LA MESURE DES TR EN ARITHMETIQUE ELEMENTAIRE :

Spécificités d'une tâche de vérification.

J.P. FISCHER

La mesure des temps de vérification "d'égalités" numériques permet d'analyser quelques spécificités des tâches de vérification : effet de positivité, effet de distance symbolique, possibilité d'inverser l'opération ... Cette dernière présente un intérêt didactique particulier, en raison de l'insistance des Instructions Officielles de 1978 sur le calcul des soustractions par addition. Les élèves inversent-ils vraiment leur calcul des soustraction ?

1. Introduction

1.1. La possibilité matérielle - grâce à l'équipement informatique - d'introduire la mesure des Temps de Réponse (TR) dans les classes conduit à accorder une place plus importante, que celle que l'enseignement "traditionnel" leur accordait, aux **tâches de vérification** qui, seules, permettent une mesure commode et précise.

Une tâche de vérification consiste, en général, en une réponse dichotomique: *juste* ou *faux*, *oui* ou *non*, *même* ou *différent*, *appartient* ou *n'appartient pas*, *est* ou *n'est pas*, etc... Quand cette réponse à la question - ou plutôt au **stimulus** car la question, qui est toujours la même, n'est pas renouvelée à chaque mesure - est *juste*, *oui*, *même*, *appartient*, *est*,... le stimulus est qualifié de **positif.** Par contre, quand la réponse est *faux*, *non*, *différent*, *n'appartient pas*, *n'est pas*, ... le stimulus est qualifié de **négatif**. Par exemple, si l'on étudie les expressions algébriques (Ranney, 1987), 5L+(9)K *est* une expression algébrique donc un stimulus positif, alors que K9)L(5+ *n'est pas* une expression algébrique donc un stimulus négatif.

Qu'il me soit permis de remercier les élèves et responsables des classes impliquées dans l'expérience rapportée, ainsi que Philippe Bernat (IREM, Nancy), Guy Agostini et Gérard Cartron (EN Montigny-lès-Metz), Claire Dupuis et Raymond Duval (IREM, Strasbourg), qui ont tous pu m'aider dans l'une ou l'autre phase de ce travail.

© *Annales de Didactique et de Sciences Cognitives*
1 (1988) (p.153- 178) IREM de Strasbourg

Or, dans le **domaine** de l'**arithmétique** élémentaire, auquel je me limiterai désormais, une tâche de vérification peut poser des problèmes spécifiques qui peuvent la différencier fondamentalement d'une **tâche de production.** Cette dernière, comme le suggère son nom, consiste à produire un résultat juste ou, plus généralement, vérifiant une certaine propriété, par exemple trouver le résultat de 72:8. Trois spécificités de la tâche de vérification vont me permettre d'illustrer cette différence entre les deux types de tâches.

(1) **Pour les stimuli négatifs, il est souvent possible de "court-circuiter" le passage par le résultat.** Par exemple, pour conclure que 9x8=71 est fausse, il n'y a pas besoin de savoir que 9x8=72: il suffit de savoir qu'un produit par 8 ne peut être impair. D'ailleurs, les recherches de Krueger et Hallford (1984) et Krueger (1986), sur des adultes il est vrai, supportent l'idée que ces derniers font intervenir des règles de parité dans leurs vérifications. Mais ceci ne signifie pas pour autant que le traitement des stimuli négatifs est toujours plus facile. Dans le domaine des calculs élémentaires, nous verrons par la suite pourquoi. Mais déjà, dans le cas de raisonnements inductifs (Lefèvre et Bisanz, 1986), où il s'agit de trouver une régularité dans une suite de nombres, on peut remarquer qu'un stimulus négatif comme 5, 10, 16, 20 semble poser bien plus de problèmes qu'un correspondant positif, par exemple 5,10,15, 20, du fait de l'existence possible d'une régularité qui n'est pas celle qui est suggérée.

(2) **Pour les opérations inverses, il se pose un problème évident de stratégie de lecture:** si un élève, qui a mémorisé la table de multiplication mais pas celle de division, doit vérifier 72:8=9 il a tout intérêt à lire l'égalité de droite à gauche et à vérifier ainsi 9x8=72, alors que s'il doit produire le résultat de 72:8 le passage par la multiplication n'est pas autant facilité ou induit.

A ce sujet, il est intéressant de noter que la seule (à ma connaissance) tentative de comparaison empirique entre tâche de production et tâche de vérification (Ashcraft, Fierman et Bartolotta, 1984) a porté sur une seule opération et a abouti à la conclusion que les deux tâches ne différaient pas essentiellement. Cette opération étant une opération directe - l'addition - il convient de s'interroger sur la validité d'un tel résultat pour les opérations inverses.

(3) Si les stimuli négatifs semblent poser des problèmes particuliers, les stimuli **positifs** en posent également ! En effet, **le fait d'indiquer la réponse juste transforme le rappel** - l'élève doit rappeler le carré de 3 par exemple - **en reconnaissance** - l'élève doit reconnaître une égalité juste, par exemple 3x3=9. Et l'expérience courante - un visage dont

on ne se rappelle pas mais que l'on arriverait facilement à reconnaître par exemple - suggère que cette distinction peut avoir son importance. D'ailleurs, à des niveaux élémentaires, Mishkin (1982) suggère que la reconnaissance est une activation de la représentation cérébrale par **la** voie sensorielle qui l'a créée, alors que le rappel est une activation de cette représentation par **une autre** voie. Egalement, un certain nombre de résultats - mon exemple aussi - font penser que le rappel nécessite plus de ressources de traitement que la reconnaissance (Craik et McDowd, 1987).

1.2. Les trois points qui viennent d'être dégagés introduisent au contenu de cet exposé. Le choix d'une tâche de vérification, dans les expériences JusteFaux (Fischer, 1987a, 1987b; Gadio, 1987; Fischer et Pluvinage, soumis), et quelques autres choix subséquents, que je vais rapporter, conduisent en effet à se poser toutes ces questions, et permettent d'y apporter des éléments de réponse.
D'abord, la présence d'une moitié de stimuli positifs conduit à se demander si l'on retrouve un **effet** général de **positivité,** un effet (*the fast-"same" effect*) qui a fait l'objet d'une littérature abondante dans les expériences de type "même-différent" où il s'agit de dire si deux stimuli perceptifs sont identiques ou différents (pour une synthèse, voir Farell, 1985). Ce problème, on peut le penser, est lié à celui posé dans le point (3) précédent. En effet, la reconnaissance n'est vraiment possible que pour les égalités justes, car il est peu probable que les égalités fausses soient stockées dans notre mémoire.
Ensuite, le choix de la réponse fausse s'avère, d'après le point (1), essentiel. En effet, si le faux résultat a la même parité que le vrai, le rejet sur la base de la parité, suggéré ci-dessus, n'est plus possible. Egalement et intuitivement, une égalité comme 6x7=100 (resp. 6x7=40) paraît "grossièrement" fausse, alors qu'une égalité comme 6x7=43 (resp. 6x7=48) paraît plus "vraisemblable". Certains psychologues ont tenté de traduire ceci en introduisant l'expression **distance symbolique**: plus la réponse fausse proposée est (symboliquement) distante de la réponse juste, plus l'égalité serait alors facile à rejeter.
Enfin, le choix de présenter, dans les expériences JusteFaux, les quatre opérations arithmétiques élémentaires pose à l'évidence le problème de l'**inversion** formelle soulevé dans le point (2). Si ce problème n'a guère été soulevé à ce jour c'est, je pense, pour des raisons conjoncturelles: les psychologues, cherchant à construire des modèles, notamment de la recherche en mémoire, préfèrent souvent les cas les plus simples, à savoir l'addition et la multiplication. Mais ceci les conduit à des résultats - par exemple le modèle de traitement d'Ashcraft (1982), ou le réseau unique (en mémoire) pour les additions et multiplications

(Geary, Widaman et Little, 1986) - dont la généralisation aux quatre opérations est loin d'être évidente et même, au vu des résultats de Fischer et Pluvinage (soumis), peu probable (d'autant que se pose aussi le problème de la représentativité des sujets testés).

