résiduel implique que l'équilibration est partielle. La présentation d'une valeur extrême dans un sens rend possible l'évocation de son extrême symétrique dans le sens opposé, mais cette éventualité semble être considérée comme exceptionnelle. Ce résultat suggère la combinaison de deux mécanismes différents : la contextualisation par l'étendue de la série, et l'évocation de marges équilibrées de part et d'autre de la norme. Si les effets de stabilité et d'équilibration ont à voir avec un modèle gaussien de la notation, c'est avec une acceptation objectiviste d'un tel modèle, puisque les effets s'observent autant dans le cas des scores que dans celui des notes.

EXPÉRIENCE 1B : LA PRÉSENTATION SIMULTANÉE

L'extension de l'étendue est vraisemblablement favorisée par l'ignorance des valeurs possibles, liée à la présentation successive. Ce type de présentation peut induire chez le sujet un

état d'incertitude, qu'il comble grâce à sa représentation de la distribution habituelle des notes. En revanche, la connaissance immédiate des limites de l'étendue, apportée par une présentation simultanée, rend cette représentation non pertinente, puisque le sujet aura alors pu conclure que le groupe qui a passé l'épreuve n'était pas homogène, ou que l'épreuve était inadaptée. Cette situation nous offre une occasion privilégiée de test de l'une de nos hypothèses, puisque si on y observe à la fois le maintien d'un effet de stabilité et l'absence d'une équilibration des marges, alors on pourra remettre en cause l'interprétation de la stabilité comme due à la position centrale, d'équilibre, de la norme sur l'étendue subjective.

Le second objectif de l'expérience 1B est d'aller plus loin dans l'étude du phénomène de stabilité de la norme, en examinant l'influence de la dénomination des catégories de réponse. Avec la présentation successive, l'échelle utilisée était verbale, et on ne peut exclure que le terme « moyen », qui est associé à la catégorie que les sujets affectent préférentiellement au critère de réussite, ait joué un rôle déterminant dans le résultat. Pour y voir plus clair, nous avons retenu deux échelles de réponse : l'échelle verbale déjà utilisée en successif et une échelle codée numériquement.

MÉTHODE

MATÉRIEL

Les quatre distributions (négative, large, étroite, positive) ont été présentées, et seul l'habillage « notes » a été reproduit.

ÉCHELLE D'ÉVALUATION

Deux systèmes de réponse en cinq catégories ont été proposés pour évaluer les notes, l'un verbal et l'autre en graduations numériques. Les catégories de l'échelle verbale sont celles qui ont déjà été utilisées avec la présentation successive. Ces catégories ont été supprimées pour l'échelle numérique, ordonnée de 1 à 5, avec les exceptions suivantes : 1 correspondait à « très mauvais », et 5 à « très bien ».

PROCÉDURE

La passation de l'expérience s'est déroulée dans des conditions analogues à celles de l'expérience 1A. La présentation simultanée des stimulus a été organisée de la façon suivante : dans un premier temps, les sujets pouvaient consulter la série aussi longtemps qu'ils le désiraient. Ensuite, l'évaluation de chaque stimulus s'effectuait de façon séquentielle, et dans un ordre identique à celui de leur présentation initiale. A cette phase de l'expérience, les sujets n'avaient plus la possibilité de revenir en arrière pour explorer de nouveau la série.

PLAN EXPÉRIMENTAL ET SUJETS

Les 56 étudiants, qui ont participé à l'expérience, ont été équirépartis en huit groupes correspondant au croisement des deux facteurs inter-sujets, étendue (négative, large, étroite, positive) et échelle d'évaluation (verbale, numérique).

HYPOTHÈSES

L'objectif de la mise en place d'une présentation simultanée est de mettre à l'épreuve la possible dissociation des effets de stabilité et d'équilibration. Dans cette situation, les conditions supposées favorables à la stabilité, structure dichotomique du contenu et de l'évaluation, et disponibilité d'une catégorie médiane de réponse, sont préservées. En revanche, la prise d'information immédiate sur l'ensemble des valeurs possibles pourra rendre inadaptée l'application d'un modèle gaussien de la notation. En conséquence, on peut s'attendre à ce que l'effet de stabilité du critère de réussite (le zéro) se maintienne, mais ne s'accompagne pas d'une équilibration subjective de l'étendue.

L'utilisation de deux systèmes de réponse, l'un verbal et l'autre codé numériquement, permettra de tester l'hypothèse selon laquelle la stabilité du critère de réussite (la « moyenne » ?) est provoquée par un effet linguistique de surface.

RÉSULTATS

La figure 3B représente les résultats obtenus, les deux échelles d'évaluation confondues : elle illustre un effet de la dissymétrie et le maintien de la stabilité du zéro en l'absence d'équilibration.

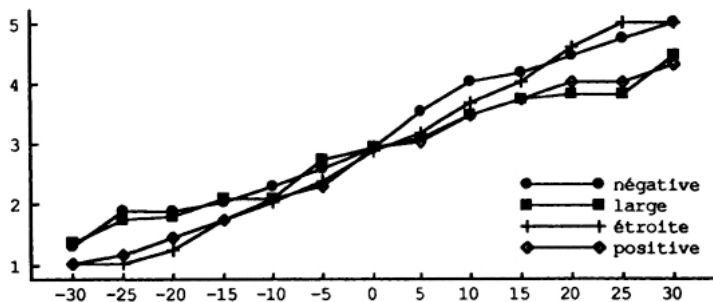


Fig. 3B. — Expérience 1B. Présentation simultanée. Interaction entre l'étendue (négative, large, étroite, positive) et les stimulus : jugements moyens des stimulus communs aux quatre séries, les deux échelles d'évaluation confondues.

