

« Mesure dynamique du potentiel d'apprentissage chez les élèves »

SAWADOGO François*

Résumé

L'évaluation dynamique de l'intelligence s'inscrit dans une nouvelle conception de l'apprentissage fondée sur la théorie de la zone proximale de développement (ZPD) de Vygotsky. L'évaluation dynamique au lieu de tester la performance actuelle du sujet mesure le potentiel d'apprentissage (PA) qui correspondrait mieux à ses capacités réelles. Notre étude montre que la valeur prédictive de l'évaluation dynamique en milieu scolaire à l'aide d'un test d'intelligence de type facteur G à savoir le D70 est significativement plus élevée que celle d'une passation classique.

Mots-clés : test, évaluation, dynamique, zone proximale de développement, potentiel d'apprentissage.

« Dynamic measurement of the potential for student learning »

Abstract

Dynamic assessment of intelligence is part of a new concept of learning based on the theory of zone of proximal development (ZPD) Vygotsky. Dynamic assessment instead of the current test performance on measures of learning potential (AP), which would better reflect its actual capabilities. Our study shows that the predictive value of dynamic assessment in schools using an intelligence test factor type G is significantly higher than that of a conventional passation D70.

Keywords: test, evaluation, dynamic zone of proximal development, learning potential.

* Université de Koudougou (Burkina Faso), Laboratoire LAPAME, Membre Laboratoire CHArt/LUTIN (Paris 8)

Introduction

L'évaluation fait l'objet de nombreuses recherches de la part des psychologues notamment en milieu scolaire. L'enjeu principal est à la fois de mesurer les performances et de faire un diagnostic des potentialités et de la capacité d'un sujet à exécuter ou à réussir certaines tâches.

Différents tests ont été élaborés pour mesurer l'intelligence des sujets. On a distingué l'évaluation classique de l'intelligence qui positionne, à partir d'un outil d'évaluation, les aptitudes d'un sujet par rapport à son groupe de référence à une période précise de sa vie.

La seconde approche plus récente intègre davantage la capacité du sujet à améliorer ses performances quand il bénéficie d'une aide : c'est l'évaluation dynamique. Elle s'inscrit dans une nouvelle conception de l'apprentissage en éducation qui défend l'idée d'apprendre à apprendre. On parle d'éducabilité de l'intelligence. L'intelligence est perfectible grâce à une éducation cognitive.

De manière classique, pour apprécier le niveau d'intelligence d'un élève, on utilise des tests qui permettent de recueillir sa performance. Celle-ci permet de situer l'élève par rapport à la moyenne de son groupe d'appartenance. Cependant, on peut se demander si elle permet d'anticiper sur sa réussite future dans les activités scolaires. Dans cet article, nous présentons une recherche qui s'inscrit dans cette nouvelle approche et qui vise à analyser la valeur prédictive de l'évaluation dynamique en milieu scolaire à l'aide d'un test d'intelligence de type facteur G, à savoir le D70. Nous abordons dans un premier temps la problématique et les fondements théoriques de la recherche, puis nous présentons l'expérimentation réalisée, les résultats obtenus et enfin nos interprétations et conclusion.

I. Problématique

En plus des évaluations scolaires habituelles (devoirs et examens), le système éducatif fait parfois appel aux tests d'intelligence dans une perspective diagnostique ou/et d'orientation. Partant du constat que l'évaluation classique introduit des biais (FEURSTEIN *et al.*, 2002), s'inscrivant dans l'approche de Vygotsky qui défend l'idée d'une zone proximale de développement (ZPD), développe l'évaluation dite dynamique qui prend en compte l'aptitude du sujet à améliorer ses performances s'il bénéficie d'une aide. L'apprentissage en milieu scolaire consiste, du point de vue de la théorie de la ZPD, à permettre à l'élève d'évoluer en fonction de ce que sa zone proximale de développement lui permet.

En effet, l'évaluation classique mesure ce que le sujet a acquis en termes de connaissances et de stratégie de résolution de problèmes. Elle fournit ainsi un indicateur des compétences stabilisées du sujet. Ce niveau initial de l'élève est souvent utilisé pour prédire ses performances futures dans de nouvelles activités. Or, cette mesure ne s'intéresse pas à ce qu'il serait capable de faire mais à ce qu'il sait faire !

On peut donc supposer que dans une approche de diagnostic ou de pronostic des performances scolaires futures d'un élève, il serait plus adapté de procéder à une évaluation de type dynamique qui s'intéresse à ce qu'il serait capable de faire plutôt que de faire passer un test de manière classique.

L'aide introduite dans le cadre d'une évaluation dynamique, selon CHARTIER (2001), permettrait également de réduire les biais culturels liés à la passation classique en familiarisant le sujet avec la tâche du test. Dans les groupes sociaux défavorisés ou des milieux culturels non occidentaux, les tests et les testing ne sont pas des faits familiers. Les enfants de ces milieux sont souvent fortement dépaysés face à une situation où ils doivent subir un test psychologique. Ainsi, l'évaluation dynamique pourrait être d'un apport considérable dans l'accompagnement et la remédiation d'élèves.

L'évaluation dynamique a-t-elle réellement une meilleure valeur prédictive que l'évaluation classique en milieu scolaire ? Telle est la question que la présente recherche se propose d'explorer à travers une passation du D70 dans deux conditions : évaluation classique versus évaluation dynamique. Pour ce faire, nous nous proposons, dans le point qui suit, d'apporter un éclairage sur les notions d'évaluation dynamique et de potentiel d'apprentissage.

II. Notions d'évaluation dynamique et de Potentiel d'Apprentissage

L'évaluation en psychologie peut être considérée comme l'ensemble des méthodes et techniques qui permettent de mesurer certaines dimensions psychologiques des individus, comme les performances de mémorisation d'un élève ou encore le degré d'émotivité d'un sujet (LOARER et CHARTIER, 1996). Le champ de l'évaluation est aussi vaste que le champ de la psychologie elle-même.