1.3. Si les trois points dégagés dans le paragraphe 1.1 introduisent bien le contenu, il n'en est pas de même pour le plan. Celui-ci est en effet un peu inattendu.
Je commencerai, dans le paragraphe **2**, par rappeler la méthodologie des "anciennes" expériences JusteFaux. Ceci me permettra alors de discuter tout de suite le problème de l'inversion formelle des divisions. En effet, dans ces "anciennes" expériences, j'avais posé des divisions à des élèves majoritairement en fin de CM2. Or, à ce niveau et à cette époque de la scolarité, cette étude me paraît plus pertinente que celle que je pourrais faire sur les résultats, obtenus au début du CM1 avec la nouvelle version de JusteFaux, rapportés par la suite. Ceci parce qu'en début de CM1 les élèves venaient à peine de voir la division et que l'on sait - je l'ai d'ailleurs vérifié empiriquement pour la soustraction dans ma thèse (Fischer, 1979) - que l'introduction scolaire d'une nouvelle notion fondamentale ne conduit pas immédiatement à des effets miraculeux.
Je passerai ensuite, dans le paragraphe **3**, à la description des nouvelles versions de JusteFaux. Cette description sera d'autant plus facile que les principaux choix méthodologiques déjà décrits sont restés les mêmes.
Dans le paragraphe **4**, je terminerai la discussion de l'inversion des opérations en m'intéressant cette fois à la soustraction. Le fait de consacrer un paragraphe entier à cette dernière me paraît d'ailleurs didactiquement opportun: en effet, les effets de la forte incitation à calculer les soustractions par addition, que l'on peut trouver dans les programmes français de 1978, commencent maintenant à pouvoir être évalués.
Enfin, dans le paragraphe **5** , j'étudierai les effets de positivité et de distance symbolique.

2. Méthode et inversion des égalités de division

2.1. *Les principaux choix méthodologiques*

J'ai déjà longuement insisté sur le fait que le choix d'une **tâche de vérification** est un choix fondamental. Au passage j'ai également souligné que les expériences de type JusteFaux portent sur l'étude des quatre opérations élémentaires: addition, soustraction, multiplication et division. Je ne reviens donc pas sur ces choix et me limite ici à présenter quelques autres choix subséquents mais néanmoins importants.

2.1.1. Un premier choix est celui de la **limitation du temps de réponse** : les élèves n'avaient que 5 secondes, au maximum, pour répondre. Cette limitation du temps de réponse est plus ou moins imposée par mon désir de faire participer à l'expérience tous les élèves des classes impliquées. Ceci m'a en effet obligé à prévoir le cas des élèves qui ne savent pas, et n'ont pas non plus le moyen de retrouver, si une égalité est juste ou fausse: il fallait interrompre - automatiquement car je ne pouvais pas toujours être à côté de l'élève en difficulté - la mesure du TR. Quant à l'ordre de grandeur - 5 secondes - du délai de réponse, il me paraît justifié par la "taille" des égalités proposées (pour les additions et les multiplications par exemple, les deux termes de la somme ou du produit sont toujours inférieurs à 10). De plus, cette limitation à 5 secondes a l'avantage statistique d'éviter une trop grande variance des TR.

2.1.2. Un deuxième choix, une fois fait celui de l'étude des quatre opérations, est celui du mode de présentation: Fallait-il "mélanger" des opérations de nature différente ? Ici j'ai choisi de ne pas trancher: la moitié des élèves (modalité <u>REG</u>roupement) s'est vue proposer 10 additions, puis 10 soustractions, puis 10 multiplications et enfin 10 divisions, en étant prévenue, avant chaque bloc de10, de la nature de l'opération qui allait suivre; l'autre moitié (modalité <u>N</u>on-<u>REG</u>roupement) s'est par contre vue proposer les opérations "mélangées" et ignorait la nature de l'opération qui allait suivre. Par ailleurs, la présentation était telle que la comparaison des modalités REG et NREG devenait la plus légitime possible.

2.1.3. Enfin, un troisième choix est celui du traitement statistique des résultats. Pour intégrer le facteur exactitude - important en milieu scolaire ! - à mes comparaisons des TR, j'utilise le test de la somme des rangs (Wilcoxon) d'une manière un peu particulière (suggérée par Leach, 1979): pour chaque égalité, je répartis les élèves suivant leur TRu (<u>T</u>emps de <u>R</u>é<u>u</u>ssite: TR correspondant à une réponse correcte) dans n classes <u>r</u>égulières de 500/n <u>c</u>enti<u>s</u>econdes (500 cs est le délai de réponse), et crée une classe extrême où sont regroupés les auteurs de réponses incorrectes ou de <u>N</u>on-<u>R</u>éponses (NR). Ceci conduit à une **comparaison intra-égalité**, non pas des TR, mais de ce que j'appelle **des performances**. Cette comparaison des performances constitue une réponse intéressante au délicat problème de l'échange vitesse-exactitude, comme cela a été montré dans Gadio (1987). En outre, je précise que le test, noté Z_{rn+1} par la suite, sera pratiqué unilatéralement car les meilleures performances dans la modalité REG me semblent prévisibles (voir le 2.2 suivant).

2.2. *L'approche du problème*

Les élèves dans la modalité REG, qui sont prévenus qu'ils vont faire des divisions, peuvent adapter leur stratégie, de lecture notamment. En conséquence, ils peuvent, sans perte de temps, transformer la vérification des items de divisions en opération de mutiplication. Comme ces dernières sont en général beaucoup plus travaillées et connues, on peut faire l'hypothèse qu'ils vont surclasser leurs camarades de la modalité NREG. Pour voir s'il en est bien ainsi, j'ai comparé les performances dans les deux modalités pour les 10 items de division, aussi bien dans l'expérience au CM2 (Fischer, 1987a) que dans l'étude génétique au CM1 et CM2 (Fischer, 1987b), les divisions n'ayant pas été présentées au CE2.

2.3. *Les résultats*

Le tableau de la page suivante résume les résultats.
Comme on peut le voir, l'hypothèse d'une supériorité des élèves de la modalité REG s'est amplement confirmée au CM2: dans l'une des expériences il y a en effet 8 égalités sur les 10 qui ont conduit à des différences significatives au seuil usuel de .05; dans l'autre, il y en a 9 sur les 10 qui sont significatives. De plus, sans aucune exception, les élèves dans la modalité REG ont toujours été plus rapides.
Au CM1, les résultats sont loin d'être aussi nets. Certes la seule égalité significative est encore en faveur de REG, mais quatre fois sur 10 les élèves de NREG ont été très légèrement plus rapides.
L'ensemble de ces résultats suggère donc l'existence d'une **tendance développementale** qui serait en faveur de l'interprétation d'une adaptation de la stratégie. En effet, l'adoption, en modalité REG, d'une stratégie qui creuse d'importantes différences par rapport à la modalité NREG, présupposerait la compréhension de la relation formelle liant division et multiplication. Or une telle compréhension doit être, à l'évidence, plus fréquente au CM2 qu'au CM1.

Remarques. Pour ce qui concerne la tendance développementale, il convient de faire deux réserves:

1) Le grand nombre de différences significatives au CM2, comparativement au CM1, n'est pas une preuve statistique de son existence.
2) Les élèves de la modalité REG au CM1 semblent avoir été plus faibles (en moyenne) que leurs camarades de NREG: la comparaison des modalités présentée dans mon autre contribution le suggère très nettement.