Experiment 1B. Simultaneous presentation. Interaction between range (negative, wide, narrow, positive) and stimuli : mean ratings of stimuli common to the four sets, the two rating scales combined.

L'effet du sens de la dissymétrie est globalement significatif (négative : $M = 3,12$, positive : $M = 2,69$; $F(1,48) = 40,88$, $p < .0001$), de même que son interaction avec les stimulus ($F(12,576) = 2,42$, $p = .005$). Le stimulus zéro n'est pas influencé de façon significative par l'extension unilatérale de l'étendue, et reçoit des évaluations identiques ($M = 2,93$), qu'il soit présenté dans une série négative ou positive. Une inférence fiduciaire permet d'établir, avec une garantie de 90 %, qu'à son niveau l'effet du sens de la dissymétrie est négligeable. L'effet d'équilibration ne se retrouve pas ici, les limites de l'étendue de variation étant ancrées aux catégories extrêmes de réponse. Sur la partie basse de l'échelle, l'ancrage est le même pour la série positive et la série étroite, et sur la partie haute de l'échelle, il est le même pour la série négative et la série étroite. L'échelle d'évaluation ne produit aucun effet significatif, ni global ni d'interaction avec le contexte ou les stimulus, et ne modifie pas la mise en place de l'effet de stabilité.

DISCUSSION

La dissociation des effets de stabilité et d'équilibration, avec une présentation simultanée, permet de réfuter l'interprétation initialement proposée par Marsh et Parducci (*ibid.*). Le critère de réussite n'occupe pas ici une position subjective centrale, et l'interprétation de sa stabilité comme due à un rééquilibrage de l'étendue n'est en conséquence pas fondée. Ce résultat n'est pas compatible avec l'application du principe d'étendue (Parducci, 1983), qui prévoit que chaque catégorie de réponse couvre une portion équivalente de l'étendue subjective. En effet, l'ancrage du centre de l'échelle à la norme, qui n'est pas en position médiane dans les séries dissymétriques, conduit à un découpage de l'étendue en intervalles inégaux : la largeur des intervalles est réduite au-dessus de zéro dans la série négative, et en dessous de zéro dans la série positive. Avec ce résultat, nous mettons l'accent sur l'aspect décisionnel impliqué dans le jugement, qui différencie ici la réussite de l'échec, et se surimpose à des ajustements entre des degrés de réussite ou d'échec, qui eux sont sensibles à la composition de la série.

Le maintien d'un effet de stabilité pour chacun des deux systèmes de réponse utilisés montre que la suppression de la dénomination des catégories, avec l'échelle en graduations numériques, n'est pas suffisante pour empêcher la mise en place de l'effet. Ce résultat ne permet toutefois pas de conclure à une absence d'effet du terme « moyen », qu'il soit effectivement présent sur l'échelle ou seulement évoqué à partir du format de l'échelle. Le fonctionnement effectif des deux échelles est le même, l'association privilégiée entre la norme et la réponse médiane étant dans tous les cas maintenue. C'est cette association qui sera mise à l'épreuve dans l'expérience 2, grâce à l'utilisation d'une échelle composée d'un nombre pair de catégories.

Dans la recherche de Marsh et Parducci (*ibid.*), l'effet d'équilibration de l'étendue apparaît avec une présentation simultanée des stimulus, alors que nous ne l'observons ici qu'avec une présentation successive. Dans cette situation, on peut supposer que l'équilibration reste pertinente pour chacun des deux habillages utilisés par Marsh et Parducci. Dans le cas

des jugements moraux, le groupe qui a commis les comportements présentés n'est pas spécifié, et on peut faire l'hypothèse que l'effet d'équilibration est une façon de rétablir la palette des comportements possibles, toutes catégories sociales confondues. La situation de jeu semble également favorable à une équilibration des gains et des pertes puisqu'elle rétablit la logique habituelle des jeux, qui sont pour la plupart à somme nulle (Bresson, 1965). Avec un habillage de type « résultats scolaires », la connaissance des valeurs effectives, que permet la présentation simultanée, conduit le sujet à réajuster son modèle de la notation. Il est vraisemblable qu'en présentation successive un tel réajustement serait possible, mais qu'il demanderait un temps plus long et une accumulation d'informations pour se produire.

EXPÉRIENCE 2 : STABILITÉ ET ANCRAGE MÉDIAN

L'objectif de l'expérience 2 est d'étudier le rôle joué par la disponibilité d'une catégorie médiane de réponse dans la mise en place de la stabilité d'une norme. En effet, dans les expériences 1A et 1B comme dans celles de Marsh et Parducci (*ibid.*), le système de réponse utilisé comporte un nombre impair de catégories. Il est possible que la catégorie médiane soit automatiquement réservée au critère de réussite, et qu'en outre elle serve de support à une dénomination implicite comme « moyenne ». Cette possibilité est mise à l'épreuve dans l'expérience 2 par l'utilisation d'un système de réponse composé d'un nombre pair de catégories, l'habillage étant inchangé et préservant la dichotomie fondamentale des évaluations en termes de réussite et échec.