Dans le présent travail, nous porterons notre intérêt particulièrement sur l'évaluation de l'intelligence. CHARTIER (1997), distingue trois grandes approches dans la mesure de l'intelligence :

- Une approche empirique qui consiste en l'élaboration d'échelles de niveau dans un but diagnostic. Les épreuves construites sur ce principe permettent de situer la performance du sujet sur une échelle et par rapport à une population donnée. Elle se fonde sur une définition pragmatique de l'intelligence développée d'abord par Alfred BINET (1857-1911) puis David WECHSLER (1896-1981) qui sont à l'origine des principaux tests de ce type. L'intelligence, pour eux, est la capacité globale de l'individu de penser d'une manière rationnelle et d'avoir des rapports utiles avec son milieu.
- Une autre approche développée par Charles Edward SPEARMAN (1863-1945) dite factorielle considère l'intelligence comme étant un ensemble qui résulte de l'interrelation entre plusieurs facteurs. On parvient ainsi à dégager un facteur général (G) qui est un indicateur de l'intelligence globale du sujet. Selon LAUTREY (2001), un rapprochement des deux tendances s'est opéré par la suite grâce aux modèles hiérarchiques qui révèlent l'existence de plusieurs facteurs

généraux associés à plusieurs types d'intelligence organisés en strates. On a ainsi principalement : un facteur général de l'intelligence fluide, un facteur de l'intelligence cristallisée et un facteur de l'intelligence visuo-spatiale. L'intelligence fluide correspond à des aptitudes cognitives de bases de l'individu qui lui permettent de s'adapter et de trouver des solutions dans des situations nouvelles. Cette forme d'intelligence est très sollicitée par les épreuves dites de facteur G « pur » tel que les Matrices Progressives de Raven. Ces épreuves font appel à l'attention ou à la mémoire de travail et à des aptitudes de traitement des règles de résolution et non à des connaissances particulières. L'intelligence cristallisée renvoie aux aptitudes acquises par l'individu sur les stratégies, les techniques et les procédures d'analyse et de résolutions. Elle est fortement liée à la richesse conceptuelle et à la complexité et du niveau d'organisation conceptuelle (LAUTREY, 2001). L'intelligence visuo spatiale se rapporte aux aptitudes de perception visuelle et de coordination spatiale des objets et des mouvements dans l'espace.

– Enfin une approche génétique qui est fondée sur la théorie piagétienne des stades du développement. PIAGET (1952), dans le cadre de ses recherches, a créé de nombreuses épreuves avec une conception structuraliste et développementale de l'intelligence. Sa théorie permettra d'élaborer des tests d'évaluation de l'intelligence tels que l'Échelle de la pensée logique (EPL) élaborée par LONGEOT et l'Échelle collective de développement logique de HORNEMANN (LAUTREY, 2001).

L'évaluation consiste alors à soumettre le sujet pris à un moment de son développement à la passation du test et à évaluer ses performances. Les résultats dépendent essentiellement du niveau de développement des structures de pensée du sujet acquises par la confrontation avec le monde physique si l'on se situe dans la théorie développementale piagétienne. VYGOSTKY (1933) qui a une approche du développement différente de celle de Piaget accorde une place importante à la médiation d'une action éducative dans le processus d'intériorisation des outils et la formation des fonctions psychologiques. Pour lui, l'action éducative, pour être efficace, doit s'insérer dans ce qu'il définit comme étant la zone proximale de développement. C'est l'écart entre ce qu'est capable de faire un enfant tout seul et ce qu'il sera à même de faire s'il bénéficie de l'aide d'un adulte ou de tout autre médiateur éducatif, (BÜCHEL, 1995). Feuerstein (HUTEAU & LAUTREY, 1999) sur la base de cette théorie de Vygostky va développer le concept d'évaluation dynamique.

Pour mesurer les aptitudes d'un sujet à suivre un apprentissage, son « learning potential », met au point des méthodes d'évaluation. Ces méthodes sont basées sur des épreuves classiques comme les matrices progressives de Raven ou la figure complexe de Rey. Dans la passation de ces tests classiques, il introduit une nouvelle dimension qui consiste en une aide du psychologue. L'examineur psychologue n'a plus une position passive mais intervient pour donner des informations sur la procédure, un feed back sur le déroulement de l'épreuve. Il devient médiateur selon l'approche de Vygotsky. L'important, dans ce type d'évaluation, est d'observer la manière dont le sujet gère ses difficultés et parvient à les surmonter. On recueille

ainsi des indicateurs sur son aptitude à apprendre, le sujet est évalué par rapport à lui-même et non par rapport à un groupe d'âge.

Ce qu'évaluent les tests classiques résulte d'apprentissages antérieurs. Les résultats à ces apprentissages sont eux-mêmes fonction de la capacité d'apprentissage du sujet et des contextes favorables ou non dans lesquels ils ont lieu. Par conséquent, il est préférable de s'intéresser aux modalités d'apprentissage du sujet (REY, 1934). La mesure du potentiel d'apprentissage est variable en fonction du paradigme expérimental utilisé. Dans le cas du paradigme test apprentissage re-test (T.A.R), il y a une première passation classique d'un test, suivie d'une phase d'aide à la compréhension de la tâche, puis une deuxième passation du même test. On apprécie alors la différence des scores au test et au re-test. Cette différence est alors appelée gain, elle correspond au potentiel d'apprentissage. Dans le cas du paradigme aide au cours du test (A.T), l'expérimentateur pendant la passation du test intervient pour aider à la résolution des items si le sujet échoue à un item.

Plusieurs paramètres sont pris en compte pour apprécier le potentiel d'apprentissage du sujet. Il y a notamment le mode de résolution des items, les manières d'échouer, le nombre d'aides sollicitées et la façon dont le sujet utilise ces aides qui lui sont apportées (LOARER & CHARTIER, 1994).