Comparaison des modalités pour les divisions

Division	CM1 (Fischer, 1987b)			CM2 (Fischer, 1987b)			CM2 (Fischer, 1987a)		
égalités	Zr501	Signif.	différence TRu	Zr501	Signif.	différence TRu	Zr1001	Signif.	différence TRu
6 : 3 = 2	.87	n.s		4.05	.001		1.67	.05	
16: 4 = 4	1.85	.05		5.18	.001		5.58	.001	
12: 3 = 5	1.26	n.s		1.83	.05		4.87	.001	
10: 2 = 4	- .42	n.s		3.29	.001		3.80	.001	
15: 5 = 3	.87	n.s		1.41	.1		4.85	.001	
14: 7 = 3	1.50	.1		3.42	.001		4.69	.001	
72: 8 = 9	1.31	.1		2.00	.05		1.64	.10	
81: 9 = 9	.53	n.s		4.47	.001		4.61	.001	
42: 6 = 9	- .92	n.s		2.73	.005		3.05	.005	
48: 6 = 7	.93	n.s		1.59	.1		3.93	.001	
Moyenne:									

Légende: = 30 cs pour REG   = 30 cs pour NREG (justifié à droite)

2.4. *Commentaire et vérifications complémentaires.*

En présentant l'hypothèse, j'ai en fait privilégié l'une des interprétations qui peut expliquer de meilleures performances dans la modalité REG. Mais les seuls résultats statistiques ci-dessus ne permettent pas d'affirmer que les élèves dans la modalité REG ont effectivement lu et calculé de droite à gauche si on n'a pas une idée de l'avantage que procure la modalité REG en général, i.e indépendamment de ce problème d'inversion de la lecture et du calcul.
Il faut donc préciser que pour la multiplication (l'autre structure multiplicative) où, à l'évidence, le problème de l'inversion ne se pose pas, j'ai enregistré les résultats suivants: au CM2b, il y a 5 égalités (sur 10) significatives en faveur de REG; au CM2a, 6; et au CM1, 1. Ces résultats suggèrent donc que l'inversion dans la modalité REG seulement n'est pas la seule source de la différence des modalités. Néanmoins ils sont parfaitement compatibles avec le fait qu'une partie de la population, de la modalité REG, a bien inversé la lecture et, en conséquence, le calcul.

2.5. *Observations cliniques*

Ces observations ont surtout été faites dans les préexpériences ou au cours de l'expérience de Fischer (1987a) où la passation était individuelle. Voici quelques observations directes, commentaires spontanés ou explications sollicitées d'élèves, tous de CM2 et tous dans la modalité REG:
Lae est une élève, scolairement performante mais qui se plaint de manquer d'entraînement (« *C'est dur quand même: l'année dernière on en faisait tous les jours, et cette année on n'en fait plus* »), qui oralise systématiquement le calcul inverse.
Nat commente spontanément après les divisions: «*J'ai multiplié les deux nombres à droite.*»

Jea, qui s'emmêle un peu dans les termes, explique : « *Je prends le résultat puis je prends le quotient - Je regarde si le diviseur est juste.... Je fais la preuve quoi !* »
Mic, un élève performant à JusteFaux, à qui j'ai demandé comment il faisait pour aller aussi vite et dont on peut regretter le silence (et la non-relance par l'expérimentateur) sur les soustractions:
« *Ben je faisais des petits trucs comme dit notre maîtresse. Par exemple pour les divisions je prenais la réponse qui était donnée et je la multipliais par le nombre qu'on devait diviser.*

Pour les additions j'avais pas trop de moyens: j'additionnais, je faisais l'opération en tête. Pour les multiplications je multipliais les deux nombres et puis je regardais. »
Ces quelques observations confirment donc qu'une partie au moins de la population, électivement dans la modalité REG, utilise bien le calcul inverse et, peut-être aussi, la lecture de droite à gauche.

2.6. *Conclusion.*

Mon intention était de montrer que les divisions, proposées dans une tâche de vérification, peuvent conduire à des stratégies de calcul, voire de lecture, essentiellement différentes de celle d'une tâche de production. Lorsque les divisions ne sont pas mélangées avec d'autres opérations, les résultats statistiques, aussi bien que les observations cliniques, vont effectivement dans ce sens, et montrent donc une limitation d'un modèle tel que celui d'Ashcraft (1982) s'il doit s'appliquer à toute l'arithmétique mentale.

3. Expérience avec une nouvelle version de JusteFaux

3.1. *Les changements (par rapport à l'ancienne version)*

La nouvelle version de JusteFaux étant destinée à l'évaluation scolaire de la connaissance des égalités numériques élémentaires, elle devait résoudre un problème d' "injustice" dans l'ancienne version: les élèves dans la modalité REG étaient globalement favorisés (dans le bilan du logiciel, il y avait une compensation théorique). Pour le résoudre, la nouvelle version prévoit deux passations pour chaque élève: l'une dans la modalité REG, l'autre dans NREG. Ceci permet en outre d'avoir deux jeux d'égalités et aussi une meilleure stabilisation statistique.
Sinon, les choix méthodologiques de l'ancienne version sont repris: tout au plus, on peut encore signaler que le nombre d'égalités par série est passé de 10 à 14, et que le délai de réponse, dans l'expérience qui va être décrite, a été de 532 cs.

3.2. *Le choix des égalités*

Je limite les faits numériques à des opérations portant sur des nombres entiers positifs inférieurs à 10 pour les opérations directes (additions et multiplications), et à des opérations dont le deuxième terme et le résultat sont inférieurs à 10 pour les opérations inverses

(soustractions et divisions).
Choix des opérations directes. Une fois ce choix fait, je ne prends en compte que l'un des deux couples de termes associés à une même paire. De plus, j'élimine les couples dans lesquels les éléments neutres 0 ou/et 1 sont impliqués et les doubles (couples de la forme (x,x)). Il reste alors 28 additions et 28 multiplications.
Dans ces 28 additions (resp. multiplications) je distingue 2 niveaux en les définissant essentiellement (et a priori) par des critères structuraux:
les additions et multiplications de niveau1 sont celles impliquant deux nombres inférieurs ou égaux à 5, ou au moins l'un des deux nombres, 2 ou 5.
Choix des réponses fausses. Le choix des réponses incorrectes aux égalités négatives se décompose ainsi:
- pour les 14 additions négatives, 7 réponses incorrectes s'écartent de 1 de la réponse correcte, et les 7 autres de 2;
- pour les 14 multiplications négatives, 7 réponses incorrectes s'écartent aussi de 1 de la réponse correcte, mais les 7 autres s'écartent d'un multiplicateur (ex.: 8x5=35).

En outre, et pour chacune des opérations, la moitié des réponses incorrectes est supérieure à la réponse correcte, l'autre moitié étant évidemment inférieure.
Choix des opérations inverses. A chaque opération directe choisie correspondent deux opérations inverses. Par exemple, à 2+3 (resp. 2x3) correspondent 5-2 et 5-3 (resp. 6:2 et 6:3). J'ai choisi arbitrairement l'une des deux (en réalité et primitivement j'ai opéré systématiquement: une fois sur deux je gardais le premier terme de l'opération directe. Mais des problèmes d'équilibrage m'ont ensuite obligé à abandonner ce choix initial dans le cas des structures multiplicatives). Enfin, le choix des réponses incorrectes pour les égalités négatives est analogue à celui des additions.

