

Rapport de recherche

Grilles ARBA-2 d'analyse des ressources et besoins d'aide

Rapport d'évaluation psychométrique

Valentino Pomini

Philippe Golay

Marjorie Blanc

Unité de Réhabilitation,
Service de Psychiatrie Communautaire,
Département de Psychiatrie, CHUV

1. INTRODUCTION

Les grilles ARBA (Analyse des Ressources et Besoins d'Aide) ont été élaborées en 2003 par le Département de l'Action Sociale et de la Santé (DASS) de Genève en étroite collaboration avec la section genevoise des institutions sociales suisses pour personnes handicapées (INSOS). L'élaboration de ces grilles d'évaluation intervient dans le cadre de la mise en place de la loi sur l'intégration des personnes handicapées (LIPH K 1 36) et plus particulièrement des questions relatives à la gestion des institutions qui les accueillent. En effet, dès 2011 le financement de ces institutions sera à la charge non plus de la Confédération mais des cantons (loi RPT concernant la répartition des tâches entre la Confédération et les Cantons). Dans ce contexte, la Fédération des Ateliers pour personnes Handicapées (FAH) a souhaité pouvoir équiper les décideurs d'un outil utilisable dans plusieurs cantons permettant d'estimer les coûts de chaque structure. L'objectif visé de la grille ARBA est donc de produire une évaluation unifiée des besoins d'aide des personnes handicapées, qui puisse devenir l'un des éléments pris en compte pour la répartition des subsides entre les différents ateliers.

ARBA est un outil d'évaluation « trans-handicaps » pouvant être utilisé dans toutes les institutions oeuvrant dans le domaine du handicap. Selon ses auteurs, l'outil se veut objectif, simple, adapté aux différents établissements d'accueil et applicable à de nouveaux types d'atteintes ou maladies non encore répertoriées (DSE-DGAS, 2006). La grille ARBA propose une évaluation quantitative la plus précise et exhaustive possible du besoin d'aide de la personne handicapée. Son objectif est d'une part de définir l'encadrement requis en partant des besoins de la personne en fonction de l'environnement dans lequel elle évolue, et d'autre part de saisir tous les besoins d'assistance liés à tous les types possibles de handicap (handicaps physiques, mentaux, psychiques).

La grille ARBA-Atelier, que nous étudions dans ce présent rapport est destinée aux ateliers protégés accueillant des personnes handicapées. Elle vise l'évaluation de tous les usagers au bénéfice d'une occupation à temps complet ou partiel au sein de ces structures. Elle ne s'adresse pas aux institutions de type résidentiel (foyer, EMS, homes) ou aux centres pour personnes mineures. Pour des raisons de commodité de langage nous l'appellerons ARBA (et non ARBA-Atelier).

Lors d'une première évaluation pilote d'ARBA sur un échantillon partiel (Pomini & Fernandez, 2006), les premières constatations effectuées révélaient quelques défauts importants de l'instrument. La section vaudoise de la Fédération des Ateliers pour personnes handicapées (INSOS-FAH-VD) et le

Service de Prévoyance et d'Aides Sociales (SPAS) du Département de la Santé et de l'Action Sociale du canton de Vaud (DSAS) ont alors mandaté l'Unité de réhabilitation du Service de psychiatrie communautaire (PCO, DP-CHUV) pour procéder à l'analyse psychométrique approfondie de la grille ARBA ainsi qu'à la validation d'une nouvelle version qui en corrigerait les défauts. Les partenaires de ce mandat se sont arrêtés sur deux buts : (1) développer une grille ARBA révisée prenant en considération les défauts identifiés lors de l'évaluation de la grille initiale et aboutissant à un outil plus simple d'utilisation ; (2) procéder à une validation empirique de l'outil sur un large échantillon d'ateliers et de moniteurs, échantillon représentatif de l'ensemble des ateliers vaudois concernés par l'utilisation de cet instrument.

2. METHODE

2.1 Plan de l'étude

Nous avons subdivisé le projet en trois phases distinctes.

- 1) L'analyse statistique et critique des données recueillies avec les grilles ARBA originales utilisées en 2006 dans les ateliers protégés du canton de Vaud (16 institutions, 2370 grilles).
- 2) L'élaboration d'une grille ARBA révisée (ARBA-R) avec un test pilote sur un échantillon plus restreint d'usagers et de moniteurs (5 institutions, 235 grilles).
- 3) La finalisation de la grille ARBA révisée en une grille ARBA-2, et son test à grande échelle sur un large échantillon d'ateliers protégés du canton de Vaud (15 institutions, 1864 grilles).

2.2 Méthode d'analyse critique des grilles

Pour chacune des trois phases de l'étude, la même procédure d'analyse critique et psychométrique a été appliquée à la version de la grille qui était testée.

Validité de contenu :

Nous avons procédé à un examen détaillé du contenu des items (formulations employées) ainsi que du système de cotation utilisé. Nous avons examiné les définitions de chacun des items en nous efforçant de clarifier ou de préciser la nature des concepts mesurés et en recherchant les formulations les plus claires et synthétiques possibles. Nous avons également cherché à mettre en évidence des biais ou failles logiques dans le fonctionnement des grilles, en nous penchant sur le système de cotation, sur la définition et l'empan des échelles de mesure employées, ainsi que sur les procédures de calcul des scores.

Des séances de concertation ont été tenues avec les moniteurs et responsables d'ateliers en vue d'obtenir leurs avis sur les défauts que nous avons repérés et les propositions de corrections que nous envisagions.

Analyse d'items :

Nous avons étudié la distribution des scores ainsi que les moyennes de chacun des items, dans le but notamment d'identifier ceux rarement sollicités. Nous avons aussi procédé à l'examen de la structure interne de la grille à l'aide d'analyses factorielles (analyses en composantes principales). Nous

souhaitions par cette méthode mettre en évidence les dimensions principales couvertes par l'instrument ainsi que les items les plus utiles dans l'évaluation de ces dimensions, en précisant entre autres la contribution spécifique de chaque item aux composantes retenues. C'est le croisement de toutes ces informations qui a permis au fil des analyses de procéder à une sélection la plus pertinente possible des items de la grille.

2.3 Recueil et gestion des données

Les moniteurs impliqués ont été formés à l'utilisation des outils et à la procédure voulue par le protocole de recherche, ceci afin de garantir une évaluation la plus précise et la plus neutre possible. Les personnels impliqués ont été informés de la possibilité d'obtenir, sur simple demande, une assistance ou des précisions supplémentaires par téléphone et par e-mail.

Toutes les données recueillies ont été soumises au secret professionnel des moniteurs et des chercheurs impliqués dans cette étude. Les grilles ARBA ont été compilées dans des bases de données anonymisées ne faisant aucune mention de l'identité des usagers ou des moniteurs. Ces fichiers sont placés sous la direction du responsable de l'étude. Ces bases de données sont la propriété du Service de psychiatrie communautaire, de l'INSOS-FAH-VD et du SPAS, partenaires dans cette recherche. Chaque institution est par ailleurs propriétaire de la base de données qui la concerne spécifiquement.

3. EVALUATION CRITIQUE DE LA GRILLE ARBA (VERSION ORIGINALE)

3.1 Présentation de la grille

La grille ARBA est composée de vingt-huit items répartis sur quatre dimensions (DSE-DGAS, 2006). La première dimension nommée « *A – Autonomie dans le travail* » évalue le niveau d'autonomie au poste de travail à l'aide d'un seul item se présentant sous la forme d'une échelle de Guttman allant de 1 à 12 points. Les trois autres dimensions (« *B – Autonomie pratique* », « *C – Rapport à soi / Rapport aux autres* » & « *D – Environnement de la place de travail et formation* ») évaluent quant à elles la fréquence des interventions des moniteurs d'atelier auprès des usagers avec des échelles de Likert à cinq points (scores de 0 à 1 par incréments de 0.25). Le score ARBA total est ensuite obtenu en sommant les scores totaux des quatre dimensions. Le système de calcul Excel limite à quatre points les totaux obtenus aux dimensions B, C & D même si la somme des scores inscrits par les moniteurs peut se révéler supérieure.

3.2 Analyse des contenus et du système de cotation

En préambule, nous désirons rappeler ici une des remarques formulées dans le rapport réalisé lors de la première étude pilote d'évaluation d'ARBA (Pomini & Fernandez, 2006). Il y était mentionné que les items d'ARBA évaluent l'aide fournie plutôt que le besoin d'aide, lequel est beaucoup plus subjectif, dépend des impressions des moniteurs, et requiert des formulations clairement orientées dans cette direction. Cette option est sans doute la bonne car elle permet d'éviter les discussions autour de la notion de besoin : ce n'est pas le besoin d'aide mais le temps effectif passé par les moniteurs à encadrer les usagers dans leur travail qui fait l'objet de l'évaluation. Dans ce sens l'usage du terme « besoin d'aide » dans le titre de la grille se révèle discutable, car il complexifie et risque de rendre confuse la tâche d'évaluation demandée aux moniteurs.

Le premier item a un statut particulier. D'une part il est coté sur une échelle différente, d'autre part il évalue davantage l'autonomie et les compétences/déficits des usagers que l'aide dispensée. Sa formulation conduit à une mesure indirecte de l'encadrement, inférant l'aide nécessaire à partir du niveau de compétences observé. Cette façon d'opérer nous paraît peu adéquate : certains usagers très compétents peuvent se voir confier des tâches complexes nécessitant un encadrement plus important. A l'opposé, des usagers avec d'importantes limitations ne nécessiteront pas de facto un encadrement plus soutenu lors de la réalisation de tâches simples. De plus, il est délicat d'amalgamer

deux types de mesures au sein d'un seul indice puisque le score total devient alors difficilement interprétable.

L'utilisation de deux systèmes de cotation répondant à des logiques différentes (A de 1 à 12 et B, C & D de 0 à 1) pose la question du poids accordé à chacune des quatre dimensions au sein du score total ARBA. Certaines dimensions méritent-elles vraiment d'avoir plus de poids dans le score ? Selon quels critères cette décision peut-elle être prise ? La lecture de la documentation sur ARBA ne fournit pas des réponses précises à ces questions. Nous n'avons donc pas été en mesure de mieux comprendre la démarche adoptée, de la légitimer voire de l'accepter, et sommes restés sur notre impression initiale d'arbitraire scientifiquement peu défendable.

La cotation des items B, C et D prend en compte la fréquence mais pas la durée des interventions. Même si un système d'évaluation verbale de la fréquence des aides fournies paraît de prime abord plus élégant et plus simple d'emploi, il ne nous paraît pas possible de faire l'économie de la durée des interventions. Des interventions relativement longues mais fréquentes contribuent au score total de façon similaire à des interventions beaucoup plus courtes mais dispensées fréquemment. Or elles pourraient bien ne pas être cotées la même chose par tous les moniteurs, certains préférant suivre à la lettre les instructions en produisant des cotations faibles (faible fréquence), d'autres n'hésitant pas à augmenter le score pour parvenir à une sorte de moyenne plus élevée (reflétant selon eux davantage la réalité). Le débat quant à l'utilisation des fréquences ou des durées dans la mesure des comportements est un vieux débat de la psychologie et de la psychométrie. Dans le cas qui nous occupe cependant, les problèmes d'interprétation liés à la mesure des fréquences nous ont paru plus importants que les bénéfices de la simplicité, notamment par rapport au fait que la conduite à tenir dans les cas complexes n'avait pas été clairement opérationnalisée.

Nous nous sommes rendus compte à l'usage des grilles Excel que la dénomination des items ne correspondait pas toujours exactement aux définitions. Cela entraîne un risque de biais systématiques dans la cotation. Par exemple, l'item « *Fréquence des absences au travail* » a pour définition « *Mesure de l'aide motivationnelle à se rendre au travail et/ou lors de cas d'absentéisme (téléphones, entretiens, etc.)* ». A l'usage, il y a de fortes chances que les évaluateurs ne reviennent pas aux définitions inscrites dans le manuel et évaluent des contenus différents de ceux réellement déterminés par les auteurs de la grille. La validité des cotations peut alors être mise en doute.

La concertation avec un panel de moniteurs travaillant dans différentes institutions du canton nous a permis de relever que la grille ARBA ne rendait pas compte de certaines composantes importantes du travail d'encadrement fourni dans les institutions, notamment la suppléance lors d'absences répétées et les démarches administratives effectuées par les moniteurs. Par ailleurs, un extrait du rapport d'activité de la FAH souligne que, selon les audits effectués, la rubrique « *autonomie au travail* » est celle qui pose le plus de difficultés d'évaluation aux moniteurs (extrait du rapport d'activité de la FAH, 2006).

La limitation à quatre points des totaux des dimensions B, C et D traduit l'intention de réserver à ces domaines des contributions plus équilibrées lors du calcul du score ARBA total car le nombre d'items composant chacune de ces trois sous-dimensions diffère (DSE-DGAS, 2006). Cette façon de faire nous apparaît cependant difficilement défendable d'un point de vue psychométrique et logique, car elle peut réduire arbitrairement un score d'encadrement élevé. L'échelle n'est plus alors en mesure de traduire l'aide effectivement fournie par les moniteurs.

A l'issue de l'analyse détaillée du contenu des items et du système de cotation, les conclusions issues du premier rapport d'évaluation ARBA (Pomini & Fernandez, 2006) se trouvaient confirmées. Elles légitiment déjà pleinement les buts du présent mandat : améliorer la définition des items et la méthode de cotation pour produire des scores plus aisément interprétables. L'analyse d'items nous a permis ensuite d'examiner de plus près la sélection des items de la grille.

3.3 Analyse des items

Parmi les 2370 usagers évalués provenant des seize ateliers protégés du canton, 63.8% étaient des hommes (n = 1487) et 36.2% des femmes (n = 845). L'âge moyen des usagers était de 43.4 ans (sd = 13.53). La distribution des âges dans l'échantillon est assez uniforme dans son ensemble, chaque groupe d'âge étant représenté plus ou moins à part égale des autres, en dehors des jeunes usagers ou des personnes aux alentours de l'âge de la retraite (v. figure 1).