On observe une grande diversité dans la mise en œuvre des paradigmes évoqués ci-dessus tant au niveau des épreuves qu'au niveau de la mesure du PA. Il existe peu de tests standardisés pour faire de l'évaluation dynamique. Cela peut s'expliquer par le fait que, dans cette approche, l'expérimentateur joue un rôle nettement plus actif que dans les situations classiques de test, d'où une grande variabilité méthodologique d'un expérimentateur à l'autre. Pour réduire cette variabilité au niveau de la distinction de différents types d'aides et la quantification des aides (BÜCHEL *et al.*, 1990) proposent d'utiliser, dans le protocole expérimental, un formulaire qui permet de noter les différents types d'aides grâce à un système de codes.

Le potentiel d'apprentissage a suscité beaucoup d'intérêt de la part des chercheurs. On trouve dans HUTEAU et LAUTREY (1999) et dans LOARER. et CHARTIER. (1996) des présentations de travaux, notamment ceux de IONESCU et JORDAN-IONESCU (1986 ; 1987). Ces derniers ont proposé un test d'apprentissage à partir des cubes de Kohs où des aides progressives sont données au sujet. Ils ont également fait une analyse développementale du PA chez des enfants de 6 à 9 ans à l'aide de ce test.

II.1. Utilisation des Matrices progressives de Raven dans la mesure du potentiel d'apprentissage

BÜCHEL, De RIBAUPIERE et SCHARNHORST (1990), ont mesuré le potentiel d'apprentissage d'enfants âgés en moyenne de 10 ans grâce au test de matrices progressives (le PM 38), construit par Raven. L'évaluation s'est faite selon le schéma test aide-retest. Les sujets ont d'abord passé de manière classique le test PM38. Ensuite, les chercheurs ont fait intervenir une phase d'apprentissage sur les modes

de résolution des items. Pour ce faire, ils utilisent des items similaires n'ayant pas le même contenu issu du Set Variation II du LPAD (Learn Potential Assessment Device) de FEURSTEIN (1979). Les enfants ont ainsi appris que pour résoudre certains items par exemple, il faut distinguer le fond et la figure, et que la stratégie ligne/colonne consiste à trouver la relation sur la première ligne ou la première colonne, puis à l'appliquer pour retrouver l'élément manquant. Après cette phase d'aide qui dure environ 40 minutes, les sujets font à nouveau une passation classique du test.

Les auteurs ont ensuite mesuré la différence entre les scores au re-test et le score au test1 pour obtenir le gain de l'apprentissage. L'un des objectifs de leur recherche était de mesurer la stabilité dans les gains obtenus par les sujets d'une passation à l'autre. Les résultats de l'étude révèle une faiblesse de la stabilité du gain. Les auteurs observent par contre une relative stabilité du nombre d'erreurs commises par sujet par rapport au nombre d'aides fournies.

II.2. Evaluation dynamique et éducation

L'intérêt de l'évaluation dynamique peut se situer dans une perspective d'éducation cognitive. Dès les années 1960, un courant d'éducation cognitive s'est développé aux Etats-Unis dans une tentative de réponse à des disparités sociales observées en milieu scolaire.

Il s'agit du programme d'éducation compensatoire dénommé « head start » 1964-1965 (LOARER *et al.*, 1995) qui visait à aider les enfants issus de couches sociales défavorisées noires et hispaniques à améliorer leurs aptitudes intellectuelles. De grands espoirs étaient placés sur ce type de programme.

Des auteurs comme HUNT (1961) et BLOOM (1964), pensaient que l'éducation compensatoire pouvait augmenter de près de 30 points un QI (LOARER *et al.*, 1995). Mais les premiers résultats obtenus par ce programme n'ont pas révélé d'amélioration significative. Ces dernières années, on observe un regain pour l'éducation cognitive. Cela est en partie dû au dysfonctionnement des systèmes d'enseignement qui ne parviennent pas à conduire un nombre important d'élèves à un niveau de qualification minimal. Près de 100 000 élèves en France par an sortent du système scolaire sans atteindre le niveau 5^e (HUTEAU, 1999). Il y a également un développement des recherches en sciences cognitives sur le caractère perfectible de l'intelligence contrairement aux thèses héréditaires qui sont délaissées (BÜCHEL *et al.*, 1995).

Ainsi apparaît un développement au niveau expérimental en termes de méthode et d'outils d'éducation cognitive. Ces méthodes visent à doter d'emblée le sujet d'outils de pensée généraux par le choix de méthode d'apprentissage pauvre en contenu (LOARER *et al.*, 1995). On va chercher à améliorer la capacité générale du sujet à apprendre et non de manière spécifique un rapport avec une discipline donnée. C'est dans cette optique que Feuerstein a développé son Programme d'enrichissement instrumental, le PEI. Ces méthodes vont mettre l'accent sur les procédures et

les stratégies, sur les facteurs cognitifs et la capacité du sujet à exercer un contrôle sur ses modes de pensée et de réflexion, à savoir une capacité de métacognition.

L'intérêt de l'évaluation dynamique pour le praticien de l'éducation peut se situer à plusieurs niveaux. Dans le milieu éducatif, en fonction de la finalité, on distingue deux grands types d'évaluation : une évaluation dite sommative qui fait le point des acquisitions de l'élève à un stade donné de son cursus scolaire, et une autre dite formative dont le but est de faire progressivement à chaque étape de l'apprentissage le point sur les capacités de l'élève dans le souci de les améliorer. L'évaluation dynamique pourrait bien s'insérer dans cette approche d'évaluation formative. En effet, l'enseignant, sans faire vraiment de la passation de test, peut, en s'inscrivant dans une démarche d'évaluation dynamique, accorder un plus grand intérêt à la manière d'apprendre de l'apprenant. Ainsi, il pourra tenir compte du potentiel d'apprentissage de l'apprenant afin de lui proposer des apprentissages qui s'intègrent dans sa zone proximale de développement et enrichir ainsi ses aptitudes cognitives (LOARER et CHARTIER, 1996).

III. Objectifs de la recherche

Dans cette recherche, notre objectif général est de développer une méthode d'évaluation du potentiel d'apprentissage en milieu scolaire sur la base du test D70.