3.3. *Equilibrage d'après les tables de Wheeler.*

Pour toutes les études annoncées en introduction, le choix des égalités s'avère particulièrement important. Prenons l'exemple de l'effet de positivité. Pour étudier directement cet effet, la méthode la plus immédiate aurait pu consister à poser les mêmes calculs avec une fois une réponse correcte, une autre fois une réponse fausse. Mais comme je ne désirais pas reprendre plusieurs fois le même calcul je suis obligé d'effectuer la comparaison sur des calculs différents. Il est donc essentiel que, a priori, les calculs ayant généré les égalités positives soient de la même difficulté que ceux ayant généré les égalités négatives.
Pour arriver à un certain équilibre, j'ai utilisé les deux tables de Wheeler (1939 et 1941) qui donnent la hiérarchie de difficulté des 100 additions de la table, et des 99 multiplications

(Wheeler n'a pas inclus 0x0).
Avec ces tables, et en faisant la somme des rangs, j'ai équilibré les additions (resp. multiplications). Pour les soustractions (resp. divisions) j'ai considéré qu'une opération a-b=c (resp. a:b=c) obtiendrait le même classement que sa correspondante b+c=a (resp. bxc=a) ou c+b=a (resp. cxb=a) qui figure dans l'échantillon. Le tableau suivant - représentant le rang moyen obtenu par les égalités choisies (avec les conventions faites pour les opérations inverses) -résume les équilibres auxquels je suis parvenu:

	ADD		SOU		MUL		DIV	
	pos	nég	pos	nég	pos	nég	pos	nég
Ecart = 1	-	68	-	68	-	68	-	71
Ecart > 1	-	71	-	71	-	71	-	68
Ensemble	70	70	70	70	69	70	69	70
Egalités 1	69	68	71	71	70	69	69	70
Egalités 2	71	71	69	68	69	70	70	69

De plus, je distingue les **soustractions additives** et les **soustractions soustractives**: les soustractions additives (resp. soustractives) sont celles pour lesquelles le deuxième terme de la différence est supérieur (resp. inférieur) à la moitié du premier. Avec ces conventions de vocabulaire, les soustractions additives ont un rang moyen de 69.5, et les soustractives de 70.0. Il y a bien entendu aussi la même proportion - une sur deux - d'égalités positives dans les soustractions additives que dans les soustractions soustractives.

3.4. *L'ordre de présentation* .

Dans la modalité NREG, les égalités 1 se présentent dans l'ordre suivant:

- Série1: 9+5=14, 5x3=14, 72:9=6, 6+8=16, 13-5=8, 18:9=2, 3+4=7, 10-6=3, 24:3=7, 3+7=10, 4x6=24, 12:2=7, 10-2=8, 7x2=16.
- Série2: 13-9=4, 6+2=9, 9x3=27, 28:7=6, 11-5=7, 8+3=11, 6x5=30, 20:5=4, 9x4=36, 7+5=10, 54:9=7, 8+9=17, 11-4=8, 8x2=16.
- Série3: 35:5=7, 8x7=56, 2+3=7, 63:7=9, 8x5=35, 6:2=3, 3+9=11, 8-5=1, 4+2=6, 13-7=6, 8x6=47, 9-4=5, 32:4=8, 9-3=4.

- Série4: 4x2=8, 12:3=3, 6x7=49, 9+7=16, 5x9=46, 21:7=3, 11-9=4, 7+2=8, 15-6=8, 7+8=13, 10:5=4, 7-5=2, 3x6=17, 12-8=4.

Maintenant le jeu des égalités 2:

- Série1: 3+5=6, 4x5=20, 56:7=8, 8+4=12, 8-6=3, 30:6=5, 5+4=9, 17-9=8, 24:4=6, 4+7=10, 8x3=27, 16:2=8, 5-2=1, 5x2=9.
- Série2: 11-3=8, 5+6=12, 7x3=21, 27:9=3, 12-7=7, 4+9=13, 6x2=10, 15:3=7, 8x4=32, 2+8=10, 36:9=4, 7+6=13, 16-7=9, 3x4=16.
- Série3: 14:7=3, 6x9=45, 2+5=7, 18:3=4, 3x2=6, 45:9=3, 6+4=11, 14-5=9, 9+2=13, 12-9=4, 7x4=29, 7-3=4, 42:6=8, 10-7=3.
- Série4: 5x7=35, 40:5=7, 9x8=73, 6+3=7, 9x2=18, 48:8=4, 6-4=2, 5+8=13, 14-6=6, 9+6=16, 8:4=2, 9-2=6, 7x9=63, 15-8=9.

L'ordre de présentation dans la modalité REG peut être déduit des précédents: il suffit d'extraire, au fur et à mesure qu'on les rencontre, dans un jeu, d'abord les 14 additions, puis les 14 soustractions, puis les 14 multiplications, enfin les 14 divisions.

3.5. *L'expérience*

La population. L'expérience concerne quatre classes de CM de la proche banlieue ou campagne de Metz, dont l'évaluation est envisagée sur deux ans, à raison d'une fois tous les six mois environ. La présente étude porte sur la première évaluation, donc au CM1. L'effectif total a été de 82 élèves, mais un a du être éliminé pour réponses trop rapides (plutôt un problème d'alcoolisme parental que de contrat) et la dernière élève à être passée par ordre alphabétique a également été éliminée (de mon étude) pour avoir un plan expérimental "parfait". Tous les résultats portent donc sur 80 élèves.

La passation. Elle s'est déroulée en décembre-janvier et se décompose en deux: les passations 1 et 2 sont espacées d'une semaine environ. Les élèves passaient, en général, par groupes de 6 dans une salle informatique équipée soit d'un nanoréseau, soit d'une configuration de TO7 ou TO7/70 isolés, sous la surveillance de l'expérimentateur ou d'un collaborateur. Avant chacune des passations, l'expérimentateur donne quelques consignes collectives à l'ensemble d'une classe. Avant la passation 1, il informe succinctement les élèves de l'ensemble du projet (évaluation tous les six mois), du fonctionnement du programme, des informations que l'on peut tirer de ces mesures. Avant chacune des passations il insiste sur le fait qu'il s'agit de répondre vite mais correctement avant tout.

Les élèves "tombent", au hasard, sur une des quatre conditions expérimentales (équitablement distribuées) lors de la première passation: les égalités 1 dans la modalité

REG, ou les 1 dans NREG, ou les 2 dans REG, ou les 2 dans NREG. La condition expérimentale de la passation 2 est alors entièrement déterminée, pour chaque élève, par celle qu'il a eue au cours de la passation 1: s'il a eu les égalités 1 (resp. 2) il aura les égalités 2 (resp. 1) et s'il a travaillé dans la modalité REG (resp. NREG), il travaillera dans NREG (resp. REG).

Les contrôles statistiques. Par rapport à la version "ancienne" de JusteFaux, la nouvelle offre des possibilités supplémentaires de contrôle statistique. En effet, chaque élève ayant passé dans les deux modalités, on peut procéder à une comparaison des TRu moyens respectifs. Comme le suggère un exemple de Bradley (1968 p.101-102) la différence entre les deux, calculée pour chaque élève, donne un ensemble de nombres positifs ou négatifs, qui peut être soumis au test des rangs signés de Wilcoxon (procédure de Leach, 1979). Je parlerai d'une **comparaison intra-élève** des modalités, la comparaison intra-égalité restant bien entendu possible aussi.

De manière analogue, on peut utiliser ce test pour la comparaison des égalités justes et fausses, pour la comparaison des soustractions additives et soustractives, ou encore pour la comparaison des différences de modalités pour deux opérations arithmétiques différentes.

Remarque.

Comme c'est dans les échecs que la proportion de réponses "au hasard", d'accidents, de réponses trop hâtives,... doit être la plus grande, on voit immédiatement l'avantage principal de la considération des TRu au lieu des TR. Secondairement cela évite aussi d'assigner un temps arbitraire aux Non-Réponses.

Cette considération des TRu a cependant l'inconvénient de réduire, considérablement pour certains élèves, le nombre de mesures sur lequel porte le calcul de la moyenne. Elle pourrait aussi conduire à une déformation du pattern des vitesses de réponse: par exemple si, dans la comparaison des égalités positives et négatives, les échecs ont été très rapides pour les égalités positives, et très lents (ou des NR) pour les égalités négatives, on voit que le pattern des vitesses de réponse serait distordu en défaveur des égalités positives.

L'analyse des temps d'échec, et des nombres de NR, montre que la description ci-dessus est bien purement théorique.