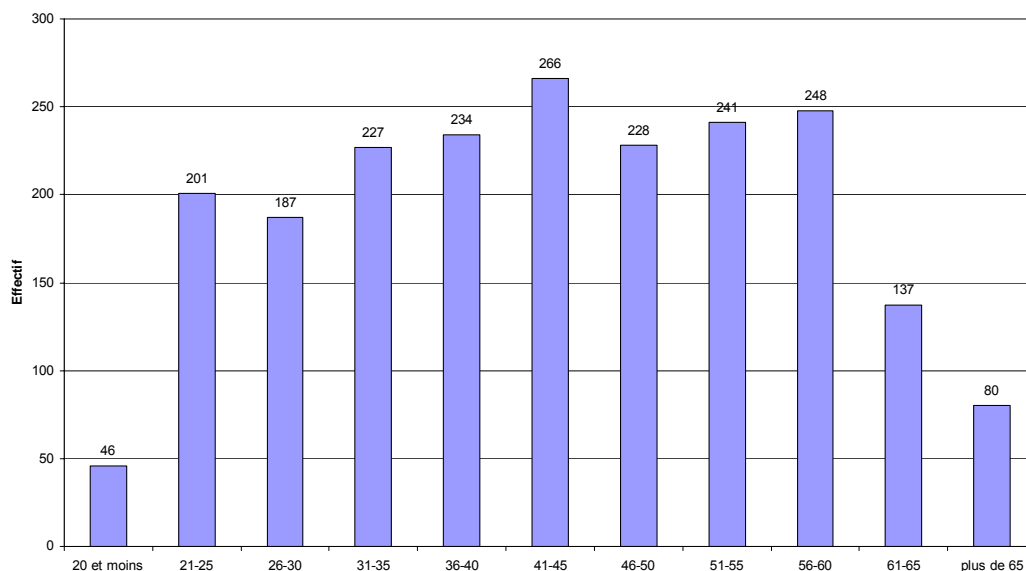


Figure 1 : distribution de l'âge des usagers

Analyse des distributions et des moyennes

La moyenne des totaux obtenus aux grilles ARBA pour l'ensemble des usagers est de 13.24 points (sd = 4.72) et la médiane de 12.75. Les valeurs extrêmes sont peu représentées et une grande partie de l'effectif se trouve proche de la moyenne, ce que nous indique la faible valeur de l'écart-type. Notons encore que 50% des scores ARBA sont situés entre les valeurs 10 et 16.5 (v. figure 2).

Nous avons regardé s'il y avait un lien entre le niveau d'encadrement et le sexe ou l'âge des usagers. Nous avons observé une différence de score statistiquement significative entre hommes et femmes ($t(2330) = -2.483, p < .05$). Le score ARBA total moyen des femmes est plus élevé de 0.51 point et se situe à 13.59. En revanche, nous n'observons pas de corrélation entre l'âge et le score total ($r = 0.005, ns$).

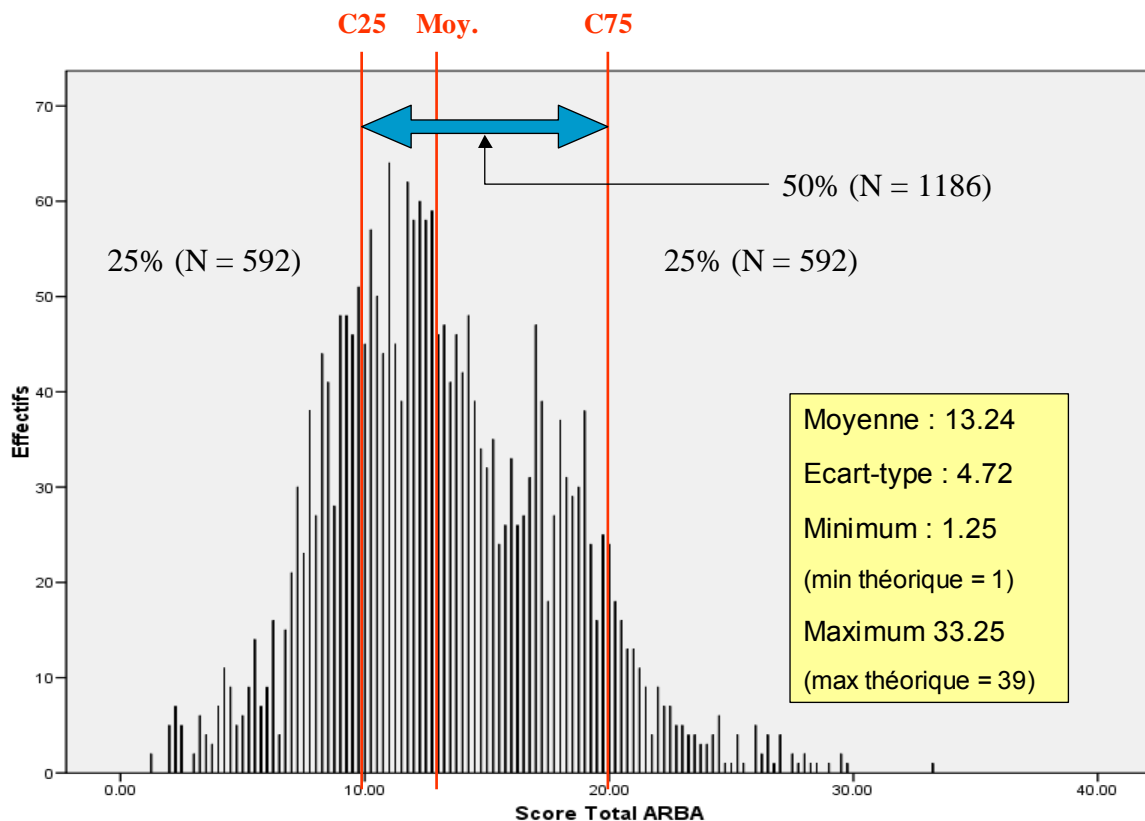


Figure 2 : distribution des scores totaux ARBA

Dans le tableau 1 figurent les moyennes des items classées par ordre décroissant. Le score à l'item d'autonomie au travail (A1) apparaît évidemment nettement supérieur aux autres puisque le mode d'évaluation utilisé est différent (respectivement une échelle de Guttman de 1 à 12 points et plusieurs échelles de Likert de 0 à 1 point). En tant que telle, cette différence n'est bien sûr pas

interprétable. De ce fait, et parce que la dimension A ne répond ni à la même logique ni au même type de contenu que les trois autres dimensions, nous avons décidé de l'exclure des analyses ultérieures portant sur les items de la grille.

Les thématiques où la demande d'aide est la plus forte sont celles relatives à la fragilité psychique (items « *fragilité, dévalorisation* », « *risque suicidaire / angoisse envahissante* ») et aux compétences liées au travail (« *compréhension des consignes* », « *stimulation au travail* », « *exécution de tâches multiples* », « *projet d'apprentissage interne ou soutien à une formation externe* »). Les items aux plus faibles scores concernent les liens avec l'extérieur de l'institution (« *fugue* » et « *appui en stage externe* ») où la plupart des usagers (respectivement 94.0% et 95.9%) obtient une note de 0.

Items	Moyennes
A1 - Autonomie dans le travail	8.91
C3 - Fragilité, dévalorisation	.43
D2 - Compréhension des consignes	.40
D6 - Stimulation au travail	.38
D1 - Exécution de tâches multiples (changements + instructions)	.37
C2 - Communication (verbale / non verbale)	.34
D4 - Projet d'apprentissage interne ou soutien à une formation externe	.32
C7 - Risque suicidaire /angoisse envahissante	.26
C5 - Gestion des conflits / comportement agressif	.25
D7 - Liens avec le réseau	.23
C4 - Fréquences des absences au travail	.18
C11 - Graves troubles neurologiques (délirants / mnésiques / épilepsie, etc.)	.12
B5 - Changement de vêtements	.11
C8 - Compulsion / addiction	.11
B1 - Utilisation des WC, hygiène	.10
D3 - Rapport au danger	.10
B3 - Gestion des quantités et de l'équilibre nourriture / boisson	.08
B7 - Déplacement à l'intérieur	.08
C12 - Surveillance santé / prise médicaments	.08
C1 - Difficultés visuelles	.07
B4 - Comportements à table	.06
B6 - Changements de position (se lever, s'asseoir, transferts)	.06
B2 - Prise de nourriture et de boisson	.05
B8 - Utilisation des transports publics / déplacements à l'extérieur	.04
C6 - Destruction de matériel	.04
C9 - Troubles du comportement sexuel	.03
C10 - Fugue	.03
D5 - Appui en stage externe	.02

Abréviations des quatre domaines d'ARBA : A = Autonomie au travail ; B = Autonomie pratique ; C = Rapport à soi / aux autres ; D = Environnement de la place de travail & formation

Tableau 1 : moyennes pour chaque item (rangées par ordre décroissant)

Si l'on considère les moyennes des totaux pour chacun des trois domaines, on voit que le domaine B (0.57) est celui qui requiert le moins d'encadrement, les deux autres domaines présentant des moyennes sensiblement proches (1.93 et 1.82 pour C et D).

Pour certains items, une très large majorité des usagers évalués obtient la note 0. Par exemple, pour l’item D5 (« Appui en stage externe »), 2274 usagers obtiennent un score de zéro (aide fournie très rare/absente) alors que 96 usagers seulement obtiennent des notes différentes de zéro (0.25, 0.5 ou 1). Le tableau 2 présente les items les plus rarement utilisés. On y indique le pourcentage de cas où la cotation est égale à zéro (aucun encadrement fourni), donc le pourcentage de cas où l’item s’avère inutile car non sollicité.

<i>Items</i>	<i>pourcentage de cas où l’item est coté 0</i>
D5 Appui en stage externe	95%
B2 Prise de nourriture et de boisson	
B6 Changements de position (se lever, s’asseoir, transferts)	
B8 Utilisation des transports publics / déplacements à l’extérieur	90%
C6 Destruction de matériel	
C9 Troubles du comportement sexuel	
C10 Fugue	
B4 Comportements à table	
B7 Déplacements à l’intérieur	85%
C1 Difficultés visuelles	
B1 Utilisation des WC, hygiène	
B3 Gestion des quantités et de l’équilibre nourriture / boisson	
B5 Changement de vêtements	80%
C8 Compulsion	
C12 Surveillance santé / prise médicaments	

Tableau 2 : items rarement sollicités

L’observation du tableau 2 montre qu’au total sept items obtiennent des scores nuls dans 90% des cas (quinze dans 80% des cas). On peut légitimement se poser la question de l’utilité de ces items, de la pertinence à les éliminer ou à les condenser, car l’information totale qu’ils véhiculent est dans l’ensemble plutôt réduite étant donné qu’ils ne sont que très rarement employés.

Structure de la grille

L’analyse factorielle est une analyse statistique qui permet de réduire l’ensemble des informations contenues par tous les items en une série de dimensions synthétiques, moins nombreuses. L’intérêt de l’analyse factorielle est de pouvoir résumer en quelques composantes clés l’information véhiculée par un nombre important d’items. Elle permet de faire émerger des facteurs ou dimensions, c’est-à-dire des groupes d’items corrélant fortement entre eux et faiblement avec les autres items. Dans le

cas qui nous occupe, une analyse factorielle des vingt-sept items peut donner théoriquement de un à vingt-sept facteurs. Une solution à un seul facteur synthétise l'ensemble des items sur une seule et unique dimension. Retenir vingt-sept facteurs serait évidemment inutile car on ne ferait que reproduire le découpage initial. Plus une corrélation entre un item et un facteur est grande, plus l'item en question est représentatif du facteur. Le choix du terme employé pour nommer le facteur dépend ainsi des items qui corréleront le plus avec lui ; ils en sont en quelque sorte les meilleurs représentants.

L'objectif d'une analyse factorielle est de rechercher le meilleur nombre possible de facteurs en veillant à ne pas perdre trop d'information et en proposant une structure la plus intelligible possible. Il existe plusieurs méthodes permettant de calculer le nombre de facteurs que l'on peut extraire d'un ensemble d'items. Sans entrer dans le détail, nous pouvons dire ici qu'elles s'appuient sur l'idée général qu'un facteur retenu doit comprendre davantage d'information qu'une variable prise isolément. Deux méthodes sont habituellement appliquées lors d'une réflexion autour des résultats d'une analyse factorielle : (1) la méthode de Kaiser qui indique le nombre des facteurs qui véhiculent plus d'information qu'une variable isolée, (2) la méthode du Scree-test de Cattell qui nous indique le seuil au delà duquel l'accumulation de l'information n'est plus suffisante pour légitimer l'inclusion d'un nouveau facteur. Les méthodes convergent souvent, mais parfois elles donnent des résultats différents, la méthode de Cattell pouvant indiquer un nombre moins important de facteurs à retenir. En tous les cas, le choix du nombre de facteurs comprend toujours au final une part de subjectivité, car aucun critère n'est absolu. De plus, une fois le nombre de facteurs choisis, le chercheur doit aussi être en mesure des les interpréter et de leur donner un nom. Tout facteur retenu doit avoir du sens aux yeux des concepteurs et des utilisateurs de l'instrument.

Pour la grille ARBA nous avons examiné en détail deux solutions (v. figure 3) : 1) une solution à trois facteurs (issue de l'application du Scree-test de Cattell), 2) une solution à sept facteurs (retenue en respectant la méthode de Kaiser). La figure 3 illustre chaque facteur par un rectangle dans lequel sont inscrits le titre donné au facteur (en gras) ainsi que les principaux items qui le composent, c'est-à-dire qui corréleront le plus avec lui.