MÉTHODE

MATÉRIEL

L'habillage qui a été utilisé est celui de notes obtenues par des étudiants à un examen, pour lequel la réussite est acquise si la note est positive. Ces notes ont été présentées à l'intérieur des mêmes séries que dans les expériences 1A et 1B.

PROCÉDURE

La méthode de présentation des stimulus qui a été retenue est, comme dans l'expérience 1B, simultanée. L'échelle d'évaluation proposée était codée numériquement, et composée de six catégories ordonnées de 1 à 6, où 1 correspondait à « très mauvais » et 6 à « très bien ».

PLAN EXPÉRIMENTAL ET SUJETS

Quarante étudiants ont été équirépartis en quatre groupes, chaque groupe étant confronté à l'une des étendues (négative, large, étroite, positive).

HYPOTHÈSE

Selon l'hypothèse d'un effet automatique de la catégorie médiane, le nombre pair de catégories de réponse pourrait se traduire par un effet de la dissymétrie des marges sur l'évaluation du zéro.

RÉSULTATS

Les résultats sont présentés dans la figure 4. Ils illustrent un effet global de la dissymétrie (négative : $M = 3,51$, positive : $M = 2,92$; $F(1,36) = 26,71$, $p < .0001$), une absence d'équilibration, ainsi qu'une interaction entre les stimulus et les séries dissymétriques : l'effet de la dissymétrie est nettement plus marqué aux extrémités de l'étendue qu'autour de zéro ($F(12,432) = 4,09$, $p < .0001$).

L'effet de la dissymétrie n'est significatif pour aucune des valeurs comprises entre -10 et $+5$, alors qu'il l'est pour toutes les autres valeurs. Sur le stimulus zéro, cet effet est réduit (négative : $M = 3,60$, positive : $M = 3,30$; $F(1,36) = 1,91$, $p = .18$), mais pas annulé. La méthode fiduciaire ne peut pas être appliquée à cette situation, puisque la différence entre les moyennes obtenues avec les séries dissymétriques ($0,30$) dépasse le critère que nous avons retenu *a priori* (5 % de l'échelle de réponse, donc $0,25$) pour conclure à un effet négligeable.

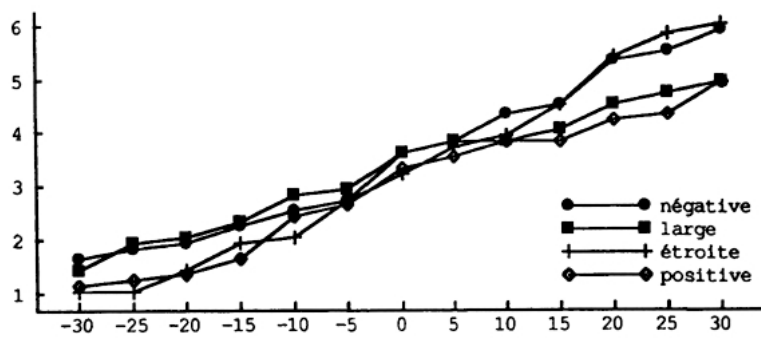


Fig. 4. — Expérience 2. Interaction entre l'étendue (négative, large, étroite, positive) et les stimuli : jugements moyens des stimuli communs aux quatre séries

Experiment 2. Interaction between range (negative, wide, narrow, positive) and stimuli : mean ratings of stimuli common to the four sets

DISCUSSION

Conformément aux considérations portant sur l'ancrage médian des normes, les résultats montrent que l'absence de catégorie de réponse médiane gêne la mise en place de l'effet de stabilité du zéro, mais elle ne suffit pas pour rétablir un effet significatif de la dissymétrie. En fait, l'évaluation du zéro donne lieu à une variabilité interindividuelle élevée : l'analyse des protocoles montre que, dans cette situation, la stabilité de la norme se manifeste, mais au choix des sujets. Selon les cas, c'est la catégorie de réponse 3 ou la catégorie 4 qui est associée au zéro. Le choix de cette catégorie est influencé par le contexte : face au zéro, 70 % des sujets répondent « 3 » si la série est positive, 40 % si elle est négative. Ne disposant pas de la catégorie médiane d'usage universel, les sujets prennent ainsi en compte les informations fournies par la série dans son ensemble pour associer au zéro l'une ou l'autre des catégories du milieu de l'échelle. La catégorie médiane facilite certainement la stabilité de la norme, mais lorsqu'elle manque, la tendance à un ancrage médian de la norme se maintient, accompagnée d'un choix et d'une forte variabilité interindividuelle.