4. Etude des soustractions

4.1. *L'inversion des égalités de soustraction*

4.1.1. *Le problème et la manière de l'attaquer.* Comme cela a déjà été souligné, les programmes français 1978 pour le Cours Elémentaire, qui sont ceux auxquels les sujets de la présente expérience ont en principe été soumis, préconisent le calcul des soustractions par

addition.
De manière précise, les Instructions Complémentaires mentionnent:

> En ce qui concerne la soustraction, le fait de calculer une différence par addition, constitue une technique qui a l'avantage de ne pas exiger des enfants d'autres compétences que celles requises pour l'addition. Si la disposition initiale (directement dérivée de a + . = b) est $\begin{array}{r} a \\ + \, . \\ \hline b \end{array}$ il conviendra d'aboutir, au moins en fin de cycle élémentaire, à la disposition habituelle $\begin{array}{r} b \\ - \, a \, . \\ \hline \end{array}$ sans que pour autant la technique de calcul soit modifiée.

Remarquons d'emblée que l'argumentation, qui s'appuie exclusivement sur un argument de "logique économique", paraît **antipsychologique**: peu de gens doivent calculer, par exemple 35-1 en passant par 1 + . = 35. Plus généralement, elle paraît refléter une certaine conception de l'apprentissage - les élèves apprennent les additions au CP, puis les utilisent au CE - qui est peu compatible avec des observations comme celles que j'ai faites sur le développement de la connaissance des faits numériques (Fischer, 1987b).
Passons maintenant à la manière d'approcher le problème. Si les élèves "adhèrent" à, i.e utilisent et comprennent, la manière de calculer suggérée par les Instructions, on peut penser qu'ils vont profiter de la présence d'une réponse pour effectuer une vérification: pour 11-5=7 par exemple, ils vont calculer tout de suite 7+5 et ne pas se demander 5+. =11. Dans ce cas nous devrions donc observer un phénomène analogue à celui qui a été observé pour les divisions: les performances dans la modalité REG devraient être bien meilleures que celles dans la modalité NREG.

4.1.2. *Résultats*. La supériorité de la modalité REG se confirme sans ambiguïté: comparativement à la modalité NREG, les TRu y sont inférieurs en moyenne de 32 cs (différence fortement significative: z=4,40); le pourcentage d'échecs y est inférieur de 3% (non significative: z=1,51). Mais le même problème que pour les divisions se repose ici: peut-on interpréter cette supériorité de REG comme une "preuve" d'inversion ? Pour répondre, on peut comparer les soustractions cette fois aux additions pour lesquelles le problème de l'inversion ne se pose pas. Or, contrairement à ce qui s'est passé pour les divisions comparativement aux multiplications, les soustractions ne semblent pas conduire à une différence des modalités plus importante que les additions. Par exemple, la comparaison intra-égalités, avec le test Zr51, conduit à **11** différences significatives sur 28 possibles, en faveur de REG, pour les soustractions, alors qu'il y en a **12** pour les additions. Il y a également une augmentation du pourcentage de réussite de 4% (significative: z=2,39), supérieure aux 3% ci-dessus. Par contre, la différence des TRu pour les additions n'est que de 26 cs: même si elle est tout aussi fortement significative (z=4,85), elle est inférieure aux 32 cs observés pour les soustractions. En dernier, une comparaison directe entre les différences des TRu dans les deux modalités pour les soustractions et les additions ne conduit **pas** à une **différence significative** (z=1,31): comme elle va néanmoins dans la direction d'une plus grande différence dans les soustractions, il est utile de rajouter (car il ne pourrait s'agir que d'un problème de pouvoir de rejet) que la comparaison analogue entre soustraction et multiplication est bien significative (fortement: z=2,80).

4.1.3. *Observations incidentes.* Contrairement à ce que j'ai rapporté sur les divisions, je n'ai pas observé d'élève inversant systématiquement (à haute voix) les soustractions, ni même faisant une remarque laissant sous-entendre une telle inversion systématique. Certes, la raison est en partie "technique": j'ai beaucoup moins observé les élèves individuellement au cours de cette recherche qu'au cours de Fischer (1987a). Néanmoins, cette raison "technique" ne m'a pas empêché de noter encore une jolie remarque sur l'inversion des ... divisions: « *Pour les divisions j'ai un truc* » (Pie) ! Mais surtout, en "rattrapant" Séb, absent lors de la passation "officielle", il m'a fait spontanément une confidence: « *Pour les divisions, je fais le contraire* ». Et cette fois-ci, contrairement à ce qui s'était passé pour Mic (voir le 2.5), je n'ai pas manqué l'occasion de rajouter: Et pour les soustractions aussi ? - « *Euh.... non.* » - Pourquoi pas ? - « *Parce que les soustractions je sais assez bien.* »
Le traitement différent des soustractions et des divisions semble ici très bien illustré. De plus, la surprise de Séb semble montrer qu'il ne s'était pas vraiment posé la question de l'inversion des soustractions, et son explication n'est guère convaincante lorsqu'on connaît la difficulté des soustractions (voir Fischer, 1987a), tout au moins pour ce qui concerne leur exécution (voir Fischer et Pluvinage, soumis). D'ailleurs les performances de Séb, au cours de la passation qui a suivi cette observation, sont conformes à ce pattern général: il a fait 5 fautes aux 14 soustractions, contre une seulement aux 14 divisions.

4.1.4. *Conclusions.* Pour les soustractions comme pour les divisions les observations cliniques ou incidentes rejoignent les observations statistiques. Mais cette fois-ci la conclusion est autre: l'ensemble des observations est en accord avec l'hypothèse que **très peu d'élèves inversent systématiquement leur vérification des égalités de sous-**

traction.

4.2. *Trois modèles de calcul des soustractions*

4.2.1. Si peu d'élèves calculent donc à partir du résultat, on peut se demander de quelle manière s'y prennent ceux qui calculent effectivement, car quelques-uns "démissionnent" dans les soustractions difficiles et quelques autres connaissent peut-être les soustractions faciles par coeur. Parmi les modèles les plus plausibles, trois me semblent intéressants à considérer ici. Ceci parce qu'ils peuvent conduire, grâce à la présence et à l'équilibrage des soustractions additives et soustractives, à des prédictions vérifiables.
Ces trois modèles sont les suivants:

- l'Addition mécanique (Am): elle se distingue de l'Addition avec adhésion (Aa) précédente par le fait que l'élève calcule mécaniquement, par exemple 11-3, par addition, i.e en cherchant 3+.=11, sans tenir compte, dans son calcul, de la réponse proposée;
- la Soustraction mécanique (Sm): les élèves utilisent une procédure de comptage, non nécessairement un par un, en arrière;
- le Choix: l'élève choisit, en fonction du rapport des deux termes de la différence, un calcul par Addition mécanique s'il est inférieur à deux, et un calcul par Soustraction mécanique s'il est supérieur.