La solution à trois facteurs permet d'expliquer 34.13% de la variance. Cela signifie qu'en réduisant la grille à trois dimensions principales, on perd environ 66% de l'information véhiculée par l'ensemble des vingt-sept items qui la composent. Bien que le découpage théorique a priori de la grille propose également trois catégories d'items, les dimensions issues de l'analyse factorielle exploratoire ne correspondent pas à ce dernier. En effet, en dehors du premier facteur qui regroupe les items de la dimension B (autonomie pratique), les autres items ne se répartissent pas dans les deux facteurs restants comme le découpage initial de la grille le prévoit. Pour cela il aurait fallu que les items C et les items D soient réunis chacun dans un facteur différent, ce qui n'est pas le cas (cf. partie gauche de la figure 3).

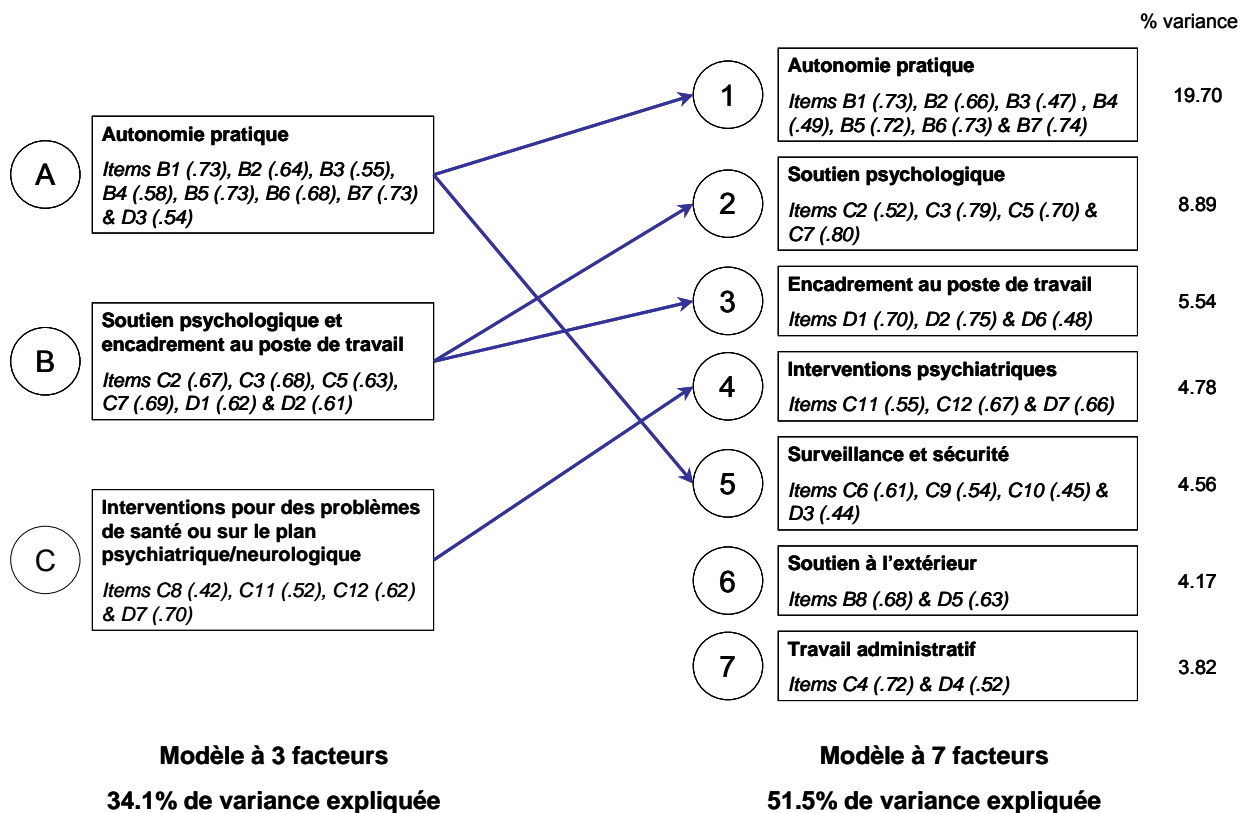


Figure 3 : analyse factorielle exploratoire avec rotation varimax (N = 2370) ; répartition des items qui corrélerent le mieux avec les facteurs

Le modèle à sept facteurs (critère de Kaiser) paraît meilleur que le précédent, puisqu'il rend compte de 51.5% de la variance, soit 17% de variance en plus (v. figure 3). Le tableau 3 reproduit toutes les corrélations observées entre l'ensemble des items et les sept facteurs retenus. On y retrouve la concentration des items B dans le premier facteur, et la dispersion des items C et D dans les six autres.

Le modèle à sept facteurs permet de conserver environ 50% de l'information véhiculée par les vingt-sept items d'ARBA, tout en gardant une structure générale relativement simple. Bien que meilleure que la précédente, cette solution reste quand même relativement moyenne. On vise en général des solutions permettant d'expliquer au moins 65-70% de la variance initiale. Néanmoins cette solution à sept facteurs nous permet de mettre à jour des dimensions tout à fait pertinentes pour la description générale du travail d'encadrement effectué par les moniteurs, comme le soutien psychologique et l'encadrement au poste de travail (qui sont confondues dans la solution à trois facteurs), le soutien à

l'extérieur ou le travail administratif (qui n'apparaissent pas dans le modèle à trois facteurs). Ainsi, bien que plus complexe, la structure à sept dimensions semble au final plus intéressante que celle à trois catégories. Il devient dès lors envisageable de réviser la grille en respectant mieux sa structuration « naturelle » empiriquement vérifiable via l'analyse factorielle.

Items	Facteurs						
	I	II	III	IV	V	VI	VII
Autonomie Pratique							
Utilisation des WC, hygiène	.734	.076	.136	.092	.109	-.048	-.023
Se nourrir, boire	.660	-.022	-.012	.070	.067	.173	.104
Gérer la quantité et l'équilibre de nourriture / boisson	.471	.168	-.005	.084	.389	.206	.003
Comportement à table	.489	.185	-.043	.049	.480	.141	-.035
Changer de vêtements	.722	.075	.142	.036	.151	-.007	-.031
Se lever / s'asseoir / transferts	.731	-.011	.059	.067	-.022	-.049	.006
Se déplacer à l'intérieur	.736	.027	.145	.203	.018	-.036	-.124
Utilisation des transports publics / déplacements	.196	-.024	.079	-.031	.100	.676	.012
Rapport à soi/autres							
Difficultés visuelles	.310	-.072	.235	.156	-.029	-.178	-.379
Communication (verbale / non verbale)	.282	.517	.466	-.005	-.001	.120	-.072
Fragilité, dévalorisation	.027	.790	.189	.182	-.026	.018	.104
Fréquences des absences au travail	-.105	-.020	.282	.254	.026	-.066	.715
Gestion des conflits / comportement agressif	.013	.703	.083	-.047	.279	.029	-.181
Destruction de matériel	.097	-.011	.111	.031	.611	-.170	-.020
Risque suicidaire / Angoisse envahissante	.026	.800	.145	.164	-.013	-.014	.104
Compulsion / addiction	.094	.272	.089	.343	.224	-.236	.092
Troubles du comportement sexuel	-.038	.086	.035	-.014	.538	.092	-.020
Fugue	.161	-.023	.110	.098	.454	.109	.290
Graves troubles neurologiques	.152	.072	.117	.553	-.048	.081	.046
Surveillance santé / prise médicaments	.170	.011	-.168	.672	-.049	.105	-.016
Environnement							
Exécution de tâches multiples	.087	.216	.669	-.146	.136	.082	.093
Compréhension des consignes	.182	.202	.750	.031	.070	.067	-.074
Rapport au danger	.390	-.009	.170	.144	.441	.079	-.145
Projet d'apprentissage ou soutien à une formation	-.130	-.051	.346	.374	.104	.178	-.518
Appui en stage externe	-.073	.064	.034	.158	.006	.631	-.036
Stimulation au travail	.181	.164	.480	.269	.263	-.112	.178
Liens avec le réseau	.081	.155	-.004	.662	.261	-.006	-.010

Tableau 3 : solution à sept facteurs après rotation varimax (N=2370) ; les saturations supérieures à /0.50/ sont en gras

3.4 Synthèse et discussion

Nous avons vu que la grille ARBA initiale répondait à deux logiques distinctes en mesurant à la fois l'autonomie de l'utilisateur (item 1) et l'aide dispensée par le moniteur (items 2-28). La sommation de ces deux types d'items en un seul score nous est apparue peu défendable car elle rend difficile l'interprétation de ce dernier. Nous avons décidé ainsi d'exclure de la nouvelle grille la première mesure et de nous concentrer exclusivement sur l'aide effective dispensée par les moniteurs. Nous avons également opéré ce choix après avoir vérifié que cet item n'apportait pas non plus d'information supplémentaire, comme nous l'ont indiqué les résultats d'une analyse factorielle complémentaire non rapportée ici. Celle-ci, effectuée sur l'ensemble des vingt-huit items montrait que l'inclusion du premier item ne changeait pas les résultats de façon substantielle, cet item n'étant par exemple pas du tout isolé dans un facteur spécifique, comme on aurait peut-être pu l'imaginer au départ.

Suite à un travail de concertation avec des moniteurs employés dans plusieurs institutions du canton, nous avons opté pour une grille composée de dix items, rédigés en regroupant le contenu des items fortement corrélés de la grille initiale. Il est en effet apparu que les dimensions mises à jour par l'analyse factorielle pouvaient correspondre aux items de la nouvelle grille, sans pour cela faire souffrir la grille d'une perte trop substantielle d'information par rapport à l'objectif poursuivi (mesure du degré d'encadrement). La formulation des nouveaux items a donc été réalisée à partir des facteurs identifiés dans la structure latente de la première grille. Nous avons délaissé le système où chaque item ne traite que d'une prestation/problématique très spécifique (p. ex : « *changer de vêtements* »), pour rédiger des items à la définition plus large qui puissent directement correspondre aux facteurs identifiés empiriquement (p. ex « aide pratique pour les actes du quotidien »).

Pour répondre aux limitations du système de cotation basé sur l'évaluation verbale de la fréquence des interventions, nous avons choisi la notion de durée cumulée, car elle prend simultanément en compte les caractéristiques de fréquence et de durée des interventions des moniteurs.

Telles ont été les trois modifications majeures de la grille, pour parvenir à un instrument plus économe, plus simple et plus valide. L'idée a été de tendre vers une version révisée qui couvre les mêmes dimensions que la grille originale et qui soit structuré selon un modèle empiriquement validé, tout en s'appuyant sur un système de cotation laissant si possible une part moindre à la subjectivité des moniteurs.

4. EVALUATION PILOTE D'ARBA-R

4.1 Présentation de la grille

La grille révisée comporte dix items (v. tableau 4). Elle reprend le contenu des sept facteurs identifiés par nos analyses factorielles exploratoires, auxquels se sont ajoutés quelques items supplémentaires considérés par notre pool d'experts comme importants pour la validité de contenu.

Items	
1.	Aide pratique pour les actes du quotidien
2.	Soutien psychologique et médical
3.	Instruction et démonstration de l'activité
4.	Co-réalisation de l'activité
5.	Surveillance et sécurité
6.	Vérification de l'activité
7.	Interventions lors de comportements
8.	Coordination avec le réseau général et
9.	Interventions motivationnelles et
10.	Travail administratif

Tableau 4 : les dix items d'ARBA-R

Le premier item a par exemple été rédigé en regroupant le contenu des items rattachés à la thématique de l'aide pratique pour les actes du quotidien. Sa définition exhaustive est la suivante : « Aide apportée pour s'orienter dans les lieux (WC, poste de travail, cafétéria, vestiaire, ...), pour se lever, se déplacer, s'asseoir, se nourrir, s'habiller et maintenir une hygiène suffisante ». Cette définition correspond aux items qui représentent le mieux le premier facteur (utilisation des WC, hygiène / se nourrir, boire / changer de vêtements / se lever , s'asseoir, transferts / se déplacer à l'intérieur).

La dimension d'encadrement au poste de travail a été subdivisée en trois items (« instructions et démonstration de l'activité », « co-réalisation de l'activité », « vérification de l'activité ») afin de permettre aux moniteurs de préciser de façon plus détaillée la nature de ces prestations. Puisque les contenus évalués par ces trois items sont particulièrement proches voire liés (ils constituent les éléments d'une séquence d'interventions), nous pouvons nous attendre à ce qu'ils entretiennent des inter-corrélations plus fortes qu'avec les autres items. Les usagers bénéficiant de prestations importantes pour l'une de ces trois composantes auront vraisemblablement des scores élevés pour les deux autres items.

Le système de cotation adopté (v. figure 4) évalue l'importance de l'aide dispensée à chaque usager dans les dix domaines définis. Il met en parallèle un système adjectival et un système d'intervalles de temps. A chaque adjectif correspond une durée d'encadrement. Les moniteurs peuvent ainsi s'appuyer sur des intervalles de durée pour qualifier de façon plus objective l'importance de l'encadrement requis. Cette alternative cherche à garantir l'adoption d'un langage commun entre évaluateurs, qui minimise les biais liés à des styles particuliers de cotation.

Un autre principe ayant guidé la mise au point du système de cotation est de pouvoir offrir des scores permettant de retranscrire de la façon la plus fidèle possible l'aide apportée. Il nous a ainsi paru nécessaire de veiller à ce que la somme des durées rapportées par six items cotés à 1 corresponde à celle retranscrite par cinq items nuls et un item coté à 6. Nous avons cherché à lier le score ARBA de façon directe à la durée de l'aide dispensée. L'idée était qu'un score total de 8 points représenterait un encadrement plus ou moins double à celui évalué par un total de 4 points. Ainsi, sans disposer formellement d'une échelle quantitative, nous avons souhaité que les scores ARBA s'en approchent, pour en faciliter le décodage et l'interprétation. Les dix dimensions sont également représentées de façon équitable sans recours à des pondérations ni des limitations d'aucune sorte.