Les objectifs spécifiques sont :

- i) mesurer l'impact des aides introduites dans un paradigme test- aide- re-tests (TAR)
- ii) analyser la valeur prédictive du potentiel d'apprentissage sur les performances scolaires.

IV. Hypothèses et Variables

IV.1. Hypothèses

Partant du postulat que les élèves après avoir bénéficié d'une aide à la phase d'apprentissage auront de meilleures performances à la seconde passation du D70, nous formulons les deux premières hypothèses suivantes :

H1 : les résultats des sujets au test 2 seront meilleurs que leurs résultats au test 1

H2 : Tous les sujets obtiendront un gain positif au test 2 par rapport au test 1

Prédiction

Par ailleurs, le score du test 2 intègre la capacité du sujet à améliorer ses compétences s'il obtient de l'aide. Ce score est donc plus révélateur des potentialités du sujet et de ses aptitudes en situation d'apprentissage scolaire. Il aura une plus grande valeur prédictive des performances scolaires. Nous faisons donc l'hypothèse suivante :

H3 : les résultats du test 2 auront une plus grande valeur prédictive des performances scolaires des sujets que ceux du test 1.

Enfin, eu égard à des études précédentes LOARER et CHARTIER (1996) dont la tendance est de penser que ce serait au niveau des élèves en difficultés que l'aide apporterait de manière significative une amélioration, nous faisons l'hypothèse suivante :

H4 : les résultats du test 2 auront une plus grande valeur prédictive des performances scolaires du deuxième trimestre notamment chez les sujets ayant des faibles notes scolaires (Moyenne Scolaire <10) par rapport aux autres (Moyenne Scolaire \geq 10)

IV.2. Les variables

– Variable indépendante

Il s'agit du niveau initial des élèves, on distinguera deux types de moyennes :

- Les élèves ayant une moyenne scolaire supérieure ou égale à 10,
- Les élèves ayant une moyenne scolaire inférieure à 10.

– Variables dépendantes

La première variable dépendante est le score au test D70. Nous distinguons le score D70 de la première passation classique (score D70_1) de celui de la seconde passation du test D70 (Score D70_2). La seconde variable dépendante est la performance scolaire du premier et du deuxième trimestre en science et en langue.

On distinguera deux types de moyennes :

- une moyenne scientifique du premier trimestre (M_Science) obtenue à partir des notes en mathématiques, en physique et chimie, en sciences de la vie et de la terre et en technologie puis une moyenne littéraire du premier trimestre (M_Langue) obtenue à partir des notes en français, en histoire et géographie, en langue vivante 1 et en langue vivante 2.
- une moyenne scientifique du deuxième trimestre (M2_Science) obtenue à partir des notes en mathématiques, en physique et chimie, en sciences de la vie et de la terre et en technologie.
- une moyenne littéraire du deuxième trimestre (M2_Langue) obtenue à partir des notes en français, en histoire et géographie, en langue vivante 1 et en langue vivante 2.

V. Méthodologie

Comme nous l'avons évoqué dans la partie théorique, il n'existe pas vraiment d'épreuve standardisée que l'on pourrait systématiquement utiliser pour une évaluation dynamique. Les chercheurs, dans la plupart des cas, ont conçu des dispositifs spécifiques avec des épreuves adaptées à leurs objectifs d'étude. Afin de pouvoir explorer les questions soulevées dans notre problématique, nous avons élaboré un dispositif expérimental adapté à notre population de collégiens et au mode de passation collective.

L'utilisation du paradigme AT indiqué dans une démarche de prise en charge individuelle n'est pas très adaptée pour notre étude. En effet, si celui-ci est très utilisé, notamment en clinique, il est difficile à mettre en œuvre dans une passation collective. Par contre, avec le paradigme TAR, la passation collective est possible et l'aide est mieux standardisée.

Après avoir retenu d'inscrire notre dispositif dans le cadre du paradigme test-aide-retest, notre préoccupation a été de trouver : d'une part un test utilisable en passation collective sur des collégiens et, d'autre part, des items adaptés pour la phase d'apprentissage. Pour le test nous avons opté pour le D70 de Kourovsky et Rennes (Manuel d'application, 1970) qui est un des tests de facteur G qui correspond à ce type de public. Le test a également fait l'objet d'une analyse très précise au niveau de ses composantes et des procédures de résolution des items (CHARTIER, 2001; DICKES & MARTIN, 1998). Le fait que ce soit un test « papier-crayon » a également été déterminant à cause de la passation collective. En plus, le test ayant été construit sur le même principe que le D48 de Pichot (Manuel d'application, 1959) avec lequel il y a une similarité d'items, nous avons la possibilité d'exploiter ce dernier test pour la phase d'apprentissage.

A l'aide de 11 items du D48, nous avons donc élaboré un cahier d'exercices pour la phase d'apprentissage. Pour ce faire, nous avons d'abord identifié la procédure de résolution de chaque item du D48. Notre classification est celle en 4 types d'items utilisée par CHARTIER (2001) pour le D70, à savoir des items spatiaux, numériques, arithmétiques et mixtes. Ensuite, nous avons retenu les 11 items pour le cahier d'exercices en veillant à une représentation des quatre types.

Conformément au paradigme TAR, nous avons prévu une première passation classique du D 70 suivie d'une phase d'apprentissage à l'aide du cahier d'exercices et enfin un re-test avec une passation du D70.

1. Les Participants

L'étude a été menée auprès de 48 élèves dont 26 filles et 22 garçons âgés de 14 à 16 ans de deux classes de 4^e de la région île de France. Toutefois suite à l'absence de 8 élèves pour diverses raisons (mortalité expérimentale) à la seconde phase de l'expérimentation, nous avons fait nos analyses sur un échantillon de 40 élèves dont 21 filles et 19 garçons. Pour les besoins d'une partie de nos analyses, nous avons considéré deux groupes expérimentaux d'élèves selon les notes scolaires. Un sous-

groupe G1 des élèves ayant une moyenne scolaire supérieure ou égale à 10 et un deuxième sous-groupe G2 ayant une note moyenne scolaire inférieure à 10.