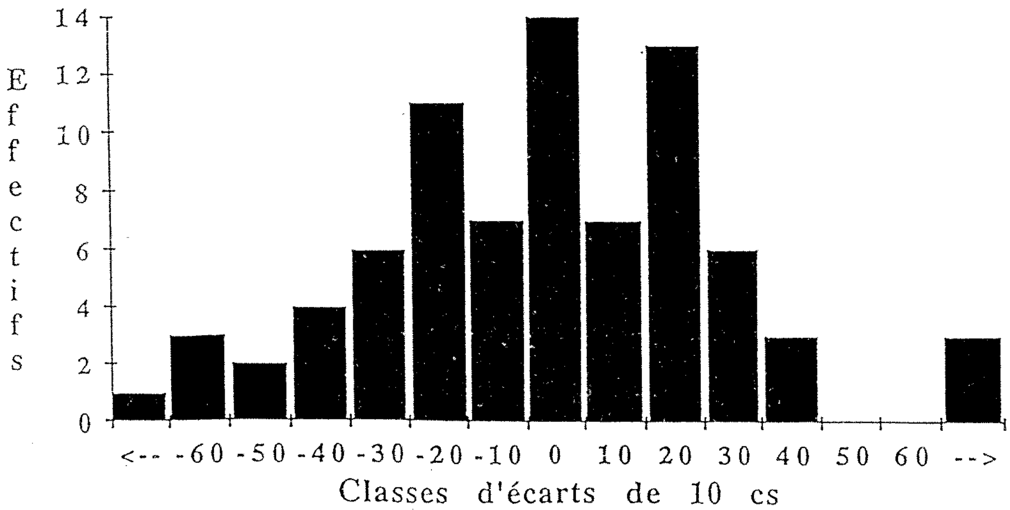
Les prédictions que l'on peut faire, si l'un de ces modèles s'avérait largement dominant, sont les suivantes:
- si Am est dominant, les soustractions additives devraient être plus faciles
- si Sm est dominant, les soustractions soustractives devraient être plus faciles
- si Ch est dominant, on devrait avoir un équilibre. Tout au moins si l'on admet que, par exemple, le calcul de 11-8 par addition est équivalent à celui de 11-3 par soustraction; ou au moins qu'il y a une équivalence globale, c'est-à-dire qu'un déséquilibre des calculs cités est compensé par un ou plusieurs autres déséquilibres.

4.2.2. La comparaison des résultats conduit à une **égalité quasi-parfaite entre soustractions additives et soustractives**. Qu'on en juge:
- les TRu sont respectivement de 275 et 276 cs;
- les pourcentages d'échecs de 23,4 et 23,8%.

Bien entendu, aucune de ces différences n'est significative, et même la faible tendance - car les soustractions additives sont à la fois très légèrement plus rapides et mieux réussies - n'est pas vraiment pertinente car le choix des réponses fausses a pu légèrement favoriser les

soustractions additives.
La conclusion est donc qu'**aucun des deux modèles Am ou Sm n'est largement dominant** dans la population étudiée. Mais ceci ne signifie pas pour autant que c'est le modèle Ch qui est dominant. En effet, si ce modèle est bien compatible avec l'équilibre des deux types de soustraction observé, il n'est pas la seule manière d'arriver à un tel équilibre. Par exemple, et théoriquement, on peut imaginer qu'une moitié des élèves suit Am et l'autre Sm: dans un tel cas, on aurait aussi l'équilibre. Ou encore, on peut imaginer qu'un tiers des élèves suit Am, un deuxième tiers Sm et le dernier tiers Ch.

4.2.3. Pour essayer de voir si l'une ou l'autre de ces possibilités théoriques s'est effectivement réalisée, j'ai analysé la distribution des écarts individuels des TRu aux deux types de soustractions. En effet, si on se place par exemple dans le cas du partage en trois tiers, on devrait avoir une ditribution trimodale des écarts individuels: en calculant l'écart dans le sens TRu aux soustraction additives - TRu aux soustractions soustractives, le tiers d'élèves suivant Am devrait se regrouper autour d'un écart négatif, le tiers d'élèves suivant Sm autour d'un écart positif, et tiers suivant Ch autour de 0. La distribution obtenue est reproduite ci-après.



A première vue cette distribution a effectivement une allure un peu trimodale. Néanmoins, pour que l'on puisse affirmer qu'elle est trimodale, il faudrait, au minimum, arriver à rejeter l'hypothèse de normalité à un seuil peu sévère comme .10.

Un χ^2 d'ajustement (Snedecor et Cochran, 1971) appliqué aux données regroupées en classes, larges de 10 cs et centrées sur les multiples de 10 (comme sur la figure, où j'ai regroupé les classes extrêmes), n'a pas permis ce rejet. Comme le regroupement en classes est arbitraire et conduit à une perte d'information, j'ai aussi testé la normalité avec le test de Kolmogorov-Smirnov (Siegel, 1956) qui peut s'appliquer aux données originales: ce test n'a pas conduit non plus au rejet de l'hypothèse de normalité. Enfin, en dernier recours, j'ai testé la symétrie et l'aplatissement: la première semble normale, mais pas le second. Néanmoins cet aplatissement non normal semble essentiellement dû à quelques cas extrêmes douteux.
Tout cela m'incite présentement à ne pas m'aventurer au-delà des conclusions soulignées dans le 4.2.2, des conclusions que le non rejet de la symétrie renforce d'ailleurs.

4.3. *Conclusion*

Même si la technique précise du calcul n'a pu être élucidée, il n'en demeure pas moins que les résultats du 4.1., en ce qui concerne le modèle Aa, et les résultats du 4.2, en ce qui concerne Am, sont compatibles avec le fait qu'**il n'y pas un phénomène massif de calcul des soustractions par addition.**

5. Effets de positivité et de distance symbolique

5.1. *Les hypothèses*

5.1.1. Plusieurs recherches sur les calculs élémentaires ont mis en évidence l'effet de positivité, tout au moins lorsque les réponses fausses proposées ne sont pas trop grossièrement erronées: Findlay (1978), Ashcraft et Fierman (1982), Hamann et Ashcraft (1985), Ashcraft et Battaglia (1978), Ashcraft et Stazyk (1981), Gonzalez et Kolers (1982), Krueger et Hallford (1984), Krueger (1986), Stazyk, Ashcraft et Hamann (1982). Hormis les trois premières, ces recherches portent cependant sur des sujets adultes (étudiants). De plus, les quatre opérations arithmétiques sont loin d'être uniformément représentées: seules les deux dernières recherches portent sur une structure multiplicative. Les opérations inverses n'apparaissent qu'une fois: il s'agit de la recherche de Findlay (1978), qui a "mélangé" additions et soustractions et qui porte d'ailleurs sur une population statistiquement peu nombreuse.

Pour donner une idée de l'ampleur de l'effet de positivité, je citerai la différence temporelle trouvée par Ashcraft et Fierman (1982), pour des additions de nombres à un chiffre et sur des élèves de $3^{ème}$, $4^{ème}$ et $6^{ème}$ année d'école: environ 17 cs en moyenne. Pour la différence des pourcentages d'erreur, les résultats d'Hamann et Ashcraft (1985) vont, en $4^{ème}$ année d'école du moins, en sens contraire. Mais, d'une part, la différence n'est pas très importante, d'autre part, sur les trois "tailles" d'erreurs étudiées - 1 ou 2, 6 ou 7, et 12 ou 13 - l'avant dernière, et surtout la dernière, sont déjà assez grossières pour de petites additions. Or comme dans la recherche de Findlay il apparaît que les erreurs de 1 conduisent à 2,5% d'erreurs de plus que les erreurs plus grossières, on peut donc penser que, chez Hamann et Ashcraft, ce sont surtout les deux dernières tailles d'erreurs qui sont à l'origine de ce résultat.

Pour justement essayer de préciser ce que l'on peut entendre par égalités *grossièrement* fausses, c'est la recherche d'Ashcraft et Stazyk (1981) qui me paraît la plus intéressante à citer, même si elle porte sur des adultes. Ashcraft et Stazyk avaient en effet proposé des additions (sur des nombres à un chiffre) fausses de 1, de 5, de 9 et de 13: l'effet de positivité a diminué avec l'écart mais ne s'est inversé que pour 13. Et cette inversion, très nette pour les "grandes" additions, où elle est de l'ordre de 20 cs, est théoriquement importante. Elle souligne d'une part qu'il vaut mieux parler d'un effet de négativité, d'autre part une limitation de certaines explications générales, par exemple celle consistant à attribuer l'effet de positivité à un « *caractère général des représentations élémentaires* » qui serait « *d'insister sur les aspects positifs des objets et événements et de négliger leurs aspects négatifs (x «n'est pas» a, b, etc.)* » (formulations de Piaget, 1977).