Besoin d'encadrement :	Durée pour une journée-type (quel que soit le taux d'occupation)
Absent	0'
Léger	1'-20'
Modéré	20'-40'
Assez important	40'-1h00
Important	1h00-1h20
Très important	1h20-1h40
Extrêmement important	1h40-2h00

Figure 4 : système de cotation choisi pour la grille ARBA-R

Le besoin d'encadrement d'un usager est déterminé par la durée totale d'encadrement dispensé en moyenne par le moniteur sur une journée où l'utilisateur est présent à l'atelier. L'évaluation se fait sur la base d'une journée-type, mais le score ARBA pour la journée-type (qui correspond à la somme des dix items) peut ensuite être transformé en fonction des jours de présence et donner lieu à un score ARBA hebdomadaire, mensuel ou annuel. Les besoins globaux d'encadrement d'un atelier sur une période donnée correspondent alors à la somme des scores ARBA de tous les usagers qui le fréquentent durant cette période.

4.2 Analyse des contenus et du système de cotation

La grille ARBA-R a été testée début 2008 dans les cinq institutions suivantes : Brico-CES-services, les Eglantines, la Fondation de Vernand, la Manufacture et POLYVAL. 235 usagers auront à cette occasion été évalués avec ARBA-R par les moniteurs de ces cinq ateliers partenaires.

Cette phase à vocation pilote aura constitué une première occasion de tester sur le terrain les modifications apportées à la grille ARBA. Cette étape aura également permis de mettre à l'épreuve les choix opérés dans la définition des concepts mesurés (évaluation axée sur l'aide apportée par le moniteur plutôt qu'autonomie de l'utilisateur).

Le contenu des items a généralement été très bien accepté par les moniteurs. Ces derniers ont notamment déclaré que les contenus proposés dans ARBA-R décrivaient de façon fidèle les prestations offertes aux usagers dans les différentes institutions. Cependant cette phase pilote a permis également de voir que certaines définitions pouvaient être affinées, et les items mieux différenciés encore les uns des autres. La rédaction des items « *instruction et démonstration de l'activité* » et « *co-réalisation de l'activité* » nous a paru par exemple devoir être légèrement modifiée afin de lever toute ambiguïté entre la démonstration des méthodes et techniques devant l'utilisateur et la co-réalisation du travail avec ce dernier. De même l'item « *vérification de l'activité* » méritait aussi une nouvelle dénomination pour éviter toute confusion avec l'item « *surveillance et sécurité* ».

Il fut néanmoins relevé que certains contenus liés aux évolutions induites par la 5^{ème} révision de la loi sur l'assurance invalidité méritaient d'être mieux pris en compte par la grille. Il s'agit principalement des prestations liées à l'élaboration de projets pour ou avec l'utilisateur, et à l'organisation d'entraînements ou de formations personnalisées dans une optique de transition et/ou de réorientation. La grille ARBA-R retranscrivant mal ces éléments, une adaptation s'imposait.

Il est également apparu que les moniteurs ne partageaient pas toujours les mêmes définitions opérationnelles des adjectifs mis en correspondance avec les durées d'encadrement. Ainsi un besoin d'encadrement estimé comme léger pouvait coïncider pour certains à quelques minutes dispensées sur la journée, alors que pour d'autres un encadrement léger équivalait au moins à vingt ou trente minutes d'aide quotidienne. Tout cela dépendait évidemment des contextes de travail, des usagers et du type d'aide fournie. En somme la double référence à des adjectifs et à des durées a entraîné un débat entre moniteurs sur la pertinence des correspondances proposées entre les durées d'encadrement et les adjectifs censés les caractériser. Contrairement au souhait initial, cette double référence fut davantage source de confusion que d'éclaircissement. A l'issue de la phase pilote, il était assez clair qu'un choix devait être opéré en faveur d'un système de cotation (adjectival/ordinal vs numérique/quantitatif) au détriment de l'autre.

Enfin, nous avons pu noter que pour un certain nombre de moniteurs l'échelle précisant les intervalles de durée ne faisait pas preuve d'une sensibilité suffisante lorsqu'ils étaient confrontés à l'évaluation d'aides brèves (5, 10 ou 15 minutes). A l'usage, les incréments de 20 minutes ont paru pour l'essentiel trop larges, limitant beaucoup la capacité de discrimination de l'outil. Une réduction des intervalles considérés s'est donc unanimement imposée.

4.3 Analyse des items

Parmi les 235 usagers évalués lors du test pilote d'ARBA-R dans les cinq institutions partenaires, 69.8% étaient des hommes (n = 164), et 30.2% des femmes (n = 71). L'âge moyen des usagers était de 44.5 ans (sd = 12.8). En dehors d'une sous-représentation de la classe d'âge 31-35 ans, voir celle des 21-25 ans, doublée d'une présence plus importante des 61-65 ans, la distribution des âges est très proche de celle observée pour l'évaluation d'ARBA (v. figures 5 et 1). On note aussi une proportion moindre de femmes que dans l'échantillon précédent.

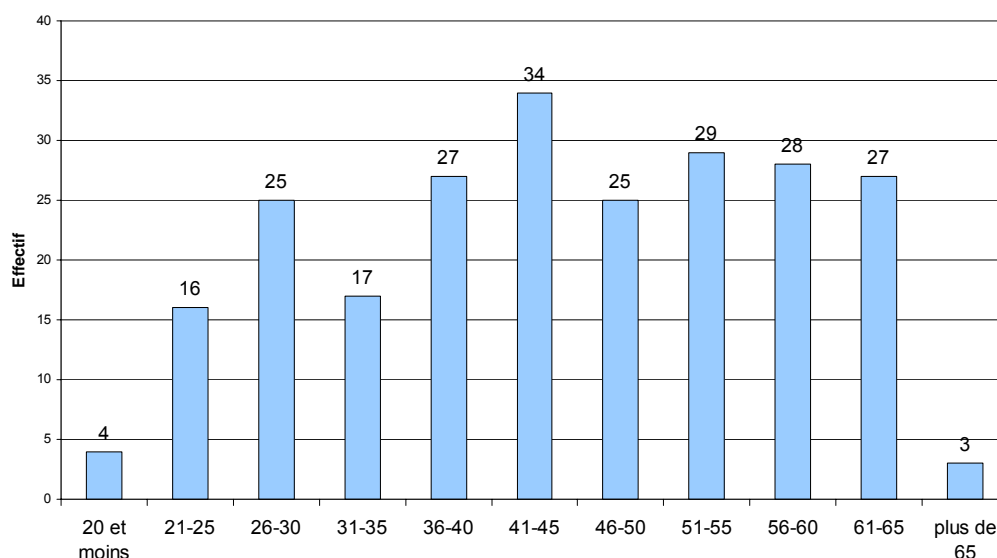


Figure 5 : distribution de l'âge

Analyse des distributions et des moyennes

Le score total moyen obtenu pour l'ensemble des usagers aux grilles ARBA-R est de 12.85 points (sd = 5.68) et la médiane de 12 points. Le 50% des effectifs obtient un score total entre 9 et 16 (v. figure 6). Le tableau 5 reprend les moyennes des scores aux différents items de la grille. L'item obtenant la moyenne la plus élevée est celui touchant à l'instruction et à la démonstration de l'activité. L'item 1 (aide pratique pour les actes du quotidien) est celui qui obtient la moyenne la plus faible. Les interventions des moniteurs concernent principalement l'encadrement au poste de travail : donner des

instructions, faire la démonstration de la tâche à accomplir, vérifier le travail ou encore participer à la co-réalisation de celui-ci. Suivent de près les interventions de type soutien psychologique ou motivationnel qui ont constitué également une partie importante de l'encadrement fourni. Enfin, les items aux scores les plus faibles concernent des interventions qui visaient à pallier aux difficultés d'autonomie pratique. On notera des moyennes néanmoins proches et peu élevées si on considère l'empan total de l'échelle (0-6).

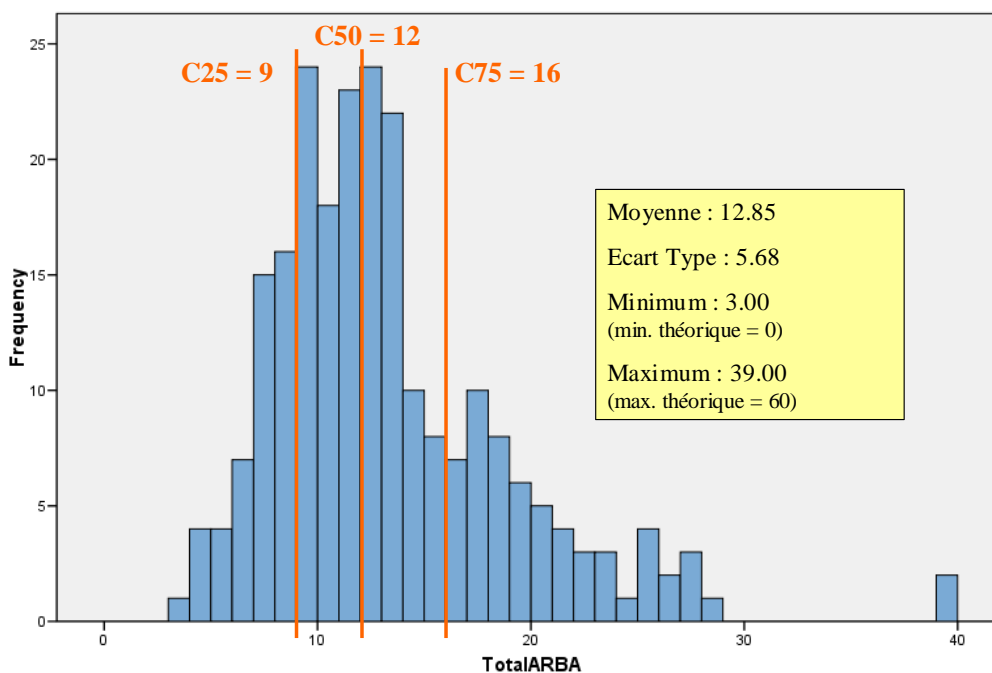


Figure 6 : distribution de l'effectif sur les scores totaux ARBA

Items	Moyennes
3. Instruction et démonstration de l'activité	1.77
6. Vérification de l'activité	1.74
2. Soutien psychologique et médical	1.60
9. Interventions motivationnelles et suppléance	1.54
4. Co-réalisation de l'activité	1.52
5. Surveillance et sécurité	1.14
7. Interventions lors de comportements perturbants	1.11
10. Travail administratif	0.97
8. Coordination avec le réseau général et déplacements	0.90
1. Aide pratique pour les actes du quotidien	0.55

Tableau 5 : moyennes des items (classées par ordre décroissant)

Nous n'avons pas observé de différence significative entre hommes et femmes ($t(233) = .253, p > .05$). Le score ARBA-R total moyen des femmes est de 12.70 points, celui des hommes de 12.91. Nous observons en revanche un lien faible entre l'âge et le score total ($r = -.151, p < .05$). Les usagers plus jeunes ont tendance à obtenir des scores totaux légèrement plus élevés. Cependant, quoique statistiquement significative, cette corrélation reste difficilement interprétable et doit être envisagée avec la prudence de circonstance.

Avec la version révisée de la grille ARBA, il n'existe plus d'item où plus de 80 ou 90% des usagers obtient un score nul (v. tableau 6). Les résultats indiquent que les items d'ARBA sont pratiquement tous sollicités pour décrire les prestations des moniteurs. Le seul item où plus de la moitié des situations est cotée à zéro est le no 1 (aide pratique pour les actes du quotidien). Il est aisé de comprendre ce résultat puisqu'une bonne partie des usagers d'ateliers protégés partenaires dans cette phase de l'étude n'est effectivement pas handicapée sur ce plan. On relève également que la quasi totalité des scores se situe entre 0 et 4. Les valeurs extrêmes de l'échelle ne sont que rarement utilisées. Ces résultats nous indiquent que la valeur maximale choisie pour l'échelle est suffisante, voire trop élevée.

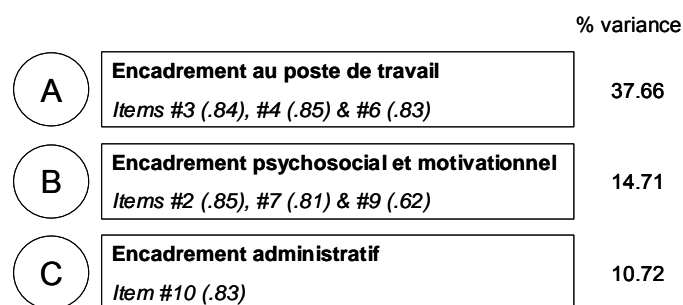
<i>Items</i>	Percentile 95	% des scores compris entre 0 et 4	% de scores nuls
1. Aide pratique pour les actes du quotidien	2.0	99.1	57.9
2. Soutien psychologique et médical	4.0	97.0	10.6
3. Instruction et démonstration de l'activité	4.0	97.9	0.9
4. Co-réalisation de l'activité	3.2	99.1	6.0
5. Surveillance et sécurité	3.0	99.1	20.9
6. Vérification de l'activité	4.0	97.4	0.9
7. Interventions lors de comportements perturbants	4.0	97.4	37.4
8. Coordination avec le réseau général et déplacements	2.0	100.0	28.9
9. Interventions motivationnelles et suppléance	4.0	98.3	12.3
10. Travail administratif	2.0	100.0	11.1

Tableau 6 : répartition des scores aux dix items de la grille révisée

Structure de la grille et homogénéité des scores

L'analyse factorielle conduit à une solution à trois facteurs, qui explique 63.1% de la variance. Les trois dimensions qui se profilent peuvent être définies comme suit : l'encadrement au poste de travail, l'encadrement psychosocial et motivationnel, ainsi que l'encadrement administratif (v. figure 7). Trois items conservent dans ce modèle un statut moins défini : l'item 1 (aide pratique pour les actes du quotidien) a tendance à saturer avec la troisième dimension mais de façon plutôt modérée ($r = 0.56$).