EXPÉRIENCE 3 : STABILITÉ ET JUGEMENT D'ÉCART

Dans les expériences 1 et 2, l'habillage des données produit une dichotomisation du matériel et l'élaboration d'une norme fonctionnelle qui structure le jugement. L'objectif de l'expérience 3 est de montrer que ce type de structuration de l'échelle et l'effet de stabilité qui en résulte peuvent être reproduits, en l'absence d'habillage explicite, dès qu'un repère est explicitement fixé par la consigne d'emploi de l'échelle. L'hypothèse est que le fonctionnement de la norme est comparatif : elle fournit un critère stable par rapport auquel chaque stimulus est évalué. En conséquence, une situation dans laquelle la tâche proposée au sujet consiste à effectuer un jugement d'écart devra provoquer un fonctionnement analogue à celui de la norme, et permettra de retrouver la caractéristique de stabilité. Pour tester cette hypothèse, nous avons mis au point un dispositif méthodologique susceptible de provoquer un jugement d'écart à une valeur retenue arbitrairement.

Si cette hypothèse est exacte, peut-on aller plus loin et supposer que la structure dichotomique du matériel en nombres relatifs est nécessaire pour provoquer l'effet ? Pour répondre à cette question, nous comparerons les effets obtenus avec un matériel dichotomisé et un matériel non dichotomisé.

MÉTHODE

MATÉRIEL

Deux types de matériel ont été utilisés : des nombres relatifs et des nombres entiers. Le matériel « nombres relatifs » est celui qui a été employé dans les expériences précédentes. Le matériel « nombres entiers » a été créé en ajoutant 57 à chacune des valeurs du premier matériel. Le nombre 57 a été arbitrairement choisi comme un nombre quelconque, afin que les distributions ainsi obtenues ne suggèrent aucune dichotomie.

CONTEXTE

Trois des distributions des expériences 1 et 2 ont été utilisées : la série contrôle étroite et les deux séries dissymétriques, négative et positive. Avec le matériel composé de nombres entiers, la série étroite variait de 27 à 87, la série négative de 7 à 87 et la série positive de 27 à 107.

ÉCHELLE D'ÉVALUATION

Deux échelles d'évaluation, composées de cinq catégories alphabétiques de A à E, ont été utilisées : une échelle susceptible d'induire un repère dichotomisant et une échelle de grandeur linéaire. Avec l'échelle à repère dichotomisant, le jugement devait exprimer l'écart à la valeur de référence choisie arbitrairement : zéro dans le cas de nombres relatifs, et 57 dans le cas de nombres entiers. Dans cette situation, la réponse A devait s'appliquer aux nombres jugés « très négatifs » ou « très inférieurs à 57 », selon le matériel, et la réponse E aux nombres jugés « très positifs » ou « très supérieurs à 57 ». Lorsque l'échelle était linéaire, aucune valeur de référence n'était fournie, la tâche consistait à évaluer la grandeur des nombres sur une échelle de A à E, où A devait être utilisé pour qualifier les nombres jugés les plus petits, et E les nombres jugés les plus grands.

PROCÉDURE

Les conditions de passation de l'expérience ont été les mêmes que dans les expériences précédentes. Les stimulus étaient présentés comme des nombres sans signification, dont il s'agissait d'évaluer la grandeur, sur l'une ou l'autre des deux échelles. Le type de présentation du contexte qui a été retenu est simultané : les sujets avaient donc la possibilité, avant la phase d'évaluation, de prendre connaissance de l'ensemble de la série.

PLAN EXPÉRIMENTAL ET SUJETS

Cent vingt étudiants ont participé à l'expérience. Ils ont été équirépartis dans l'une des 12 conditions expérimentales, définies en fonction du croisement des trois facteurs inter-sujets : matériel (nombres relatifs, nombres entiers), étendue (négative, étroite, positive) et échelle d'évaluation (à repère dichotomisant, linéaire).

HYPOTHÈSES

Le dispositif méthodologique que nous proposons devra permettre de démontrer que le jugement d'écart reproduit le fonctionnement stable de la norme, et donc qu'une interprétation de l'effet de stabilité basée sur la mise en place d'un jugement d'écart est plausible. L'objectif est de reproduire avec l'échelle à repère dichotomisant un effet de stabilité analogue à celui des expériences 1 et 2. L'utilisation de deux types de matériel, des

nombres relatifs et des nombres entiers, permettra d'étudier les possibilités d'application d'un jugement d'écart sur la base d'une simple convention, sans l'appui d'une dichotomisation de surface des stimulus.

RÉSULTATS

Les données recueillies avec l'échelle à repère dichotomisant et avec l'échelle linéaire ont fait l'objet de traitements disjoints. Dans les deux cas, la présentation simultanée n'a entraîné aucun effet d'équilibration, ce qui confirme les résultats de l'expérience 1.

Jugements d'écarts

La figure 5 représente les jugements moyens obtenus avec l'échelle à repère dichotomisant, les deux types de matériel confondus. Dans ce cas, la valeur de référence (0 ou 57) reste stable quel que soit le sens de l'extension de l'étendue.