Ces précisions données, je peux maintenant en revenir à ma propre expérience. Aucune des réponses fausses ne semblant l'être "grossièrement", je ferai l'**hypothèse 1** que, **pour chacune des quatre opérations, les performances sont meilleures pour les égalités positives** (comparativement aux négatives correspondantes) de mon échantillon.

5.1.2. Pour l'effet de distance symbolique, comme on a déjà pu l'entrevoir à propos des erreurs "grossières", les choses se compliquent un peu. En effet, les quelques recherches ayant étudié cet effet ont souvent choisi des distances bien "espacées" comme 1, 5, 9, 13, et ont parfois eu du mal à le retrouver, notamment pour la multiplication (Stazyk et al., 1982). Se pose aussi le problème des réponses violant la parité (Krueger et Hallford, 1984; Krueger, 1986).

Pour donner une idée de l'ampleur de l'effet, je citerai encore la recherche d'Ashcraft et

Fierman (1982): ils ont trouvé que des additions fausses de 1 étaient plus lentes d'environ 30 cs que des additions fausses de 5 ou 7. Pour les multiplications, Stazyk et al., avec des adultes, ont bien introduit (expérience 3) des réponses fausses de la table comme je l'ai fait. Mais leur technique n'a pas été identique à la mienne: ils présentaient le calcul 20 ou 60 cs avant la réponse. Je préfère donc me référer à un "papier" de Duffy et Fisher (rapporté dans Krueger et Hallford, 1984 p. 179): ces chercheurs ont trouvé que les produits faux sont rejetés plus rapidement (115 millisecondes) si la réponse n'est pas un multiple de l'un des deux termes.

En dépit donc de certaines "complications" et du manque d'évidences (nettes et fines) de l'effet de distance symbolique pour certaines opérations ou pour certaines populations, mon **hypothèse 2** sera que **les élèves sont plus performants dans les additions, soustractions et divisions fausses de 2 que dans celles fausses de 1**, mais que, par contre, **ils sont plus performants dans les multiplications fausses de 1 que dans celles fausses d'un multiplicateur.**

5.2. *Les résultats*

Le tableau suivant rassemble tous les résultats.

		effet de positivité			effet de distance symbol.		
Opér.	Crit	posit.	négat.	différ.	nég. 1	nég. >1	différ.
ADD	Ech.	15	13	- n.s.	15	11	+ p<.05
	TRu	234	253	+ p<.0002	255	251	+ n.s.
SOU	Ech.	24	23	- n.s.	29	17	+ p<.0002
	TRu	268	283	+ p<.001	290	276	+ p<.002
MUL	Ech.	09	14	+ p<.01	09	18	+ p<.0002
	TRu	226	260	+ p<.0002	262	257	- n.s.
DIV	Ech.	18	19	+ n.s.	24	14	+ p<.0002
	TRu	254	278	+ p<.0002	280	276	+ n.s.

Pour la lecture de ce tableau, il faut donner les précisions suivantes:

- les Echecs sont indiqués en pourcentages et portent respectivement sur 1120 (=14x80) mesures par case pour l'effet de positivité, 560 (=7x80) mesures par case pour l'effet de distance symbolique;

- les Temps de Réussite sont en centièmes de seconde;
- le signe + (resp. -) montre que la comparaison va (resp. ne va pas) dans le sens de l'hypothèse;
- p<s précise le seuil s de signification (test bilatéral), et n.s. signifie non significatif, les seuils extrêmes pris en considération étant s=.20 et s=.0002.

5.3. *Les commentaires*

5.3.1. Pour l'**effet de positivité**, on voit que dans la colonne des différences le signe - apparaît deux fois. Mais la tendance contraire à l'hypothèse 1 qu'il traduit n'est pas significative. Comme par ailleurs tous les résultats non significatifs (qui concernent exclusivement les échecs) sont accompagnés d'un résultat fortement significatif pour les TRu, on peut conclure que, globalement, l'hypothèse 1 est confirmée: **les performances sont meilleures pour les égalités positives.**
Néanmoins, l'hypothèse 1 mérite incontestablement d'être affinée dans le sens que laissait prévoir l'examen de la littérature: l'effet facilitant de la positivité est clair et net pour les TRu, en particulier pour la multiplication, mais, hormis pour ces dernières, il ne se voit quasiment pas pour les échecs.

5.3.2. Pour l'**effet de distance symbolique**, le commentaire prend à peu près la même forme. La seule fois où il apparaît un signe - dans la colonne des différences, la tendance contraire à l'hypothèse 2 qu'il traduit n'est pas significative. Comme tous les résultats non significatifs (exclusivement au niveau des TRu) s'accompagnent toujours d'un résultat significatif pour les échecs, on peut aussi conclure que l'hypothèse 2 s'est, globalement, confirmée: **les performances sont moins bonnes pour les égalités fausses de 1, sauf pour les multiplications.**
Néanmoins, il convient de préciser que c'est surtout au niveau des échecs que le résultat est clair et net, alors qu'au niveau des TRu il n'existe que pour les soutractions.

5.3.3. Mais le commentaire serait incomplet si l'on ne discutait pas simultanément les deux effets. D'autant que les résultats obtenus sont, pour les opérations inverses du moins, particulièrement jolis: les égalités positives viennent chaque fois, au niveau des échecs, s'intercaler entre les égalités fausses de 1 (moins souvent réussies que les positives) et les égalités fausses de 2 (plus souvent réussies). Ceci renforce un peu l'hypothèse 1 dans la mesure où, même pour les échecs, il semble y avoir un petit effet facilitant de positivité si l'on choisit des réponses fausses symboliquement très proches de la réponse juste. Et, dans cette dernière perspective, 2 s'avère déjà une trop grande distance, alors qu'un multiplicateur s'avère, dans le seul cas des multiplications évidemment, un choix qui semble idéal.

6. Conclusions

6.1. Dans la présente contribution, j'ai essayé de montrer qu'une tâche de vérification, en arithmétique élémentaire, présentait quelques spécificités par rapport à une tâche de production.

Les spécificités sur lesquelles j'ai insisté sont: la possibilité de vérifier les égalités impliquant les opérations inverses (soustraction et division) en partant du résultat proposé, l'effet de positivité selon lequel les égalités positives (justes) conduisent à de meilleures performances, et l'effet de distance symbolique selon lequel une égalité négative (fausse) est d'autant plus difficile à rejeter que la réponse fausse proposée est symboliquement proche de la réponse juste.

En général, avec éventuellement quelques affinements, mes différentes observations **confirment** assez nettement ces spécificités, à l'exception toutefois de la vérification de la soustraction.

Mais cette confirmation ne doit pas faire croire que la mesure des TR dans une tâche de vérification est un exercice vraiment particulier dans lequel, par exemple, le réflexe jouerait un rôle prépondérant: d'une part, parce que les corrélations que j'ai pu faire entre un classement à JusteFaux et le classement scolaire habituel sont toujours significativement positives, d'autre part parce que Geary et Widaman (1987) ont montré que les composantes du traitement détectées avec cette méthode sont en continuité avec les habiletés intellectuelles identifiées par les méthodes d'analyse factorielle.

6.2. L'exception des soustractions m'a conduit à analyser le calcul des soustractions d'un peu plus près: les résultats trouvés, les observations recueillies, et les interprétations proposées, suggèrent que **les élèves de CM1 n'utilisent pas** systématiquement et massivement **la technique de calcul par addition** fortement recommandée par les programmes officiels auxquels ils ont été soumis.