L'item 5 (surveillance et sécurité) a une position hybride car il corrèle de façon égale avec les deux premières dimensions (respectivement $r = 0.48$ et $r = 0.41$) ; l'item 8 enfin (coordination avec le réseau général et déplacements) a également un statut hybride, avec une tendance à saturer sur les deux dernières dimensions ($r = 0.47$ et $r = 0.51$).



Modèle à 3 facteurs - 63.1% de variance expliquée

Figure 7 : synthèse des dimensions obtenues

Ces résultats apparaissent comme cohérents avec le modèle théorique sous-jacent à l'élaboration de la grille. Le modèle issu de l'analyse factorielle exploratoire souligne l'indépendance de la composante administrative par rapport aux aides plus directes dispensées au quotidien. Quant à l'item 1, il occupe résolument une position solitaire et indépendante des autres items, ce à quoi on pouvait s'attendre vu le type de handicap auquel il est assez spécifiquement attaché.

items	1	2	3	4	5	6	7	8	9	10
<i>1. Aide pratique pour les actes du quotidien</i>	1.00									
<i>2. Soutien psychologique et médical</i>	0.20	1.00								
<i>3. Instruction et démonstration de l'activité</i>	0.17	0.17	1.00							
<i>4. Co-réalisation de l'activité</i>	0.28	0.13	0.68	1.00						
<i>5. Surveillance et sécurité</i>	0.29	0.23	0.38	0.42	1.00					
<i>6. Vérification de l'activité</i>	0.14	0.12	0.60	0.61	0.38	1.00				
<i>7. Interventions lors de comportements perturbants</i>	0.24	0.59	0.27	0.35	0.44	0.30	1.00			
<i>8. Coordination avec le réseau et déplacements</i>	0.25	0.36	0.22	0.19	0.37	0.13	0.30	1.00		
<i>9. Interventions motivationnelles et suppléance</i>	0.26	0.44	0.42	0.44	0.34	0.44	0.56	0.26	1.00	
<i>10. Travail administratif</i>	0.16	0.10	0.00	0.17	0.03	0.04	0.10	0.20	0.08	1.00

Tableau 7 : corrélations inter-items (N = 235) ; en gras les corrélations > .50

Au niveau des corrélations inter-items, les trois items qui présentent les inter-corrélations les plus élevées ($r \geq .60$) sont ceux relatifs à l'aide spécifique apportée au poste de travail (v. tableau 7). Ce résultat est logique. Ces trois items sont issus d'un seul et même facteur identifié initialement dans la grille ARBA originale et que l'on a retrouvé tout naturellement ici : ils représentent le séquençage en trois moments distincts d'une aide fournie autour des activités réalisées par les usagers à leur place de travail. Au moment de l'élaboration d'ARBA-R, sous l'impulsion des moniteurs, nous avons décidé de séparer les trois items malgré leur ressemblance sémantique et leur appartenance à un seul et même facteur. Par ailleurs l'item 2 (soutien psychologique et médical) corrèle positivement avec l'item consacré aux interventions lors de comportements perturbants ($r = .59$), lequel corrèle à son tour assez bien avec l'item 9 touchant les interventions motivationnelles et la suppléance en cas d'absence ($r = .56$). Il est donc là aussi normal que ces trois items se retrouvent dans le même facteur, intitulé pour l'occasion encadrement psychosocial et motivationnel (v. figure 7). Les items 1, 8 et 10 sont faiblement corrélés avec les autres items, comme on pouvait s'y attendre après l'examen de la structure factorielle (cf. figure 7).

La contribution de chacun des items au score total (corrélations item-total; on définit le total comme la somme de tous les items à l'exclusion de l'item corrélé) est présentée au tableau 8. Tous les items participent de façon équilibrée et pas trop élevée à l'évaluation de l'aide dispensée par les moniteurs. Ceci correspond bien au type de résultat espéré. L'item « *travail administratif* » est celui qui contribue le moins au score total. On notera que cet item est l'un des moins discriminatifs de l'échelle (avec les items 8 et 1) : 91.9% des usagers y ont un score égal à 0 ou 1. Sa faible contribution au score total est certainement liée à cela. On relèvera aussi que les deux autres items les moins discriminatifs sont également ceux qui présentent les corrélations item-total les plus basses.

	<i>Total (sans l'item)</i>
1. Aide pratique pour les actes du quotidien	0.35
2. Soutien psychologique et médical	0.44
3. Instruction et démonstration de l'activité	0.56
4. Co-réalisation de l'activité	0.61
5. Surveillance et sécurité	0.54
6. Vérification de l'activité	0.52
7. Interventions lors de comportements perturbants	0.61
8. Coordination avec le réseau général et déplacements	0.41
9. Interventions motivationnelles et suppléance	0.63
10. Travail administratif	0.15

Tableau 8 : corrélations item-total (N = 235)

ARBA présente une très bonne homogénéité interne (α de Cronbach = 0.81 ; cet indicateur, dimensionné entre 0 et 1, est élevé si les corrélations inter-items le sont aussi, il montre que tous les items contribuent de façon homogène au score global). En supprimant l'item 10 (« *tâches administratives* ») qui est apparu comme le plus faible de la série, l'homogénéité interne de l'échelle varie très peu (α = 0.82). Dès lors, malgré sa contribution apparemment minimale au score total, l'élimination de cet item n'améliore pas significativement l'homogénéité de l'échelle.

4.6 Synthèse et discussion

Les commentaires des moniteurs suite à cette phase pilote nous ont indiqué que les choix opérés au niveau du contenu se sont révélés judicieux : selon eux la nouvelle grille traduit de manière économique mais complète le travail réalisé auprès des usagers. Les définitions des items sont transparentes : il n'est plus nécessaire de recourir au manuel pour compléter la grille. Ces avis ont confirmé également la pertinence des dimensions retenues lors de l'élaboration de la grille.

L'analyse psychométrique a mis en évidence que tous les items de la grille sont utilisés à une fréquence satisfaisante. Il n'y a plus comme auparavant une série d'items systématiquement cotés à zéro. Par ailleurs, la nouvelle grille reflète une structure empirique assez claire, simple et logique, même s'il y a trois items dont la position dans cette structure tridimensionnelle est moins claire. Enfin la forte homogénéité de l'échelle n'incite pas à en supprimer des éléments.

L'expérience issue de cette étude pilote a montré toutefois que des aménagements relativement simples méritaient d'être apportés. Le double système de référence (adjectival-quantitatif) peut être source de confusion. Il est apparu à la réflexion pertinent de le supprimer et de ne conserver que la cotation quantitative (durées exprimées en minutes). Le manque de sensibilité de l'échelle des durées devait être corrigé par un choix d'intervalles plus fins. Les aides peu fréquentes dispensées en l'absence de l'utilisateur (travail administratif et réseau) sont difficiles à évaluer sur une base journalière (p. ex. tenue de dossier, réunion de réseau). Après concertation, il a paru plus adapté d'évaluer ces interventions sur une base mensuelle. Cette alternative devrait se révéler plus naturelle à l'usage, elle devrait aussi éviter aux moniteurs de faire des calculs pour rapporter à une base journalière une évaluation de type mensuel.

5. EVALUATION D'ARBA-2

5.1 Présentation ARBA-2

Les principaux changements par rapport à ARBA-R ont été effectués essentiellement dans la structure de l'instrument et ont moins touché la composition ou la teneur même des items qui s'était révélée à l'usage satisfaisante. C'est à l'issue d'une réflexion commune avec les moniteurs impliqués dans l'évaluation d'ARBA-R que la grille ARBA-2 a été mise au point. Pour tenir compte de la dimension « *encadrement administratif* » (3^e facteur) qui semblait plus facile à évaluer sur une base mensuelle que journalière, pour tenir compte aussi des remarques sur le manque de prise en considération du travail accompli pour aider les usagers dans leurs projets professionnels, nous avons décidé de décomposer clairement la grille en deux parties indépendantes. La première partie (items 1-8) est destinée à évaluer l'encadrement direct alors que la deuxième partie (items A et B) porte sur les prestations administratives, d'orientation et de réseau (v. tableau 9).

<i>Items</i>
Encadrement direct
1. <i>Aide pratique pour les actes du quotidien</i>
2. <i>Soutien psychologique et médical</i>
3. <i>Instructions et démonstration de l'activité</i>
4. <i>Co-réalisation de l'activité</i>
5. <i>Surveillance et sécurité</i>
6. <i>Vérification du travail fourni</i>
7. <i>Interventions lors de comportements perturbants</i>
8. <i>Interventions motivationnelles et suppléance</i>
Prestations administratives, d'orientation et réseau
A. <i>Démarches d'orientation, réinsertion et formation</i>
B. <i>Travail administratif et réseau</i>

Tableau 9: structure et items d'ARBA-2

ARBA-2 aboutit désormais à deux scores indépendants, chacun étant la somme des items qui le constituent, à savoir : (1) un score d'encadrement direct (ARBA-ED, items 1-8), (2) un score de prestations administratives, d'orientation et de réseau (ARBA-AOR, items A-B).

Le système de cotation a été également modifié : le recours aux adjectifs a été abandonné au profit de la seule référence en minutes. Les huit premiers items rendent compte de l'encadrement direct journalier moyen fourni par un moniteur à un usager (v. figure 8). L'échelle adoptée est étalonnée de 0 à 60 minutes par incréments de 5 minutes. Il s'agit de la durée totale d'encadrement dispensé en moyenne par le moniteur sur une journée où l'utilisateur est présent dans l'atelier. Le score ARBA-ED pour l'encadrement direct peut être ensuite transformé selon les besoins en un score hebdomadaire, mensuel ou annuel, en multipliant le score journalier par le nombre de jours travaillés sur la période considérée.

Cocher la durée en minutes correspondant le mieux à la durée totale de l'encadrement dispensé en moyenne sur une journée de travail type et quelle que soit la présence effective de l'utilisateur.												
0'	5'	10'	15'	20'	25'	30'	35'	40'	45'	50'	55'	60'

Figure 8 : système de cotation choisi pour les items 1-8 d'ARBA-2 (encadrement direct)

Le score d'encadrement administratif, d'orientation et réseau (ARBA-AOR) correspond à la somme des durées indiquées pour les items A et B. Il s'agit des durées moyennes des prestations administratives et d'accompagnement dans un travail de réseau et d'orientation professionnelle fournies pour un usager sur un mois. Sur la base des discussions avec un panel de moniteurs, la durée maximum des interventions a été fixée à trois heures par mois, par incréments de 30 minutes (v. figure 9). Comme pour l'encadrement direct, le score ARBA-AOR peut être transformé en scores trimestriels, semestriels ou annuels par simple multiplication du score mensuel initial.

Cocher la catégorie correspondant le mieux à la durée totale des prestations dispensées à l'utilisateur lors du dernier mois.						
0'	½ heure	1h	1h30	2h	2h30	3h

Figure 9 : système de cotation choisi pour les items A-B d'ARBA-2 (prestations administratives, d'orientation et de réseau)

Pour la phase test à grande échelle de l'outil, nous avons proposé aux utilisateurs de la grille ARBA-2 un espace de commentaires libres ainsi que trois questions supplémentaires qui ne figuraient pas dans ARBA-R. Pour répondre aux deux premières questions, la durée d'encadrement totale (score ARBA-ED : items 1-8) était automatiquement calculée et apparaissait à l'écran après la compilation de la grille. Dans la première question on demandait au moniteur de noter si cette durée lui semblait correcte, surestimée, ou sous-estimée. Si elle ne lui paraissait pas correctement évaluée, une

seconde question l'invitait à insérer son estimation personnelle du temps d'encadrement total dédié par jour à l'utilisateur. Ces deux questions avaient pour objectif de contrôler la qualité des évaluations de l'aide directe réalisées avec ARBA-2, en les comparant aux estimations « intuitives » des moniteurs. La troisième question concernait le besoin d'aide. Des moniteurs nous avaient signalé que, faute de temps ou d'une disponibilité suffisante, ils n'étaient pas toujours en mesure de dispenser des prestations qu'ils jugeaient pourtant nécessaires ou utiles. Il a donc été systématiquement demandé d'indiquer dans cette troisième question la durée d'encadrement devant être idéalement dispensé à l'utilisateur (encadrement idéal).

5.2 Analyse des contenus

La grille ARBA-2 a été testée entre mai et juin 2008 auprès de quinze des seize institutions ayant participé à l'évaluation de la grille ARBA originale. Le contenu des items a généralement été très bien accepté. Personne n'a formulé de remarque majeure lors des séances de formation ou à d'autres moments de l'évaluation. La nouvelle structure en deux parties a aussi été favorablement accueillie. Cette modification a permis de faciliter la démarche d'évaluation et de mieux rendre compte de cet aspect particulier du travail d'encadrement.

Nous avons réparti en trois catégories les commentaires libres fournis par les moniteurs à l'issue de l'évaluation de certains usagers (v. tableau 10). Outre un certain nombre de précisions relatives aux causes des interventions et à la nature de celles-ci, ces commentaires mettent en évidence plusieurs interrogations ou inquiétudes liées à l'utilisation de la grille, notamment en ce qui concerne la représentativité de l'évaluation au moment où elle est conduite. La notion de journée type ne semble pas toujours facile à délimiter, en particulier lorsque les usagers fréquentent de façon chaotique les ateliers, passent d'un atelier à l'autre, voire sont suivis à l'extérieur. La particularité de certaines situations a engendré un certain nombre de questions concrètes. Elles nous ont obligé à formuler des réponses précises, qui figureront dans le manuel d'utilisation et permettront à l'avenir d'améliorer encore la fidélité interjuges de l'instrument. Le manuel d'utilisation n'en sera que plus complet et efficace car il favorisera une meilleure harmonisation des pratiques d'évaluation. Heureusement, les commentaires touchant la définition même des items sont globalement positifs et ne nous conduisent pas à opérer des changements.