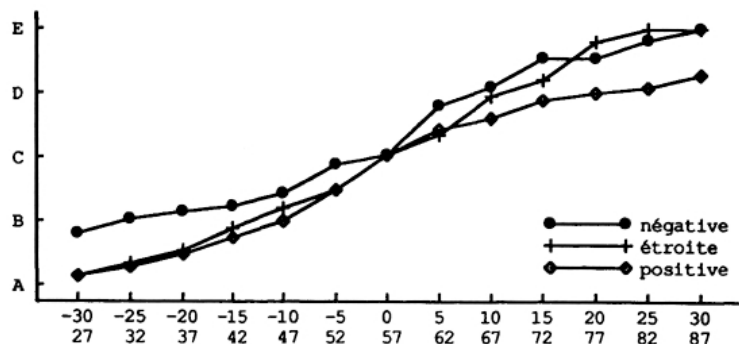


Fig. 5. — Expérience 3. Échelle à repère dichotomisant.
Interaction entre l'étendue (négative, étroite, positive)
et les stimulus : jugements moyens des stimulus communs
aux trois séries, les deux types de matériel confondus

Experiment 3. Dichotomous rating scale.
Interaction between range (negative, narrow, positive)
and stimuli : mean ratings of stimuli common
to the three sets, the two materials combined

L'effet de l'étendue est significatif : la série négative ($M = 3,30$) étant globalement surévaluée comparativement à la série positive ($M = 2,77$) ($F(1,54) = 62,18, p < .0001$). Une inter-

action entre les stimulus et les séries dissymétriques apparaît, telle que l'effet de la dissymétrie est plus marqué aux extrémités de l'étendue qu'au niveau du point de référence ($F(12,648) = 2,97, p = .0005$). Le sens de la dissymétrie ne modifie pas de façon significative l'évaluation du repère (0 ou 57 : $M = 3,00$), et cet effet peut être considéré comme négligeable avec une garantie de 95 %. Le type de matériel utilisé ne modifie aucun des effets obtenus : l'interaction entre les stimulus et les séries dissymétriques a rigoureusement la même allure que les nombres soient relatifs ou entiers. Ce résultat confirme d'une part la mise en place d'un jugement d'écart au repère, aussi bien sur un matériel déjà relatif que sur la base de l'application d'une règle de décision purement conventionnelle, et d'autre part la stabilité de ce repère.

Échelle linéaire

L'échelle de grandeur linéaire produit des effets contextuels conformes au principe d'étendue (Parducci, 1983), et ceci quel que soit le type de matériel. La série négative ($M = 3,30$) est globalement surévaluée comparativement à la série positive ($M = 2,62$) ($F(1,54) = 69,10, p < .0001$), et cet effet de contraste est significatif pour chacune des valeurs des stimulus. La valeur 0 ou 57, par exemple, est effectivement surévaluée lorsqu'elle est présentée dans la série négative ($M = 3,35$) comparativement à la série positive ($M = 2,70$) ($F(1,54) = 19,17, p < .0001$).

DISCUSSION

Le dispositif méthodologique, mis en place avec l'échelle à repère dichotomisant, permet de retrouver, en l'absence de contenu, des résultats analogues à ceux qui sont observés lorsque l'habillage induit une norme de dichotomisation. Cette similitude des effets appuie l'interprétation selon laquelle la stabilité peut être provoquée par l'intermédiaire d'un jugement d'écart. Cette dichotomisation de l'évaluation entraîne un ancrage stable et se surimpose à des ajustements plus précis, qui rendent compte des écarts à l'intérieur de chaque classe d'événements.

EXPÉRIENCE 4 : ÉQUILIBRATION SANS STABILITÉ

L'objectif de l'expérience 4 est de poursuivre, en l'absence d'habillage explicite, la décomposition des effets de stabilité et d'équilibration, en analysant les facteurs impliqués dans le rééquilibrage de l'étendue. L'hypothèse, qui sous-tend cette expérience, est que la structure dichotomique en nombres relatifs du matériel peut produire une équilibration des extrêmes, à la seule condition que l'étendue effective ne soit pas connue, c'est-à-dire avec une présentation successive. Dans cette option, l'effet d'équilibration nécessiterait simplement que soit identifiée la structure du matériel, et résulterait d'un appariement des valeurs absolues. La présence d'un repère stable, qu'il soit induit par l'habillage des données ou l'échelle d'évaluation, n'aurait aucune influence dans la mise en place de cet effet. Rappelons que la première interprétation de la stabilité des normes, donnée par Marsh et Parducci (*ibid.*), reposait sur la mise en place d'une équilibration de l'étendue de part et d'autre de la norme. Nous avons démontré par la suite que l'effet de stabilité pouvait être dissocié de celui d'équilibration. L'expérience 4 s'appliquera, à l'inverse, à provoquer une équilibration de l'étendue de part et d'autre d'une valeur qui ne fera plus office de norme stable. Cette alternative implique qu'il n'est pas exclu que le sujet équilibre les extrêmes autour de zéro, sans ancrer l'échelle de réponse sur ce même zéro. Si notre hypothèse est exacte, ce paradoxe devra en parfaire la démonstration.

MÉTHODE

MATÉRIEL

Pour tester l'influence de la structure dichotomique du matériel dans l'équilibration de l'étendue, les nombres relatifs ont été présentés en l'absence d'habillage. Trois distributions ont été utilisées : la série contrôle étroite et les deux séries dissymétriques, négative et positive.

PROCÉDURE

La méthode de présentation du contexte qui a été retenue est successive : la phase d'évaluation débutait donc sans que les sujets aient eu accès

à l'ensemble de la distribution. La tâche était d'évaluer les stimulus présentés comme des nombres sans signification, sur l'échelle de grandeur linéaire en cinq catégories déjà utilisée dans l'expérience 3.

PLAN EXPÉRIMENTAL ET SUJETS

Les 30 étudiants qui ont participé à l'expérience ont été équirépartis en trois groupes : étendue négative, étroite ou positive.