J'ai suggéré ailleurs (Fischer, 1987a) une explication possible de ce fait: bon nombre d'élèves connaissent insuffisamment les additions pour que le passage par ces dernières présente un réel avantage pour eux. Mais d'autres explications complémentaires peuvent être avancées:

- le calcul "normal" et conscient est, sous la pression du temps, "court-circuité" par une reconnaissance (voir mon autre contribution) qui est parfois erronée. En témoigne cette réponse d'un sujet assez brillant (un an d'avance scolaire) qui a, néanmoins et comme beaucoup de ses camarades, échoué à 11-3=9. Lorsque je lui ai demandé, a posteriori, comment il a fait pour se tromper, il a expliqué: «**Normalement** je fais neuf plus trois, douze .... »;

- les élèves ont des procédures soustractives déjà trop bien "ancrées" pour que l'apprentissage scolaire puisse les modifier: le cas de Séb, rapporté dans le 4.1.3, illustre très bien une telle explication. En effet, Séb explique, en dépit des nombreuses fautes qu'il fait, qu'il n'inverse pas les égalités de soustraction parce qu'il les *sait assez bien*. Une autre observation peut la corroborer: les nombreuses erreurs pour les soustractions fausses de 1 - 29%, i.e nettement le pourcentage d'erreurs le plus fort dans le tableau précédent - est en accord avec l'utilisation de procédures soustractives, apprises très précocément (Fischer, 1981), lesquelles conduisent souvent à des erreurs de 1 (Beattie, 1979; Fischer, 1979)[(1)].

6.3. Les résultats, souvent plus psychologiques que didactiques, peuvent cependant, comme j'ai essayé de le montrer pour la soustraction, présenter un intérêt didactique immédiat au niveau de l'évaluation globale. D'autres inférences sont d'ailleurs possibles: le fait que les additions et soustractions fausses de 2 sont beaucoup mieux réussies que celles fausses de 1 suggère que les raisonnements corrects de parité n'ont pas joué un rôle important. Ceci ne prouve pas pour autant qu'ils sont totalement étrangers à des élèves de CM1: il se peut, par exemple, que les élèves qui, de manière intuitive, en auraient été capables, se retrouvent préférentiellement parmi ceux qui ont calculé exactement.

Mais au-delà de ces résultats immédiats, il me paraît important, si l'on veut utiliser la méthode des TR (dans une tâche de vérification) dans les classes, de s'assurer de sa sensibilité et de sa fiabilité en milieu scolaire. Car ce dernier implique des conditions (bruit, dérangement, etc ...) qui ne sont pas toujours celles des laboratoires de psychologie où sont typiquement étudiés les effets que j'ai rapportés. Or, sur ce plan, **les résultats** trouvés, la différence des pourcentages d'échec entre les égalités fausses de 1 et celles fausses de 2 (ou d'un multiplicateur) notamment, me **paraissent plutôt convaincants**.

(1) Il est intéressant de noter que:

1. Fischer (1979) a observé, au CE1 et avant l'apprentissage scolaire de la soustraction, une **procédure** soustractive **erronée** (voir p.65) et fréquemment utilisée (voir p.35), expliquant parfaitement l'échec de la vérification de 11-3=9;
2. Fischer et Pluvinage (soumis) soutiennent que les élèves, de CM2 en tout cas, ont souvent recours à une **mémoire procédurale** pour les soustractions;
3. Gadio (1987) a trouvé, en $6^{\text{ème}}$, **61% d'échecs** dans la vérification de 11-3=9.

REFERENCES

Ashcraft M.H., 1982. The development of mental arithmetic: A chronometric approach. *Developmental Review*, **2**, 213-236.

Ashcraft M.H. et Battaglia J., 1978. Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. *Journal of Experimental Psychology: Human Learning and Memory*, **4**, 527-538.

Ashcraft M.H. et Fierman B.A., 1982. Mental addition in third, fourth, and sixth graders. *Journal of Experimental Child Psychology*, **33**, 216-234.

Ashcraft M.H., Fierman B.A. et Bartolotta R., 1984. The production and verification tasks in mental addition: An empirical comparison. *Developmental Review*, **4**, 157-170.

Ashcraft M.H. et Stazyk E.H., 1981. Mental addition: A test of three verification models. *Memory & Cognition*, **9**, 185-196.

Beattie I.D., 1979. Children's strategies for solving subtraction-fact combinations. *Arithmetic Teacher*, **27**, 14-15.

Bradley J.V., 1968. Distribution-free statistical tests. Englewood Cliffs: Prentice-Hall.

Craik F.I.M. et McDowd J.M., 1987. Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **13**, 474-479.

Farell B., 1985. "Same"-"Different" judgments: A review of current controversies in perceptual comparisons. *Psychological Bulletin*, **98**, 419-456.

Findlay J.M., 1978. What form of memory do schoolchildren use whilst performing mental arithmetic ? In M.M. Gruneberg, P.E. Morris, R.N. Sykes (Eds), *Practical aspects of memory* . London: Academic Press.

Fischer J.P., 1979. La perception des problèmes soustractifs aux débuts de l'apprentissage de la soustraction. Thèse: Université Nancy I.

Fischer J.P., 1981. Développement et fonctions du comptage chez l'enfant de 3 à 6 ans. *Recherches en Didactique des Mathématiques*, **2**, 277-302.

Fischer J.P., 1987a. L'automatisation des calculs élémentaires à l'école. *Revue Française de Pédagogie*, **80**, 17-24.

Fischer J.P., 1987b. Les faits numériques à l'école: une étude développementale par les TR. *Psychologie Scolaire*, **60**, 7-24.

Fischer J.P. et Pluvinage F., soumis. Complexités de compréhension et d'exécution des opérations numériques élémentaires.

Gadio I., 1987. L'automatisation du calcul au début de l'école secondaire. (Manuscrit: DEA de Didactique des Mathématiques). Strasbourg: non publié.

Geary D.C. et Widaman K.F., 1987. Individual differences in cognitive arithmetic. *Journal of Experimental Psychology: General*, **116**, 154-171.

Geary D.C., Widaman K.F. et Little T.D., 1986. Cognitive addition and multiplication: Evidence for a single memory network. *Memory & Cognition*, **14**, 478-487.

Gonzalez E.G. et Kolers P.A., 1982. Mental manipulation of arithmetic symbols. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **8**, 308-319.

Hamann M.S. et Ashcraft M.H., 1985. Simple and complex mental addition across development. *Journal of Experimental Child Psychology*, **40**, 49-72.

Krueger L.E., 1986. Why 2x2=5 looks so wrong: On the odd-even rule in product verification. *Memory & Cognition*, **14**, 141-149.

Krueger L.E. et Hallford E.W., 1984. Why 2+2=5 looks so wrong: On the odd-even rule in sum verification. *Memory & Cognition*, **12**, 171-180.

Leach C., 1979. Introduction to statistics: A nonparametric approach for the social sciences. New York: Wiley.

LeFevre J.A. et Bisanz J., 1986. A cognitive analysis of number-series problems: Sources of individual differences in performance. *Memory & Cognition*, **14**, 287-298.

Mishkin M., 1982. A memory system in the monkey. *Philosophical Transactions of the Royal Society of London B*, **298**, 85-95.

Piaget J., 1977. Recherches sur l'abstraction réfléchissante : 1/ L'abstraction des relations logico-arithmétiques (tome 34 des EEG). Paris : PUF.

Ranney M., 1987. The role of structural context in perception: Syntax in the recognition of algebraic expressions. *Memory & Cognition*, **15**, 29-41.

Siegel S., 1956. Nonparametrics statistics for the behavioral sciences. New York: Mc Graw-Hill.

Snedecor G.W. et Cochran W.G., 1971. Méthodes statistiques. Paris: Association de Coordination Technique Agricole (6ème éd.).

Stazyk E.H., Ashcraft M.H. et Hamann M.S., 1982. A network approach to mental multiplication. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **8**, 320-335.

Wheeler L.R., 1939. A comparative study of the difficulty of the 100 addition combinations. *Journal of Genetic Psychology*, **54**, 295-312.

Wheeler L.R., 1941. A comparative study of the difficulty of learning the multiplication combinations. *Journal of Genetic Psychology*, **59**, 189-206.