Nous avons également relevé la pertinence de l'item concernant l'estimation du besoin d'encadrement à proprement parler. En effet plusieurs commentaires ont signalé le fait que dans certains cas les structures ne pouvaient pas offrir l'encadrement escompté pour un accompagnement satisfaisant de l'utilisateur. Ici la différence entre aide effective et besoin d'aide apparaît comme tout à fait cruciale ; elle rejoint la réflexion que nous avons amorcée sur l'intitulé même de la grille dès la première évaluation critique d'ARBA (cf. p.6)

Catégorie	Commentaires
<i>Les usagers et les structures</i>	<p>vieillesse, grande autonomie, grande dépendance, troubles mnésiques, troubles de la communication, troubles du comportement, troubles psychiques ou physiques spécifiques, etc., situations pour lesquelles les structures peuvent difficilement faire face (scores extrêmes par rapport aux scores moyens observés dans l'atelier) ;</p> <p>emploi du temps (taux d'occupation partiels, horaires particuliers) ; différence entre taux d'engagement et présence effective, définition de la journée-type ;</p> <p>fréquentation par le même usager de plusieurs ateliers au sein de l'institution ; usagers travaillant hors ateliers mais suivis à l'extérieur par des moniteurs de l'institution ;</p> <p>arrivée récente dans l'atelier, formation ou apprentissage en cours.</p>
<i>L'utilisation de la grille</i>	<p>étalonnage trop large pour les items des prestations administratives ;</p> <p>difficulté de compréhension de certaines consignes ;</p> <p>clarté des items et de leur définition, facilité d'emploi de la grille.</p>
<i>Réalités et besoins</i>	<p>inadéquation entre niveau de fonctionnement de l'utilisateur et possibilités réelles d'encadrement ; ressources en temps et personnel insuffisantes ;</p> <p>besoin de plus de temps d'écoute et d'échange, de plus d'activités physiques, d'accompagnement, de temps pour élaborer des projets, etc.</p>

Tableau 10 : commentaires libres des moniteurs lors de l'utilisation d'ARBA-2

5.3 Analyse des items

Parmi les 1864 usagers évalués, 63.8% sont des hommes (n = 1189) et 36.2% des femmes (n = 675). L'âge moyen des usagers est de 43.8 ans (sd = 13.4). La distribution de l'âge dans l'échantillon présente des caractéristiques très similaires à celle obtenue lors de l'analyse de la première grille (v. figure 10). Les personnes évaluées fréquentent en moyenne leur institution depuis 11.2 ans (sd = 12.3), ceci à raison de 4.3 journées de présence par semaine (sd = 1.2). 69.4% des usagers se rendent tous les jours de la semaine à leur atelier. L'ancienneté dans l'institution est significativement liée à l'âge ($r = .451, p < .01$) ; les usagers les plus âgés fréquentent généralement leur institution actuelle depuis plus longtemps que les personnes plus jeunes. Les usagers ne diffèrent pas en fonction de leur sexe dans le nombre d'années passées dans l'institution ($t(1840) = -.135, p < .01$). De même, femmes et hommes ne diffèrent pas au niveau de leur âge moyen ($t(1843) = -.666, p < .01$).

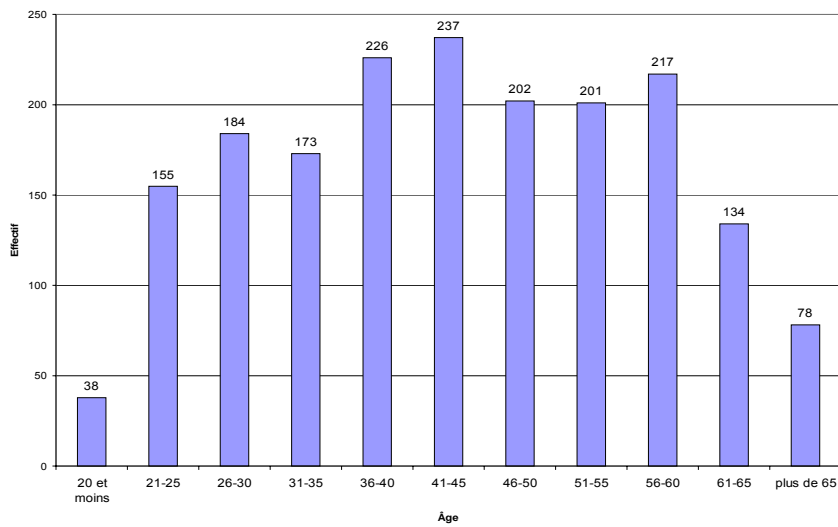


Figure 10 : distribution de l'âge (N = 1864)

Analyse des distributions et des moyennes

Le score total moyen obtenu pour l'ensemble des usagers à la dimension « *encadrement direct* » est de 12.39 points (sd = 10.01) et la médiane de 10. Le 50% des effectifs obtient un score entre 6 et 15 points (v. figure 11).

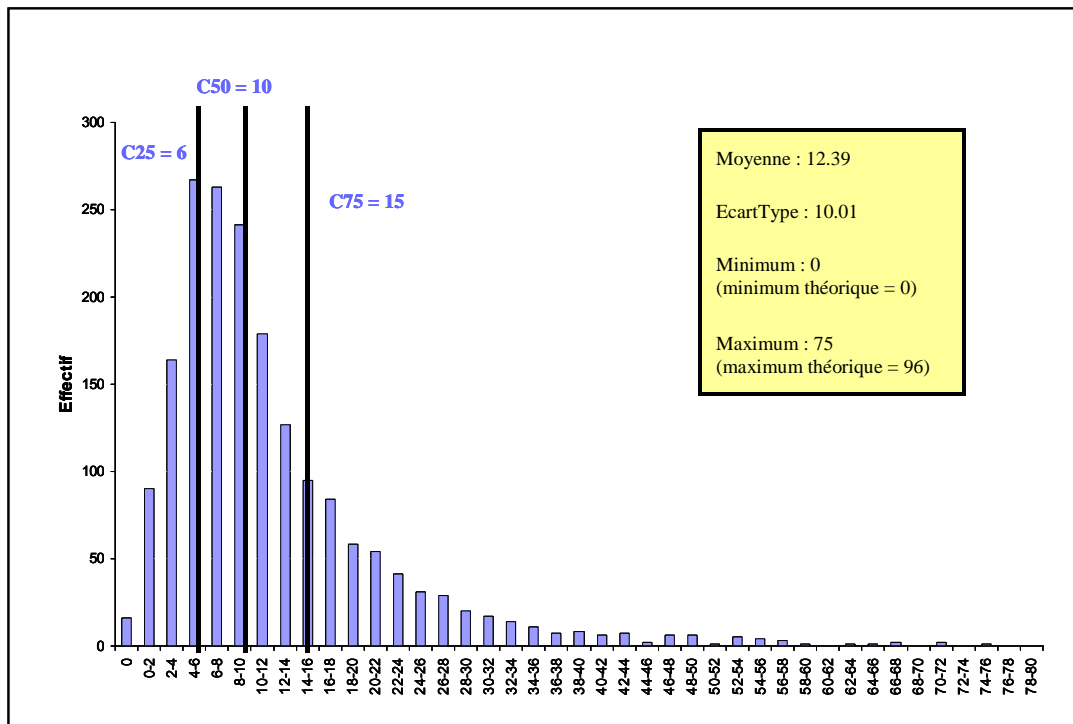


Figure 11 : distribution des scores totaux d'encadrement direct

Le score total pour les prestations administratives, d'orientation et réseau est de 3.02 points (sd = 2.22) et la médiane de 2. La moitié des sujets obtiennent des scores entre 1 et 4 points (v. figure 12).

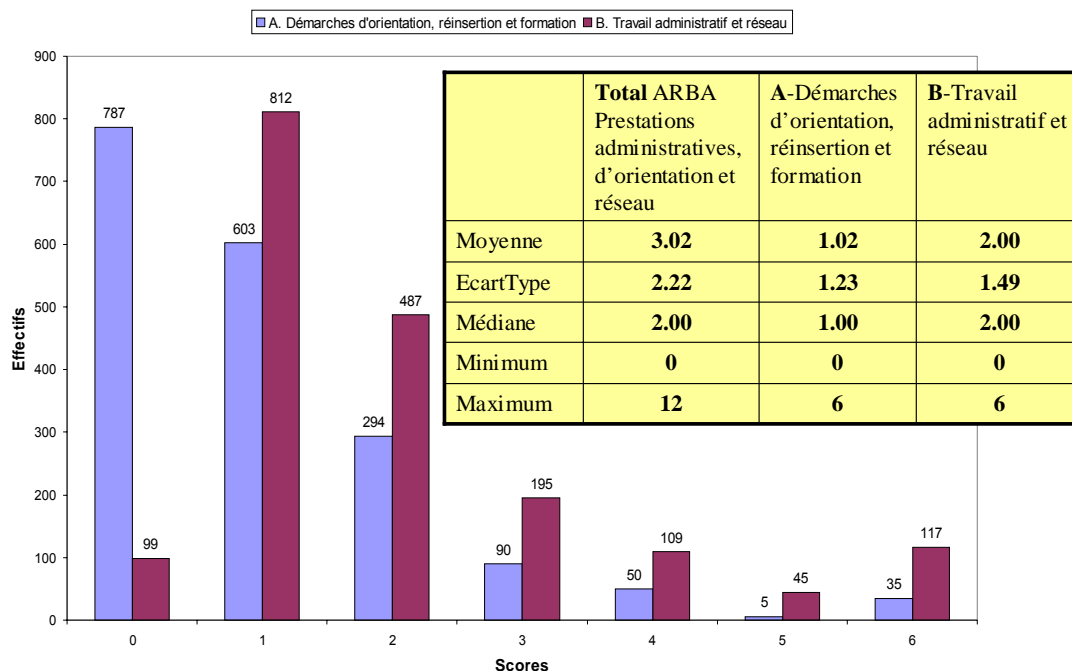


Figure 12 : distribution des scores aux deux items des prestations administratives, d'orientation et de réseau

Les moyennes aux différents items figurent au tableau 11. L'item obtenant la moyenne la plus élevée est celui du soutien psychologique et médical. L'item « aide pratique pour les actes du quotidien » est celui qui obtient la moyenne la plus faible.

Items	<i>Moyennes</i>
2. Soutien psychologique et médical	2.17
3. Instruction et démonstration de l'activité	1.84
4. Co-réalisation de l'activité	1.81
6. Vérification du travail fourni	1.62
8. Interventions motivationnelles et suppléance	1.61
7. Interventions lors de comportements perturbants	1.39
5. Surveillance et sécurité	1.08
1. Aide pratique pour les actes du quotidien	0.86
B. Travail administratif et réseau	2.00
A. Démarches d'orientation, réinsertion et formation	1.02

Tableau 11 : moyennes pour chaque item (rangées par ordre décroissant)

De façon générale, les interventions des moniteurs se situent principalement au niveau de l'encadrement des activités liées au travail et dans le soutien psychologique et médical.

Le score ARBA-2 moyen pour l'encadrement direct des hommes est de 11.80 points alors que celui des femmes est de 13.42. La différence s'avère statistiquement significative ($t(1227.168) = -3.224$, $p < .05$). Nous avons également observé une différence statistiquement significative entre hommes et femmes pour les scores de prestations administratives, d'orientation et de réseau ($t(1862) = -2.494$), mais qui n'est pas très élevée en valeur absolue. En effet, les femmes obtiennent en moyenne des scores plus élevés de 0.27 point seulement. Relevons encore que cette différence entre sexes avait déjà été mise en évidence au niveau des scores de la grille initiale.

Nous observons un lien faible, bien que statistiquement significatif, entre l'âge des usagers et les deux sous-totaux d'ARBA-2 (encadrement direct : $r = -.142$, $p < .01$, prestations administratives, d'orientation et de réseau : $r = -.184$, $p < .01$). Ces corrélations sont toutefois trop faibles pour affirmer que les usagers plus jeunes ont véritablement tendance à obtenir des scores ARBA plus élevés.

Comme pour ARBA-R, il n'existe aucun item où une large partie des usagers obtient un score nul (v. tableau 12). Tous les items de la grille sont régulièrement sollicités par les moniteurs. Une partie importante des scores est localisée sur la première moitié du barème. En effet, 95% des usagers obtiennent des scores localisés entre 0 et 6. Les scores compris entre 8 et 12 sont plutôt rares mais néanmoins occasionnellement utilisés. Ils légitiment le maintien d'un empan total de 60 minutes sectionné en 12 points, car cela permet de rendre compte de certaines situations très particulières voire exceptionnelles, sans devoir modifier l'outil.

<i>items</i>	<i>Percentile 95</i>	<i>% scores compris entre 0 et 8</i>	<i>% scores nuls</i>
1. Aide pratique pour les actes du quotidien	4.0	98.3	66.1
2. Soutien psychologique et médical	6.0	97.4	15.3
3. Instructions et démonstration de l'activité	4.0	99.7	8.2
4. Co-réalisation de l'activité	6.0	98.1	23.2
5. Surveillance et sécurité	4.0	99.1	37.8
6. Vérification du travail fourni	4.0	99.7	15.6
7. Interventions lors de comportements perturbants	6.0	97.4	46.4
8. Interventions motivationnelles et suppléance	5.0	98.8	26.7
		% scores compris entre 0 et 4	
A. Démarches d'orientation, réinsertion et formation	3.0	97.9	42.2
B. Travail administratif et réseau	6.0	91.3	5.3

Tableau 12 : répartition des scores aux dix items de la grille ARBA-2

Structure de la grille et homogénéité des scores

L'analyse factorielle conduite sur l'ensemble des dix items permet de retenir une solution à trois facteurs qui semble tout à fait bien traduire les dimensions sous-jacentes de la grille. Cette solution permet de garder 63.4% de l'information contenue dans les dix items. Les trois dimensions pouvant être identifiées sont l'encadrement psychosocial et motivationnel (items 7,2,1,5,8), l'encadrement au poste de travail (items 3,6,4), ainsi que les prestations administratives, d'orientation et de réseau (items A-B ; v. tableau 13).