Dans le groupe des sujets ayant une moyenne supérieure ou égale à 10, on aura des sous-groupes pour chaque type de matières. Ainsi, nous avons les sous-groupes d'élèves ayant une moyenne en sciences supérieure ou égale à 10 au premier trimestre (M_Science \geq 10) et au deuxième trimestre (M2_Science \geq 10). Il en est de même pour les matières littéraires où il y a les sous-groupes des élèves ayant une moyenne en langues supérieure ou égale à 10 au premier trimestre (M_Langue \geq 10) et au deuxième trimestre (M2_Langue \geq 10). Pour le groupe des sujets ayant une moyenne inférieure à 10, il y a les sous-groupes des élèves ayant une moyenne en sciences inférieure à 10 au premier trimestre (M_Science < 10) et au deuxième trimestre (M2_Science < 10), puis ceux des élèves ayant une moyenne en langues inférieure à 10 au premier trimestre (M_langue < 10) et au deuxième trimestre (M2_langue < 10).

Tableau 1 : Sous- groupes expérimentaux

Sujets	Groupes	Moyennes scolaires
40 élèves de 4 ^e	G1	M_Science \geq 10 M2_Science \geq 10 M_Langue \geq 10 M2_Langue \geq 10
	G2	M_Science < 10 M2_Science < 10 M_langue < 10 M2_langue < 10

V.1. Matériel

Caractéristique du D70

Le D70 est un test d'efficiences intellectuelle qui mesure une intelligence générale de type facteur G. La tâche consiste pour le sujet à trouver le domino manquant dans un ensemble de dominos. Il existe une règle de progression pour trouver la solution. Le test est composé de quarante quatre (44) items. Une analyse de Dickes et Martin (CHARTIER (2001) permet de regrouper selon le mode de résolution en quatre types ces items :

- Les items spatiaux qui se résolvent grâce à un traitement spatial par l'utilisation entre autres de règles de symétrie, de répétition et d'inversion ;
- Les items numériques qui nécessitent l'utilisation de règles d'incrémentations soit par l'ajout ou le retrait ;
- Les items mixtes où le sujet doit faire appel à la fois à une règle numérique et à une règle spatiale pour la résolution ;
- Les items arithmétiques qui se résolvent par l'application d'une opération arithmétique simple du type $a+b = c$.

Tableau 2 : Répartition des items du D70 (44 items au total)

	Items spatiaux	Items numériques	Items mixtes	Items arithmétiques
Effectif	20	12	9	3
Pourcentage	45,5	27,3	20,4	6,8

Source : CHARTIER (2001)

V.2. Procédure

La méthode de test-aide-retest a été mise en œuvre de la manière suivante : Une passation classique du D70 suivi d'un apprentissage à l'aide du cahier d'exercices (11 items du D48), puis un re-test du D70 dans des conditions classiques.

Les passations classiques du D70

Les passations classiques du D70 se sont faites collectivement, chaque élève disposant d'un cahier du D70 et d'une feuille de réponses. A chacune de ces phases de test classique, les élèves ont été divisés en 2 groupes de 25 en moyenne. La consigne utilisée est celle du manuel du D70 : « Dans ce test, chaque dessin représente un groupe de dominos. Le nombre de points sur chaque moitié de domino peut varier de 0 à 6. Regardez chaque groupe et recherchez la valeur du domino qui manque. Inscrivez sur la feuille de réponses les chiffres correspondant à ce domino ».

La page des 4 items exemples est projetée à l'aide d'un transparent, tous les élèves ont ainsi une même présentation sur les exemples.

La phase d'apprentissage

La phase d'apprentissage qui a suivi immédiatement le test 1 s'est faite à l'aide du cahier d'exercice élaboré à partir de 11 items du D48. Le cahier d'exercices est construit comme évoqué dans le point élaboration du dispositif de manière à avoir une représentation proportionnelle des différents types d'items contenus dans le D48 et dans le D70. Nous avons retenu 3 items pour chacun des types spatial, numérique et arithmétique et pour le type mixte 2 items.

L'aide apporté consiste à faire découvrir aux élèves : l'existence de différences entre les items ; les différents types d'items possibles (spatial, numérique, arithmétique, mixte) et le caractère aléatoire de la présentation de ces types d'items.

Cet apprentissage s'est fait en alternant des phases collectives et individuelles. Nous expliquons aux élèves que nous allons faire ensemble d'autres exercices qui leur permettront de mieux comprendre comment résoudre ce genre de problèmes. Pour ce faire, nous avons élaboré une fiche d'exercices (cf. annexes) contenant quatre items de types différents 3 (spatial), 34 (numérique), 26 (mixte), 43 (arithmétique). Chaque élève est invité à essayer de trouver la valeur du domino manquant dans ces dessins. A l'issue de cet exercice individuel qui dure environ 5 minutes, nous passons à une phase collective de correction. Les items sont successivement projetés à l'aide d'un rétroprojecteur au tableau. On procède à un recueil des procédures de

quelques élèves en suscitant des échanges pour aboutir à la bonne stratégie. Après s'être assuré que le type d'item (spatial pour le 3) est bien compris des élèves, on passe à l'item suivant. Les sept autres items du cahier d'exercices sont ensuite projetés de manière alternée en expliquant aux élèves que dans les tests les types sont mélangés. Une résolution collective de chaque item est faite avant de passer à un autre. Une semaine plus tard, nous procédons au re-test, la seconde passation classique du D70.

En résumé, chacun de nos sujets fait d'abord une passation classique du D70 suivi d'une séance d'apprentissage à l'aide d'items du D48 puis une semaine plus tard fait une seconde passation du D70. A l'issue de ces expérimentations nous avons obtenu les résultats que nous présentons ci-dessous.