HYPOTHÈSE

L'hypothèse testée ici est que la conjonction de deux facteurs, l'incertitude sur les extrêmes induite par la présentation successive et la structure dichotomique du matériel, provoquera un rééquilibrage de l'étendue en l'absence d'une stabilité du zéro. Cette hypothèse s'appuie sur les résultats des expériences précédentes : l'expérience 1 en ce qui concerne la facilitation de l'effet d'équilibration avec une présentation successive, l'expérience 3 pour ce qui est de l'effet de l'échelle de grandeur linéaire, qui permet d'obtenir une modification de l'évaluation du zéro en fonction de l'étendue.

RÉSULTATS

Paradoxalement, mais conformément aux considérations portant sur l'effet de la structure du matériel et de l'incertitude concernant son étendue, la figure 6 illustre à la fois une équilibration des extrêmes et une variabilité de l'évaluation du zéro.

L'effet global de la dissymétrie est significatif : la série négative ($M = 3,06$) étant surévaluée par rapport à la série positive ($M = 2,44$) ($F(1,27) = 12,62$, $p = .001$). Le stimulus zéro est nettement influencé par le sens de la dissymétrie (négative : $M = 3,10$, positive : $M = 2,30$; $F(1,27) = 10,05$, $p = .004$). L'équilibration de l'étendue est vérifiée au niveau statistique, puisque, en comparaison de la série étroite, la série positive donne lieu à une surévaluation globale des six stimulus négatifs (posi-

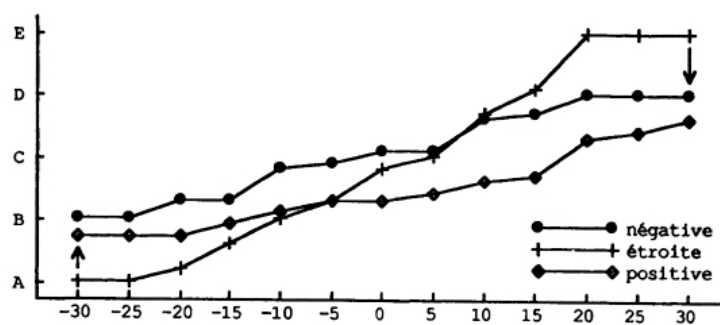


Fig. 6. — Expérience 4. Interaction entre l'étendue (négative, étroite, positive) et les stimulus : jugements moyens des stimulus communs aux trois séries. L'équilibration est indiquée par les flèches

Experiment 4. Interaction between range (negative, narrow, positive) and stimuli : mean ratings of stimuli common to the three sets. Arrows indicate the symmetrizing of the endpoints

tive : $M = 1,90$, étroite : $M = 1,52$; $F(1,27) = 7,65$, $p = .01$), et la série négative à une sous-évaluation globale des six stimulus positifs (négative : $M = 3,73$, étroite : $M = 4,30$; $F(1,27) = 5,38$, $p = .03$).

DISCUSSION

L'expérience 4 remet en question l'interprétation de l'effet d'équilibration de l'étendue initialement envisagée par Marsh et Parducci (*ibid.*). Lorsqu'il apparaissait conjointement à une norme stable, cet effet était supposé rendre compatible la position subjective de la norme sur l'étendue avec sa position centrale sur l'échelle d'évaluation. Le maintien d'un rééquilibrage de l'étendue, en l'absence d'un repère qui fonctionne comme une norme stable, constitue un résultat paradoxal, mais prévisible compte tenu des caractéristiques du matériel. La structure dichotomique du matériel permet un appariement des valeurs absolues, qui favorise une égalisation subjective des marges de l'étendue lorsque celles-ci ne font pas l'objet d'une présentation préalable. La mise en place d'une équilibration en l'absence de norme stable est en contradiction avec l'application du principe d'étendue (Parducci, 1983), selon lequel chaque catégorie de

réponse couvre une portion égale de l'étendue subjective. En effet, puisque l'étendue subjective évoquée par les séries dissymétriques est rééquilibrée, le zéro occupe une position médiane dans chacune de ces deux séries, et son évaluation ne devrait en conséquence pas donner lieu à un ajustement. Le maintien d'un effet d'étendue sur cette valeur conduit à une variabilité de l'intervalle couvert par les catégories affectées aux nombres négatifs et aux nombres positifs. Ce résultat peut être expliqué par le caractère partiel de l'effet d'équilibration, qui se limite à l'évocation d'une valeur extrême sans que soit reconstituée la portion de distribution qui sépare la marge effective de la marge évoquée.