Items	Facteurs		
	I	II	III
1. Aide pratique pour les actes du quotidien	0.72	0.06	0.03
2. Soutien psychologique et médical	0.75	0.24	0.15
3. Instructions et démonstration de l'activité	0.09	0.84	0.10
4. Co-réalisation de l'activité	0.29	0.72	0.07
5. Surveillance et sécurité	0.70	0.23	0.09
6. Vérification du travail fourni	0.23	0.80	0.04
7. Interventions lors de comportements perturbants	0.82	0.11	0.06
8. Interventions motivationnelles et suppléance	0.63	0.39	0.09
A. Démarches d'orientation, réinsertion et formation	-0.03	0.04	0.87
B. Travail administratif et réseau	0.25	0.12	0.73

Tableau 13 : solution à 3 facteurs après rotation varimax ; les saturations supérieures à |0.50| sont en gras

La structure observée avec ARBA-R se trouve confirmée ici avec un échantillon d'utilisateurs beaucoup plus conséquent. De plus, les items A et B (anciens items 8 et 10 d'ARBA-R) figurent désormais dans un facteur qui leur est propre.

Comme on pouvait s'y attendre d'après les résultats de l'analyse factorielle, les plus fortes intercorrélations concernent les items 2, 7, 8 et 5 qui traitent d'aspects liés au fonctionnement psychosocial, ainsi que les items 3, 4, et 6 qui touchent plutôt l'encadrement au poste de travail (v. tableau 14). L'aide pratique au quotidien reste un item assez autonome et peu corrélé avec les autres. Pour terminer, les items A et B entretiennent des corrélations très faibles avec les huit autres items évaluant l'encadrement direct et sont eux-mêmes modérément intercorrélés.

La contribution des items 1 à 8 au score ARBA-ED (corrélations item-total) est présentée au tableau 15. Nous y constatons des contributions relativement homogènes pour chacun des huit items composant la dimension d'encadrement direct de la grille.

Comme souhaité, tous les items de cette première partie contribuent de façon moyenne aux scores d'encadrement direct. D'ailleurs l'échelle de l'encadrement direct présente une très bonne homogénéité interne sur la mesure de Cronbach ($\alpha = 0.83$).

Items	1	2	3	4	5	6	7	8	A	B
1. Aide pratique pour les actes du quotidien	1.00									
2. Soutien psychologique et médical	0.43	1.00								
3. <i>Instruction et démonstration de l'activité</i>	0.16	0.29	1.00							
4. <i>Co-réalisation de l'activité</i>	0.35	0.39	0.49	1.00						
5. Surveillance et sécurité	0.48	0.44	0.31	0.31	1.00					
6. <i>Vérification du travail fourni</i>	0.19	0.33	0.53	0.50	0.34	1.00				
7. Interventions lors de comportements perturbants	0.40	0.62	0.20	0.28	0.51	0.31	1.00			
8. Interventions motivationnelles et suppléance	0.34	0.54	0.33	0.40	0.42	0.44	0.53	1.00		
A. Démarches d'orientation, réinsertion et formation	0.06	0.12	0.14	0.09	0.09	0.04	0.04	0.08	1.00	
B. Travail administratif et réseau	0.17	0.28	0.15	0.21	0.23	0.23	0.25	0.26	0.32	1.00

Tableau 14 : corrélations inter-items

En ce qui concerne la seconde partie de la grille (score ARBA-AOR), les corrélations item-total pour chacun des deux items est de 0.32. La mesure alpha de Cronbach est à 0.48, ce qui correspond à une homogénéité modérée. Ce résultat moins probant que pour l'échelle d'encadrement direct n'est guère étonnant dans la mesure où la dimension des prestations administratives, d'orientation et de réseau n'est composée que de deux items faiblement intercorrélés.

Enlever un des items n'aurait guère de sens : l'échelle se réduirait alors à une seule et unique question. Ce qui pourrait éventuellement être envisagé est au contraire un enrichissement par quelques items supplémentaires : on y gagnerait sans doute au niveau de la mesure d'homogénéité ou des corrélations item-total, mais on y perdrait sur le plan de la concision. Il s'agit évidemment d'une question de choix stratégique. En l'occurrence on peut estimer aujourd'hui que le principe d'économie doit primer sur des indices d'homogénéité ou des corrélations item-total. A l'usage cependant, s'il s'avère que ce type de prestations mérite d'être plus détaillé, il sera toujours temps de considérer l'élargissement de cette échelle, avec l'avantage de pas devoir toucher aux qualités psychométriques actuelles de l'échelle d'encadrement direct.

	<i>Total (sans l'item)</i>
1. Aide pratique pour les actes du quotidien	0.49
2. Soutien psychologique et médical	0.66
3. Instruction et démonstration de l'activité	0.45
4. Co-réalisation de l'activité	0.54
5. Surveillance et sécurité	0.59
6. Vérification du travail fourni	0.53
7. Interventions lors de comportements perturbants	0.61
8. Interventions motivationnelles et suppléance	0.64

Tableau 15 : corrélations item-total (ARBA-ED) pour la première partie d'ARBA-2

5.4 Analyses complémentaires

Les questions complémentaires posées en fin de grille sur la qualité des estimations fournies par ARBA-2 ont également fait l'objet d'analyses statistiques. De plus, comme autre mesure de la validité des évaluations fournies par ARBA-2, nous avons comparé deux classements des mêmes usagers par rapport au temps d'encadrement qui leur était dédié. Le premier classement était établi sur une base plutôt intuitive par les moniteurs eux-mêmes, le second résultait des scores ARBA-2. Nous avons profité de vérifier sur des échantillons plus restreints le niveau de fidélité test-retest de l'instrument à deux semaines d'intervalle. Pour terminer, nos analyses ont porté sur l'évaluation du besoin d'encadrement que l'on a souhaité comparer à l'encadrement réellement fourni, ceci afin d'observer s'il existe ou non un écart entre ces deux types de mesure.

Qualité des estimations

Les moniteurs pensent-ils que la grille ARBA-2 fournit des estimations correctes du temps d'encadrement direct des usagers ? Une très large majorité des grilles complétées (86%, soit 1612 grilles) aboutit à une réponse positive. ARBA-2 semble apporter des estimations du temps dédié aux usagers jugées correctes par les moniteurs (v. figure 13). Le nombre de grilles surestimant ou sous-estimant la durée d'encadrement réelle est quasi identique (124 sous-estimations & 128 surestimations) ; il ne représente que le 14% du total des évaluations.

Le score moyen ARBA d'encadrement direct (ARBA-ED) est égal à 12.39 points (sd = 10.01 ; cf. figure 11) alors que la variante corrigée de ce même score est de 11.94 (sd = 9.51). Le score ARBA-

ED corrigé correspond à l'estimation fournie par un moniteur lorsqu'il pensait que le score calculé par la grille ne reflétait pas bien la durée d'encadrement réelle. Quand les moniteurs considéraient comme correct un score issu de l'évaluation des huit premiers items, nous avons évidemment repris ce score pour l'utiliser comme score corrigé et nous permettre ensuite d'effectuer les comparaisons. La moyenne des écarts entre les deux types de scores se monte à 2.22 minutes (3.58%). La différence est faible (0.45 point ARBA), en dessous du seuil de discrimination de l'échelle.

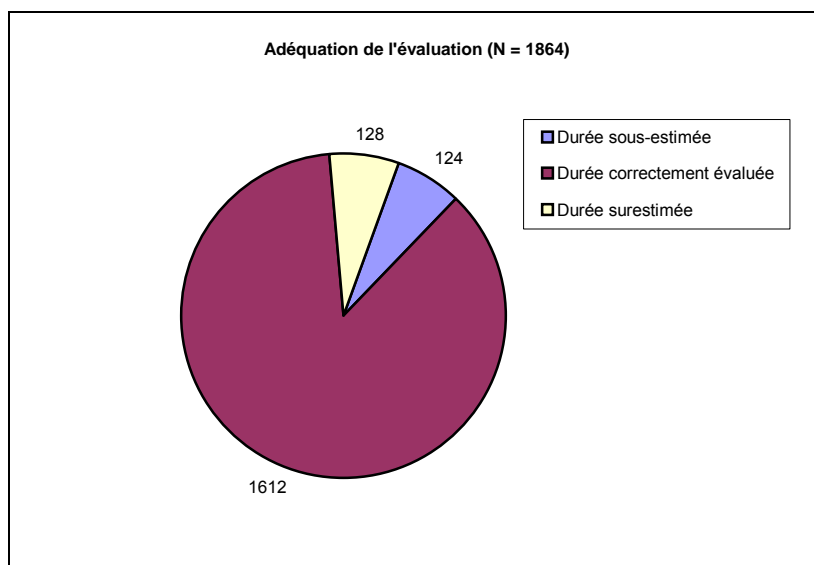


Figure 13 : pourcentages des évaluations du temps d'encadrement jugées correctes, sous-estimées et surestimées

Nous avons également examiné la magnitude des écarts institution par institution (v. figure 14). Nous ne constatons aucune différence significative pour dix des quinze institutions ayant participé à l'étude. Pour quatre des cinq institutions restantes, l'erreur est inférieure ou égale à 1 point ARBA (5 minutes). Nous constatons enfin une différence de 4.45 points sur une institution. Il nous est difficile de dire au regard des données à notre disposition si cet écart est dû à une particularité du mode de fonctionnement de la dite structure. Il est également possible que l'erreur ait été en partie accentuée en raison du score moyen très élevé observé à cet endroit. De façon plus générale, une large majorité des grilles produit une évaluation jugée correcte par les moniteurs et l'écart moyen constaté est relativement faible surtout si l'on exclut de nos analyses l'institution ayant produit les écarts les plus importants.

Une autre façon d'évaluer la qualité des estimations fournies par ARBA-2 a été de procéder à une comparaison supplémentaire entre le classement des usagers sur l'aide reçue et leurs scores ARBA. Nous avons demandé à des moniteurs travaillant dans des ateliers comportant un nombre modéré de patients (entre dix et quinze environ) de classer leurs usagers, en leur attribuant des rangs en fonction

de l'aide qui leur est habituellement dispensée, partant de l'utilisateur demandant le plus d'attention pour terminer avec celui qui nécessite le moins d'encadrement. Nous avons ensuite corrélé les rangs donnés par les moniteurs avec les scores ARBA-ED. Cette comparaison a été réalisée dans six ateliers (chacun provenant d'une institution différente), dont les effectifs étaient compris entre sept et seize usagers ($m = 13.67$ usagers, $sd = 3.61$).

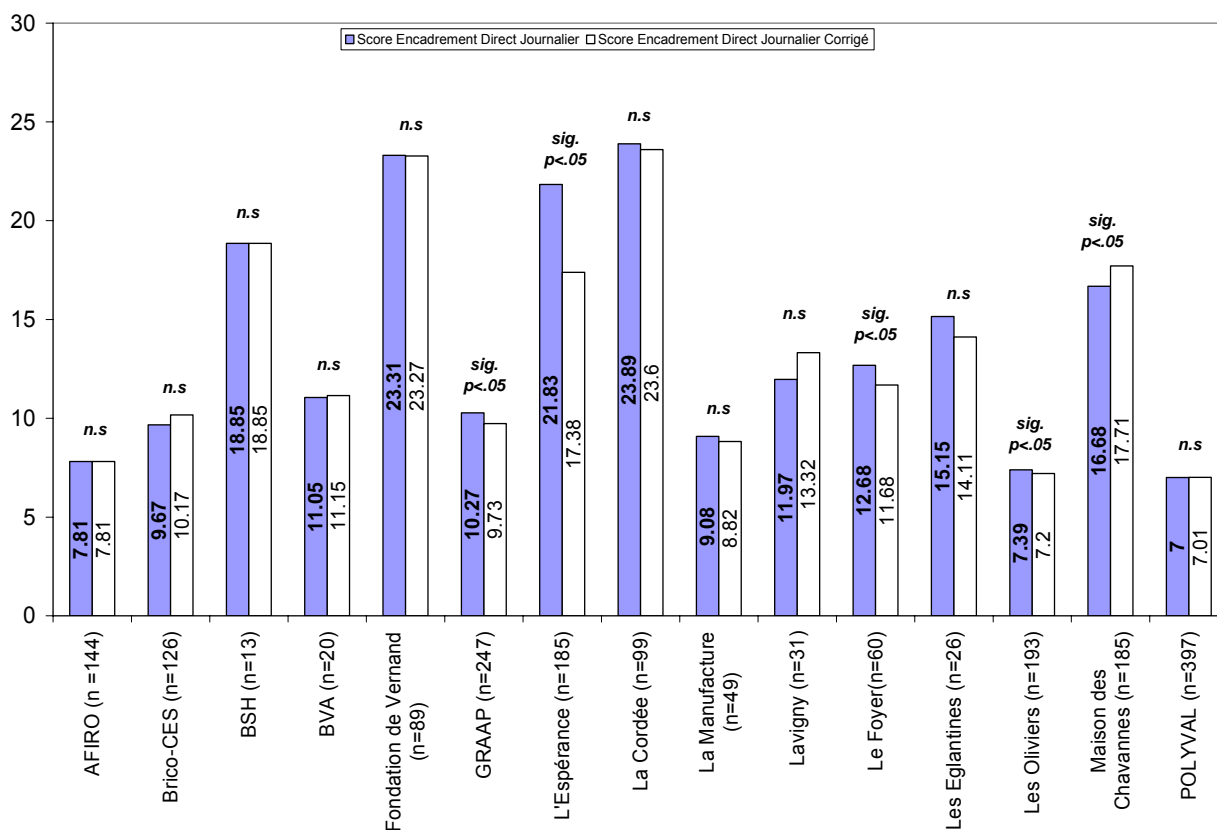


Figure 14 : différences entre scores bruts et corrigés, institution par institution

La corrélation *Rho de Spearman* moyenne sur ces six ateliers est de 0.71 ($sd = 0.24$). Ce résultat est satisfaisant : la grille ARBA permet d'obtenir grosso modo le même classement des personnes handicapées selon l'aide qu'elles reçoivent que le classement subjectif effectué par les moniteurs sans faire appel à une quelconque mesure.