VI. Résultats et discussions

VI.1. Mesure du potentiel d'apprentissage : comparaison des scores aux deux tests du D70 (score D70_1 & score D70_2)

Au niveau des deux scores au D70, nous avons fait deux hypothèses sur le principe que les élèves après avoir bénéficié d'une aide à la phase d'apprentissage amélioreront leurs performances au re-test :

- l'une (H1) prédisait que les résultats des sujets au test 2 seront meilleurs que leurs résultats au test 1 à savoir que $\text{score D70}_2 > \text{score D70}_1$;
- l'autre hypothèse (H2) postulait que tous les sujets obtiendront un gain positif au test 2 par rapport au test 1 c'est-à-dire que chacun d'eux aura : $\text{score D70}_2 - \text{score D70}_1 \geq 0$.

Les moyennes des scores D70_1 et D70_2 sont respectivement de 17,8 pour le premier test et de 21,8 points pour le second. La corrélation entre les deux scores est de .65, elle est significative au seuil .01. La comparaison des deux moyennes donne une différence significative pour un $t = 4,06$ au seuil $\alpha : .01$. Le score moyen des élèves au test 2 est donc supérieur à celui du test 1. Le calcul des écarts entre le score D70_2 et le score D70_1 donne pour 29 sujets un gain positif, pour 3 sujets un gain égale à 0 et pour les 8 autres des écarts négatifs. Les écarts entre les deux scores varient de 0 à 19 points en valeur absolue (cf. graphique 1) si l'on compare à la performance initiale de chaque sujet. Le traitement statistique donne un gain moyen de 4 points avec un écart type de 6, 18.

Les résultats de notre expérimentation chez des élèves de 4ème de collège âgés de 14 à 16 ans donnent un score du D70 obtenu à la suite d'une procédure d'évaluation dynamique supérieur au score initial obtenu par une passation classique du D70 significatif à .01. Ces résultats confirment notre première hypothèse (H1) qui prédisait une amélioration des scores au test 2.

Ainsi, il y a eu une amélioration des performances au test des élèves entre les deux passations. Le gain obtenu signifierait que la phase d'apprentissage à l'aide d'items

du D48 similaires à ceux du D70 introduite entre les deux passations a permis aux élèves d'améliorer leurs performances. Ce qui serait conforme à la théorie sur les évaluations dynamiques qui partent du principe de l'existence d'une zone proximale de développement chez l'enfant défini par Vygotsky. Celui a considéré que l'introduction d'une aide dans une procédure d'évaluation permet d'obtenir de meilleures performances du sujet liées à son potentiel d'apprentissage.

La passation de D70 dans une procédure d'évaluation dynamique que nous avons faite a permis de placer les élèves dans des conditions proches des situations habituelles d'apprentissages en milieu scolaire. Nos résultats, de ce point de vue, ont une valeur écologique qui autoriserait à faire des extrapolations à partir du potentiel d'apprentissage observé chez un élève lors d'une évaluation dynamique (avec le D70) sur son potentiel dans le cadre d'apprentissage purement scolaire. Toutefois, n'ayant pas pu avoir un groupe témoin qui n'aurait fait que deux passations classiques du D70 sans phase d'apprentissage, nous sommes obligé de relativiser nos résultats sur le gain lié à l'apprentissage, car cette amélioration des performances est également due à l'effet du re-test.

Notre deuxième hypothèse (H2) qui prévoyait que tous nos sujets auront un gain positif au re-test n'est pas confirmée. Car si le gain moyen de 4 points est positif, ce n'est pas le cas pour tous les sujets pris individuellement. Huit (8) de nos sujets ont des écarts négatifs. Ce qui signifie que l'aide apportée n'a pas eu pour effet d'améliorer les performances de ces sujets mais à plutôt entraîner une régression. Un tel effet qui ne va pas dans le sens de la logique de l'aide pourrait s'expliquer par le fait que la phase d'aide aurait perturbé la compréhension du sujet du fait des explications de l'expérimentateur. Mais vu que la majorité des sujets (32) n'ont pas régressé, cette explication est peu plausible.

L'autre hypothèse explicative serait l'interférence du facteur motivation que la passation collective ne nous permettait pas de contrôler. En effet, si le sujet à la seconde passation ne s'applique pas pour faire le test, cela peut biaiser l'évaluation de sa performance.

VI.2. Valeur prédictive du potentiel d'apprentissage : Relation scores des tests & performances scolaires

Pour les relations entre les scores aux tests, nous avons fait l'hypothèse (H3) suivante : les résultats du test 2 auront une plus grande valeur prédictive des performances scolaires des sujets que ceux du test 1. Donc : la liaison entre les scores D70_2 et moyennes scolaires > la liaison scores D70_1 et les moyennes scolaires. Le calcul des corrélations entre les scores aux tests et les résultats scolaires donne les résultats suivants :

Tableau 4 : coefficients de corrélation entre les scores au test et les moyennes des notes scolaires.

Variables	M_SCIENCE	M_LANGUE	M2_SCIENCE	M2_LANGUE
ScoreD70_1	0,39*	0,27	0,27	0,23
ScoreD70_2	0,59**	0,45**	0,40*	0,46**

* la corrélation (Pearson) est significative au seuil .05

** la corrélation (Pearson) est significative au seuil .01

Les résultats montrent que la liaison entre les scores au test D70 et les moyennes scolaires est assez faible pour le test 1. Par contre, la liaison est plus importante entre les notes scolaires et les scores du test 2. Si pour le score D70_2 on a des liaisons significatives avec les différentes moyennes des notes scolaires, pour le score D70_1 on observe qu'une seule liaison significative avec la moyenne en sciences du premier trimestre.

L'analyse des résultats révèle d'assez fortes liaisons entre le score D70_2 et les différentes moyennes des notes scolaires significatives au seuil .01 pour la moyenne des matières scientifiques du premier trimestre (M_sciences) et les moyennes littéraires du premier trimestre et du deuxième trimestre (M_langue et M2_langue) et au seuil .05 pour la moyenne scientifique du deuxième trimestre (M2_sciences). Pour toutes les moyennes scolaires considérées, les liaisons avec le score D70_2 sont nettement plus élevées que celles avec le score D70_1. En effet, avec le score D70_1, on n'a qu'une seule liaison significative celle avec la moyenne sciences du premier trimestre au seuil .05. La valeur prédictive des performances scolaires de l'évaluation dynamique au vu de nos résultats semble donc être plus grande que celle de l'évaluation classique dite statique.