DISCUSSION GÉNÉRALE

Si les normes de jugement sont conçues, dans la tradition psychophysique, comme le résultat d'une accumulation de faits en mémoire, notre recherche met l'accent sur la notion de dichotomie surimposée qui était jusqu'à présent négligée par le courant de recherche auquel nous nous référons. Dans la vie quotidienne, le fonctionnement des échelles de jugement est pourtant le plus souvent dichotomique : on accepte ou on rejette d'emblée des objets, des événements, en fonction de nos attentes, de nos exigences. Cette catégorisation initiale a l'avantage d'intervenir très rapidement face à un continuum numérique. Dans le cas de résultats scolaires, selon que la note est supérieure ou inférieure au critère nécessaire pour réussir, l'étiquette « réussite » ou « échec » peut lui être apposée. C'est dans un deuxième temps qu'on est sensible au contexte immédiat, et qu'une réussite pourra être jugée plus ou moins satisfaisante. Des exemples du fonctionnement dichotomique des échelles de jugement ne manquent pas. On peut même sérieusement se demander si tout jugement de valeur n'est pas fondamentalement dichotomique, même si l'apparence est contraire. La particularité des situations expérimentales est de proposer des tâches suffisamment simples pour qu'un seul critère soit applicable aux objets, ce qui n'est en général pas le cas lorsqu'il s'agit de personnes, de lectures, de spectacles ou d'événements – mais on peut parier que lorsque la dimension de l'espace de description grandit, la taille des échelles utiles se réduit, d'où l'intérêt d'un « point neutre » et la pro-

nable pertinence d'une extension multidimensionnelle de notre approche.

Lorsque, comme dans la recherche de Marsh et Parducci (*ibid.*) ou dans notre expérience 1A, la stabilité d'une norme s'observe conjointement à une équilibration de l'étendue, la conception psychophysique des normes de jugement n'est pas remise en cause. Dans ce cas, l'hypothèse d'un « ancrage naturel », développée par Marsh et Parducci, reste plausible : l'effet de stabilité peut s'expliquer par la nature dichotomique des contenus et par l'équilibration subjective de l'étendue qu'ils suscitent. Simplement, il faut considérer que si le neutre joue alors le rôle d'un niveau d'adaptation (Helson, 1947 ; Kahneman et Miller, 1986) c'est pour un contexte reconstitué autour de lui. Ce qui est perçu comme une limitation de l'effet de contexte. Mais la possibilité, mise en évidence par la recherche présente, de dissocier les effets et de les reproduire en l'absence d'habillage explicite offre une alternative pour dépasser cette interprétation initiale, et permet de dégager les sources de variation propres à chacun des effets. La dissociation montre que ces effets sont déterminés par des facteurs différents, qui tiennent pour l'un à une dichotomisation des évaluations, et pour l'autre à la structure de surface du matériel.

Les expériences 1B et 3 (échelle à repère dichotomisant) montrent que la stabilité des normes n'est pas automatiquement liée à un rééquilibrage du contexte, mais qu'elle est provoquée par une structuration dichotomique des jugements induite par l'habillage des données ou la consigne de réponse. Selon le type de tâche, le sujet se voit communiquer des normes conventionnelles qui devront guider son activité (expérience 3, échelle à repère dichotomisant), ou bien dispose de normes personnelles qui s'établissent en repères (expériences 1 et 2). Dans tous les cas, l'incitation à dichotomiser le matériel crée de la stabilité, et entraîne l'application d'une règle claire de mise en correspondance entre la norme et les stimulus. Comme le montre l'expérience 2, la disponibilité d'une catégorie médiane de réponse homogénéise le choix de la catégorie affectée au critère, et de ce fait renforce l'effet statistique de stabilité, mais ne conditionne pas le processus.

L'expérience 4 souligne l'originalité d'un matériel relatif, qui permet la reconnaissance du positif et du négatif sans relation nécessaire avec un critère de réponse stable, et facilite

l'évocation des extrêmes. L'utilisation d'une stratégie très économique, basée sur l'appariement des valeurs absolues des nombres, permet d'expliquer l'ensemble des résultats obtenus lorsque l'équilibration n'est pas gênée par une connaissance immédiate des marges de l'étendue. L'extension unilatérale de la marge des possibles est doublement conditionnée : par la structure du matériel et le mode de présentation des stimulus. En situation de non-connaissance des possibles, la présentation initiale d'un stimulus extrême dans un sens suggère l'éventualité d'une valeur opposée équivalente.

Par cette recherche, nous pensons avoir établi que les effets de stabilité et d'équilibration observés par Marsh et Parducci : 1 / n'excluent pas une influence du contexte immédiat, 2 / se retrouvent avec d'autres objets d'évaluation, et 3 / ne relèvent pas d'un seul et même mécanisme. Les résultats que nous avons obtenus démontrent que ces effets dépendent de l'existence d'une norme d'évaluation préétablie, de la structure des stimulus et de la forme de leur présentation. Associé implicitement au centre de l'échelle de réponse, le critère dichotomisant peut, par représentation, être replacé au centre du contexte immédiat en cas de dissymétrie de celui-ci. Toutefois, il apparaît que le recentrage du contexte est un effet qui nécessite une mise en forme particulière des stimulus, et que le critère peut aussi se maintenir en position excentrée par rapport au contexte. Cette dissymétrie des écarts par rapport à un critère dichotomisant se retrouve dans de nombreuses situations quotidiennes, comme celles par exemple de concours, ou d'examens très sélectifs où seuls les meilleurs candidats sont retenus, ou à l'inverse dans des situations peu sélectives. Elle correspond alors à une réalité fonctionnelle, et le recentrage du contexte n'a pas lieu de se réaliser. C'est après la décision que les écarts peuvent être relativisés, mais leur évaluation se fait alors, la plupart du temps, au prix d'une distorsion des échelles.