VI.3. Variation de la liaison des scores au test D70 avec les performances selon le niveau scolaire

Postulant que c'est chez les élèves en difficultés que l'aide apportera de manière significative une amélioration, nous avons fait l'hypothèse (H3) suivante : les résultats du test 2 auront une plus grande valeur prédictive des performances scolaires du deuxième trimestre notamment chez les sujets ayant des faibles notes scolaires (Moyenne Scolaire <10) par rapport aux autres (Moyenne Scolaire \geq 10). C'est-à-dire que :

- la liaison entre score D70_2 et M2_sciences (<10) > la liaison entre score D70_2 et M2_sciences (\geq 10) ;
- la liaison entre score D70_2 et M2_Langue (<10) > la liaison entre M2_langue (\geq 10).

Pour apprécier cette variation selon le niveau scolaire, nous avons traité les résultats en scindant notre échantillon en sous-groupes à partir des critères de niveau défini dans la partie méthodologie : les sous-groupes de sujets ayant une moyenne supérieure ou égale à 10 pour chaque type de matière, et les sous-groupes de sujets

ayant une moyenne inférieure à 10. Nous obtenons des corrélations très faibles entre le score D70_1 et les moyennes scolaires et légèrement plus élevées avec le score D70_2. Mais aucune de ces corrélations n'est significative statistiquement.

Tableau 5 : Différentes corrélations observées selon les groupes de performances scolaires

Variables	M_Science	M_Langue	M2_Science	M2_Langue	M_Science	M_Langue	M2_Science	M2_Langue
	<10	<10	<10	<10	<10	>10	>10	>10
ScoreD70_1	0,01	-0,06	-0,14	0,54	0,16	0,13	0,35	0,16
ScoreD70_2	0,38	0,03	-0,15	0,53	0,33	0,21	0,40	0,41

On observe de manière paradoxale des corrélations négatives entre ScoreD70_1 et M_langue <10, entre score D70_1 et M2_Science<10 et entre score D70_2 et M2_Science<10. A partir du postulat que c'est aux élèves en difficultés que l'aide apporte de manière significative une amélioration, nous avons fait l'hypothèse H3 selon laquelle les résultats du test 2, score D70_2, auraient une plus grande valeur prédictive des performances scolaires que ceux du test 1, score D70_1, chez les sujets ayant des faibles notes scolaires (moyennes scolaires <10) par rapport aux autres (moyennes scolaire ³10). Nos résultats sur ces différences de la valeur prédictive du score D70 en fonction du niveau des élèves ne donnent aucune corrélation significative. Notre hypothèse n'est donc pas confirmée. Comment expliquez-vous cela ?

On observe également de manière paradoxale des corrélations négatives entre ScoreD70_1 et M_langue <10, entre score D70_1 et M2_Science<10 et entre score D70_2 et M2_Science<10. Ceci signifierait que la relation entre les scores au D70 et cette moyenne scientifique du deuxième trimestre est inversement proportionnelle. Ce résultat est probablement lié à un biais introduit soit par des notes très inhabituelles obtenues par les élèves au deuxième trimestre dans une des matières scientifiques en l'occurrence les mathématiques. L'analyse des corrélations entre les scores D70 et les matières spécifiques donne effectivement des corrélations très faibles pour les notes obtenues en maths au deuxième trimestre.

Dans le souci d'affiner nos analyses sur le lien entre les scores aux tests et les performances au niveau scolaire, nous nous sommes intéressé aux liaisons spécifiques avec des disciplines comme les mathématiques, la physique chimie et le français. Nous obtenons des corrélations très différentes selon la matière considérée.

Tableau 6 : récapitulatif des coefficients de corrélation entre les scores au test et les notes scolaires dans des disciplines spécifiques

Variables	Maths	Maths_2	PC	PC2	Français	Français 2
ScoreD70_1	.46**	.13	.41**	.14	.31	.28
ScoreD70_2	.42**	.26	.57**	.47**	.50**	.53**

* la corrélation (Pearson) est significative au seuil .05

* la corrélation (Pearson) est significative au seuil .01

Nous observons pour le score D70_2 des liaisons significatives avec la plupart des matières sauf avec les mathématiques au deuxième trimestre. En effet, au deuxième trimestre la liaison avec les mathématiques n'est pas du tout significative avec un coefficient de corrélation très faible .26. Avec le score D70_1 on n'a des liaisons significatives que dans deux cas : la liaison score D70_1 et Mathématiques du premier trimestre avec un coefficient de corrélation .46 significatif au seuil .01 ; l'autre liaison entre la physique chimie et le score D70_1 est également significative avec un coefficient de corrélation .41 au seuil .01.

L'analyse des corrélations avec des disciplines spécifiques à savoir mathématiques, physique chimie et français confirme le fait que la liaison est plus forte entre les performances scolaires et le score D70_2 qu'avec le score D70_1. Cependant, nous avons observé que paradoxalement, les corrélations avec le français sont légèrement plus élevées qu'avec les maths. Or, les items du D70 étant plus proches des problèmes de type mathématique, nous nous attendions à une tendance opposée. Cette situation pourrait tout simplement signifier que le test D70 fait effectivement appel à des fonctions générales de l'intelligence du sujet sans être liées à un type de raisonnement purement mathématique.