RÉSUMÉ

Les effets d'étendue de l'échelle de variation des stimulus, qui font partie des effets de contexte de référence sur les jugements perceptifs, peuvent subir une altération lorsque les sujets disposent d'une norme dichotomisante préétablie. Ces altérations, initialement mises en évidence par Marsh et Parducci (1978), consistent en un ancrage du centre de l'échelle de réponse aux stimulus

« neutres » qui jouent le rôle de critères de dichotomisation, et en une équilibration de l'étendue subjective de part et d'autre du « neutre ». Les quatre expériences présentées ici montrent que ces altérations 1 / ne suppriment pas totalement l'effet de contexte habituellement observé, 2 / s'observent avec différents types de matériels utilisables comme objets d'évaluation, et 3 / relèvent de mécanismes divers qui sont liés soit à la nature des systèmes d'évaluation soit à celle du matériel-stimulus. Cette diversité peut servir de guide pour l'analyse des effets analogues qui se produisent dans diverses situations de la vie quotidienne.

Mots-clés : normes de jugement, contexte, dichotomie, étendue subjective.

BIBLIOGRAPHIE

- Bresson F. — (1965) Les décisions, in P. Fraisse et J. Piaget (Édit.), *Traité de Psychologie Expérimentale*, vol. VIII, Paris, PUF, 221-306.
- Fabre J.-M. — (1993) *Contexte et jugement : de la psychophysique à la responsabilité*, Lille, Presses Universitaires de Lille.
- Harvey O., Campbell D. — (1963) Judgments of weight as affected by adaptation range, adaptation duration, magnitude of unlabelled anchor, and judgmental language, *Journal of Experimental Psychology*, 65, 12-21.
- Haubensak G. — (1992a) The consistency model : A process model for absolute judgments, *Journal of Experimental Psychology : Human Perception and Performance*, 18, 303-309.
- Haubensak G. — (1992b) The consistency model : A reply to Parducci, *Journal of Experimental Psychology : Human Perception and Performance*, 18, 314-315.
- Helson H. — (1947) Adaptation-level as frame of reference for prediction of psychophysical data, *American Journal of Psychology*, 60, 1-29.
- Johnson D. M. — (1949) Learning function for a change in the scale of judgment, *Journal of Experimental Psychology*, 39, 851-860.
- Kahneman D., Miller D. T. — (1986) Norm theory : Comparing reality to its alternatives, *Psychological Review*, 93, 136-153.
- Lépine D., Rouanet H. — (1975) Introduction aux méthodes fiduciaires : inférence sur un contraste entre moyennes, *Cahiers de Psychologie*, 18, 193-218.
- Manis M. — (1967) Context effects in communication, *Journal of Personality and Social Psychology*, 5, 326-334.
- Marsh H. W., Parducci A. — (1978) Natural anchoring at the neutral point of category rating scales, *Journal of Experimental Social Psychology*, 14, 193-204.
- Mellers B. A. — (1986) « Fair » allocations of salaries and taxes, *Journal of Experimental Psychology : Human Perception and Performance*, 12, 80-91.
- Parducci A. — (1965) Category judgment : A range-frequency model, *Psychological Review*, 72, 407-418.
- Parducci A. — (1974) Contextual effects : A range-frequency analysis, in E. C. Carterette et M. P. Friedman (Edit.), *Handbook of perception*, vol. II, New York, Academic Press, 127-141.
- Parducci A. — (1983) Category ratings and the relational character of judgment, in H.-G. Geissler, H. F. Buffart, E. L. Leeuwenberg et V. Sarris (Edit.), *Modern issues in perception*, Amsterdam, North-Holland, 262-282.

- Parducci A. — (1995) *Happiness, pleasure and judgment : The contextual theory and its applications*, Mahwah (NJ), Lawrence Erlbaum.
- Parducci A., Calfee R. C., Marshall L. M., Davidson L. P. — (1960) Context effects in judgment : Adaptation-level as a function of the mean, midpoint and median of the stimuli, *Journal of Experimental Psychology*, 60, 65-77.
- Parducci A., Fabre J.-M. — (1995) Contextual effects in judgment and choice, in J.-P. Caverni, M. Bar-Hillel, F. H. Barron et H. Jungermann (Edit.), *Contributions to decision making I*, Amsterdam, Elsevier Science Publishers, 97-109.
- Tversky A., Kahneman D. — (1971) The belief in the « law of small numbers », *Psychological Bulletin*, 76, 105-110.
- Tversky A., Kahneman D. — (1982) Judgments of and by representativeness, in D. Kahneman, P. Slovic et A. Tversky (Edit.), *Judgment under uncertainty : Heuristics and biases*, New York, Cambridge University Press, 84-98.
- Wedell D. H. — (1990) Methods for determining the locus of context effects in judgment, in J.-P. Caverni, J.-M. Fabre et M. Gonzalez (Edit.), *Cognitive Biases*, Amsterdam, North-Holland, 285-302.
- Wedell D. H. — (1995) Contrast effects in paired comparisons : Evidence for both stimulus-based and response-based processes, *Journal of Experimental Psychology : Human Perception and Performance*, 21, 1158-1173.
- Wedell D. H., Parducci A. — (1988) The category effect in social judgment : Experimental ratings of happiness, *Journal of Personality and Social Psychology*, 55, 341-356.

(Accepté le 10 septembre 1998.)