Conclusion

Les praticiens dans le secteur de l'éducation, conseillers d'orientation psychologues entre autres, ont comme préoccupation récurrente, d'avoir des éléments permettant de pronostiquer sur la capacité d'un élève à suivre telle ou telle formation. C'est également une préoccupation des parents et des élèves. Très souvent ils s'appuient sur les notes antérieures ou bien sur les résultats d'évaluation de type sommative sur le bilan des acquisitions de l'élève, ou encore une évaluation par des tests des capacités intellectuelles ou d'intelligence de l'élève. Ces évaluations de type statique ne prennent pas nécessairement en compte le potentiel d'apprentissage du sujet à savoir si ce dernier bénéficie d'une aide. Par conséquent ces évaluations restent partielles sur la capacité réelle du sujet. Elles ne reflètent que ce que le sujet est capable de faire tout seul à un moment donné.

Dans notre étude, nous avons, à travers une expérimentation fondée sur le principe de l'évaluation dynamique, mesuré non seulement les performances d'élèves de 4^e au test d'intelligence D70 produites dans des conditions classiques de test (score D70_1), mais également leurs performances au même test après qu'ils ont bénéficié d'une aide à la compréhension des procédures de résolution aux items du test. Les résultats de l'expérimentation montrent que les élèves ont considérablement amélioré leurs performances au test après la phase d'apprentissage. En effet, le score D70_2 est supérieur au score D70_1 de manière significative.

Dans une approche pronostique du potentiel intellectuel d'un élève, il serait ainsi préférable d'utiliser une procédure d'évaluation dynamique plutôt qu'une mesure par une passation classique de test. Cette idée est confortée par les résultats obtenus sur les liaisons entre les deux scores au D70 des élèves et leurs performances

scolaires. Sur l'ensemble des matières considérées pour ces élèves de 4^e, nous avons des liaisons positives significatives entre le score D70_2 obtenu après la phase d'apprentissage et les moyennes scolaires. A l'inverse, avec le score D70_1, les liaisons avec les moyennes scolaires ne sont pas significatives sauf avec la moyenne scientifique du premier trimestre, et même dans ce cas la liaison avec le score D70_2 est plus forte. Néanmoins, n'ayant pas pu avoir un groupe contrôle qui n'aurait fait que deux passations classiques du D70 sans phase d'apprentissage pour mesurer la part d'un simple retest, ces résultats sont à relativiser.

Si, de manière globale, pour l'ensemble des sujets la valeur prédictive des performances scolaires du score du D70 obtenu à l'issue de l'évaluation dynamique est plus élevée que celui obtenu par la passation classique initiale, on n'observe pas de différence significative selon le niveau des élèves. Toutefois, cette absence de différence significative pourrait résulter du fait que constituer des sous-groupes selon le niveau pour un nombre total d'élèves si petit (40) ne permet pas un traitement statistique pertinent.

Références bibliographiques

- BÜCHEL F. P., 1995.** L'éducation cognitive. Le développement de la capacité d'apprentissage et son évaluation, Neuchâtel, Delachaux & Niestlé.
- BÜCHEL F. P., De RIBAUPIERRE A., & SCHARNHOST U., 1990.** Le diagnostic du potentiel d'apprentissage par LPAD : une étude de fidélité, *European journal of psychology of education*, Spécial issue, 5 :135-158.
- BERNAUD J. L., 2000.** préface, in D. Pasquier, Manuel du test d'évaluation dynamique de l'éducabilité (TEDE), éd. cabinet avenir &entreprise.
- CHARTIER P., 2001.** Les apports de la recherche en psychologie aux tests d'intelligence : quelles conséquences pour la pratique ? *L'Orientation scolaire et professionnelle*, 4 : 511-533
- CHARTIER P., 2001.** Les apports de la recherche en psychologie aux tests d'intelligence : quelles conséquences pour la pratiques ?, *l'orientation scolaire et professionnelle*, 30 (4) : 511-533.
- CHARTIER D., 1997.** Les tests d'intelligence, *Evaluer l'intelligence*, Paris, Inetop,
- DICKES P., & MARTIN R., 1998.** Les composantes de l'intelligence générale du D70, *Psychologie et Psychométrie*, 19,1.
- FEURSTEIN R., FALIK L. & RAND Y., 2002.** Dynamic assessments of cognitive modifiability. ICELP Press, Jerusalem: Israel
- HUTEAU M., & LAUTREY J., 2003.** Evaluer l'intelligence, *Psychométrie cognitive*, Paris, Puf.
- HUTEAU M., 1995.** Manuel de psychologie différentielle, Paris, Dunod.
- LOARER E., & CHARTIER D., 1994.** Le potentiel d'apprentissage est-il général ou spécifique au domaine d'apprentissage ? , In HUTEAU, M., (coord), les techniques psychologiques d'évaluation des personnes, Issy-les -Moulineaux, EAP.
- LOARER E. & CHARTIER D., 1996.** L'évaluation dynamique des aptitudes : révolution ou gadget ? , *Psychologie française*, 41 : 35-46.
- CHARTIER P., & LOARER E., 2008.** Evaluer l'intelligence logique : Approche cognitive et dynamique, Dunod
- LOARER E., HUTEAU M., CHARTIER D., & LAUTREY J., 1995.** Peut-on éduquer l'intelligence ? L'évaluation d'une méthode d'éducation cognitive, Berne, Lang.

LAUTREY J., 2001. L'évaluation de l'intelligence état actuel et tentatives de renouvellement, in HUTEAU, M., (coord), Les figures de l'intelligence, Paris, EAP : 19-42.

MANUEL D'APPLICATION, 1959. Test D48. Paris : Editions du Centre de Psychologie Appliquée

MANUEL D'APPLICATION, 1970. Test D70. Paris : Editions du Centre de Psychologie Appliquée

PAOUR J. L., JAUME J. & De ROBILLARD O., 1995. De l'évaluation dynamique à l'éducation cognitive : repères et questions. In BÜCHEL, F., 1995. L'éducation cognitive. Le développement de la capacité d'apprentissage et son évaluation, Neuchâtel, Delachaux & Niestlé.

REY A., 1934. D'un procédé pour évaluer l'éducabilité, Archives de Psychologie, Neuchâtel, n° 24, 96 p

YIGOTSKY L.S., 1933. Pensée et langage (traduction de Sève, F., 1997), Rééditions : Paris, La Dispute.