Fidélité test-retest

Afin de vérifier la fiabilité des résultats de la grille ARBA-2, nous avons décidé d'évaluer à deux reprises quarante-huit usagers. Quatre usagers sélectionnés au hasard ont été évalués deux fois dans chaque institution. Le délai retenu entre les deux évaluations était d'environ une à deux semaines. Ce délai ne devait pas être trop court pour éviter des biais de mémoire trop évidents, ni trop long afin d'éviter que des changements trop importants (p. ex. accident, difficultés personnelles ou familiales, changement d'atelier, etc.) puissent se produire dans l'intervalle et venir légitimement modifier les scores. Si un des usagers sélectionnés avait rencontré d'importants changements entre les évaluations, consigne avait été donnée de sélectionner à la place et au hasard, un autre usager du même atelier. Nous avons exclu trois grilles de l'analyse : un atelier nous a informés avoir récemment déménagé vers un nouveau site, ce qui a produit de l'avis des moniteurs des changements majeurs dans le quotidien de deux des usagers réévalués. Un deuxième atelier nous a signalé avoir commis une erreur lors de la deuxième évaluation. L'évaluation de la fidélité test-retest a porté au final sur quarante-cinq usagers ; ses résultats sont présentés au tableau 16.

	<i>ICC</i> ¹	<i>R de Pearson</i> ²	<i>Rho de Spearman</i> ³
Totaux			
Encadrement direct (8 items)	0.856**	0.869**	0.839**
Prestations administratives, d'orientation et réseau (2 items)	0.684**	0.695**	0.846**
Items		<i>Rho de Spearman</i> ³	
1. Aide pratique pour les actes du quotidien		0.616**	
2. Soutien psychologique et médical		0.798**	
3. Instructions et démonstration de l'activité		0.734**	
4. Co-réalisation de l'activité		0.751**	
5. Surveillance et sécurité		0.749**	
6. Vérification du travail fourni		0.812**	
7. Interventions lors de comportements perturbants		0.676**	
8. Interventions motivationnelles et suppléance		0.694**	
A. Démarches d'orientation, réinsertion et formation		0.892**	
B. Travail administratif et réseau		0.712**	

¹Corrélation intra-classe, ²coefficient de linéarité, ³coefficient de monotonie, *p<0.05, **p<0.01

Tableau 16 : coefficients de fidélité test-retest à 2 semaines (N = 45) ; les corrélations supérieures à |0.50| sont en gras

Les coefficients de fidélité test-retest sont satisfaisants, cela tant au niveau des items qu'au niveau des scores totaux. Toutes les corrélations calculées sont statistiquement significatives au seuil de 1%. Les évaluations ARBA-2 possèdent une stabilité dans le temps suffisante sur l'intervalle de deux semaines. Malgré le faible nombre d'usagers évalués à deux reprises au regard de l'échantillon total,

les statistiques présentées peuvent être considérées comme fiables, avec un taux d'erreur alpha de 5% et un taux d'erreur beta de 5%¹.

Evaluation de l'encadrement idéal

Les usagers pourraient dans certaines situations bénéficier de prestations supplémentaires pour développer ou acquérir de nouvelles compétences. Par manque de temps, les moniteurs ne sont pas toujours en mesure de fournir l'ensemble des prestations qui selon eux sont nécessaires pour permettre aux usagers de fonctionner de façon optimale dans leur environnement de travail. Nous avons donc demandé aux utilisateurs de la grille d'estimer quelle serait la durée d'encadrement quotidien total dont devraient idéalement bénéficier les usagers évalués. Dans le cas où cet encadrement idéal correspondait parfaitement à l'encadrement total quotidien calculé via les scores ARBA-ED, nous avons indiqué aux moniteurs qu'ils devaient reporter dans cette rubrique la valeur d'encadrement direct telle que chiffrée par la grille.

Malheureusement les données récoltées indiquent une confusion au niveau de la compréhension des consignes. Un nombre important de grilles comporte des scores nuls ou inférieurs aux scores de l'encadrement réellement dispensé. Un certain nombre d'utilisateurs a sans doute raisonné en termes de temps additionnel et non pas en termes de durée d'encadrement idéal sur une journée. Même s'il est possible d'exclure ou de corriger une partie de ces données, rien ne nous permet de garantir que les grilles présentant des scores en apparence cohérents n'aient pas, elles aussi, été complétées sur la base d'une compréhension erronée de la consigne. Il ne nous a donc pas été possible de les exploiter de manière fiable. Pour prendre en compte à l'avenir l'évaluation par les moniteurs de l'encadrement idéal dont auraient besoin les usagers, il sera donc nécessaire de reformuler l'item afin d'éviter des erreurs liées à une mauvaise interprétation de la consigne.

5.5 Synthèse et discussion

Il nous semble avec ARBA-2 avoir abouti à une formule adéquate pour le contenu et le système de cotation des items. Les références adoptées permettent de coter la très grande majorité des situations et laissent une certaine marge aux moniteurs pour rendre compte des situations particulières nécessitant un encadrement plus soutenu. L'échelle de mesure proposée pour les prestations administratives, d'orientation et de réseau paraît également correctement dimensionnée. Une augmentation de l'empan pourrait être envisagée afin de mieux tenir compte d'autres types de fonctionnement dans d'autres cantons. Cette modification n'aurait toutefois pas de conséquence sur les résultats des évaluations telles que menées dans le canton de Vaud.

¹ Erreur alpha : probabilité de 5% de considérer par erreur que la corrélation trouvée n'est pas due au hasard alors qu'elle l'est ; erreur beta : probabilité de 5% de considérer que la corrélation trouvée est due au hasard alors qu'elle ne l'est pas. Les chiffres sont ceux habituellement appliqués en statistiques pour l'erreur alpha. Le critère fixé pour l'erreur de type beta est quant à lui plus sévère que les valeurs habituellement retenues (20%).

Les dimensions issues de l'analyse factorielle correspondent au modèle obtenu avec la grille ARBA-R. Les prestations des moniteurs peuvent être décomposées selon trois grands axes : (1) l'encadrement du patient sur son poste de travail, (2) le soutien psychosocial et motivationnel, (3) les activités administratives, d'orientation et de réseau. Le pourcentage de variance expliquée par les trois facteurs est assez bon, l'homogénéité des scores ED et AOR satisfaisante. Les résultats de l'étude pilote d'ARBA-R sont donc confirmés avec ARBA-2 sur un échantillon nettement plus conséquent.

Le système de correction des scores invitant les moniteurs à estimer si la durée calculée par la grille leur semble adéquate nous a permis de vérifier la qualité des résultats obtenus. L'écart constaté en moyenne est faible, en dessous du seuil de discrimination de l'échelle. Pensé avant tout comme un indicateur de contrôle utile à la validation psychométrique de la grille, ce système de correction n'apparaît plus désormais nécessaire à un usage routinier de celle-ci.

L'évaluation des besoins d'encadrement idéal pourrait quant à elle être conservée, moyennant une modification de la consigne. Pour rendre cette tâche plus claire, il s'agirait de demander aux moniteurs d'indiquer le temps additionnel nécessaire par rapport au temps d'encadrement effectif calculé par le système informatique. Il s'agira de mettre un zéro si l'encadrement effectif se révèle satisfaisant et n'appelle pas un supplément de temps. Cette alternative permettra peut-être d'empêcher la confusion associée à l'ancienne consigne. Il est sans doute plus naturel pour estimer la durée idéale d'encadrement de dire combien de temps supplémentaire (ou à retrancher) il faut par rapport au temps déjà fourni que d'énoncer à nouveau une durée totale. Inscrire un zéro semble aussi plus rapide que devoir reporter une durée calculée informatiquement. Il serait également opportun de permettre aux moniteurs de préciser dans quels domaines ces prestations supplémentaires méritent d'être dispensées. On pourrait ainsi demander aux moniteurs de spécifier si le temps additionnel doit être consacré aux domaines de l'encadrement au poste de travail, de l'encadrement psychosocial et motivationnel ou enfin sous forme de prestations administratives, d'orientation et de réseau. Cette procédure serait plus riche que la mention d'une durée unique et s'appuierait sur les dimensions sous-jacentes à l'instrument.

Pour terminer, nous pensons conserver le champ destiné aux commentaires libres, et le déplacer en fin de grille pour bien signaler que ces commentaires portent sur l'ensemble des rubriques de l'instrument. Il ressort d'ailleurs des commentaires laissés par les moniteurs qu'un certain nombre de précisions et d'instructions supplémentaires mériteraient de figurer dans un manuel d'utilisation. Il apparaît en effet nécessaire d'expliquer clairement aux utilisateurs la méthode d'évaluation retenue pour l'usage régulier de la grille (p. ex. sondage mensuel d'une fraction des usagers ou recensement annuel général). Les consignes devront être précisées afin de garantir un traitement standardisé et uniforme d'un certain nombre de situations qui apparaissent peut-être particulières aux yeux de certains moniteurs mais qui s'avèrent en même temps récurrentes dans les institutions.

6. POSSIBILITES D'EXPLOITATION DES GRILLES ARBA

6.1 Possibilités de comparaisons inter et intra-institutionnelles

Il est possible avec la grille ARBA de dresser le profil de chaque institution. Celui-ci permet de rendre compte de la nature et de la quantité des prestations offertes aux personnes handicapées qui y travaillent. A partir de tels profils, il devient par exemple possible de comparer, dimension par dimension, les scores obtenus dans diverses institutions, ou de mettre en lien les scores d'une institution avec les scores moyens d'institutions du même type voire avec des scores moyens cantonaux (cf. figure 15).

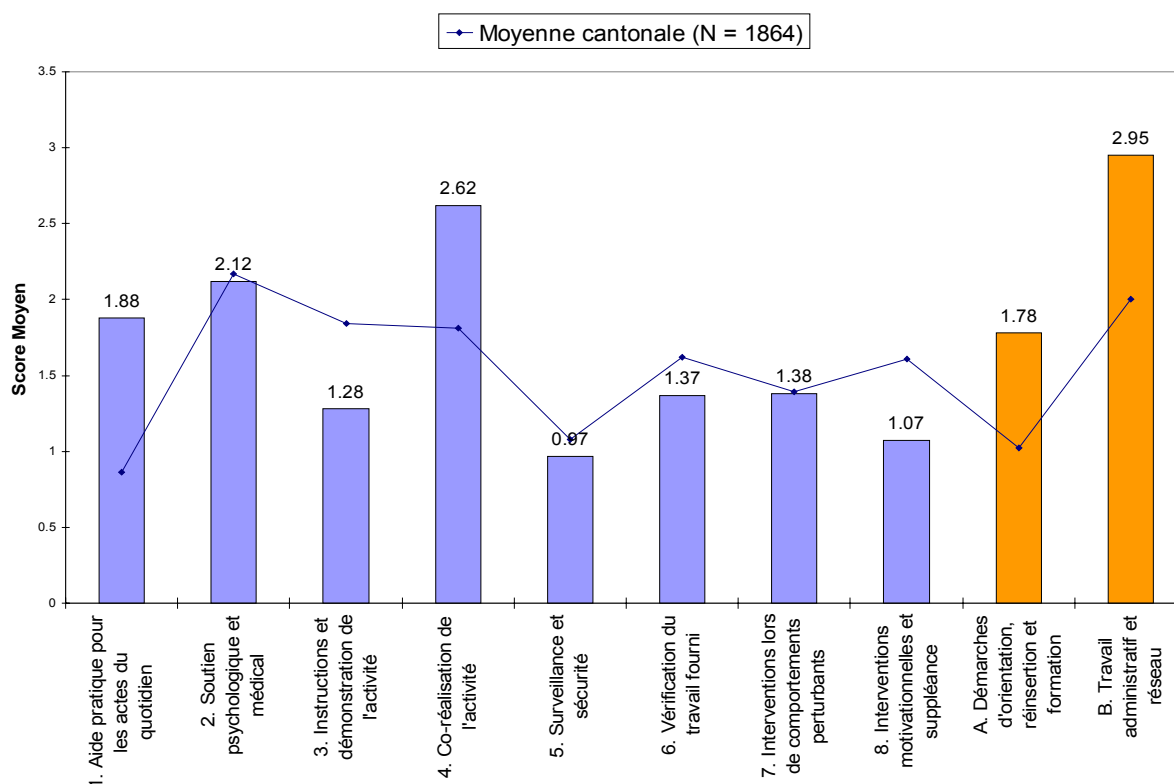


Figure 15 : exemple du profil d'une structure comparé aux résultats moyens cantonaux

Sur la figure 16 sont présentés les scores moyens des dix items constituant le profil du Brico-CES Service (score moyen encadrement direct = 9.67) mis côte à côte avec ceux du GRAAP (score moyen d'encadrement direct = 10.27). Ce procédé offre la possibilité d'identifier les spécificités fines de différentes structures qui peuvent par ailleurs présenter des scores moyens très proches (9.67 points pour Brico-CES et 10.27 points pour le GRAAP). En l'occurrence on s'aperçoit dans l'exemple cité

qu'un effort particulier est mis au GRAAP sur le soutien psychologique et médical des usagers, sur les interventions motivationnelles et le travail administratif, alors que l'encadrement du Brico-CES met davantage l'accent sur l'aide fournie à la place de travail (vérification du travail, instructions, surveillance et sécurité) ainsi que sur l'aide pratique pour les actes du quotidien.

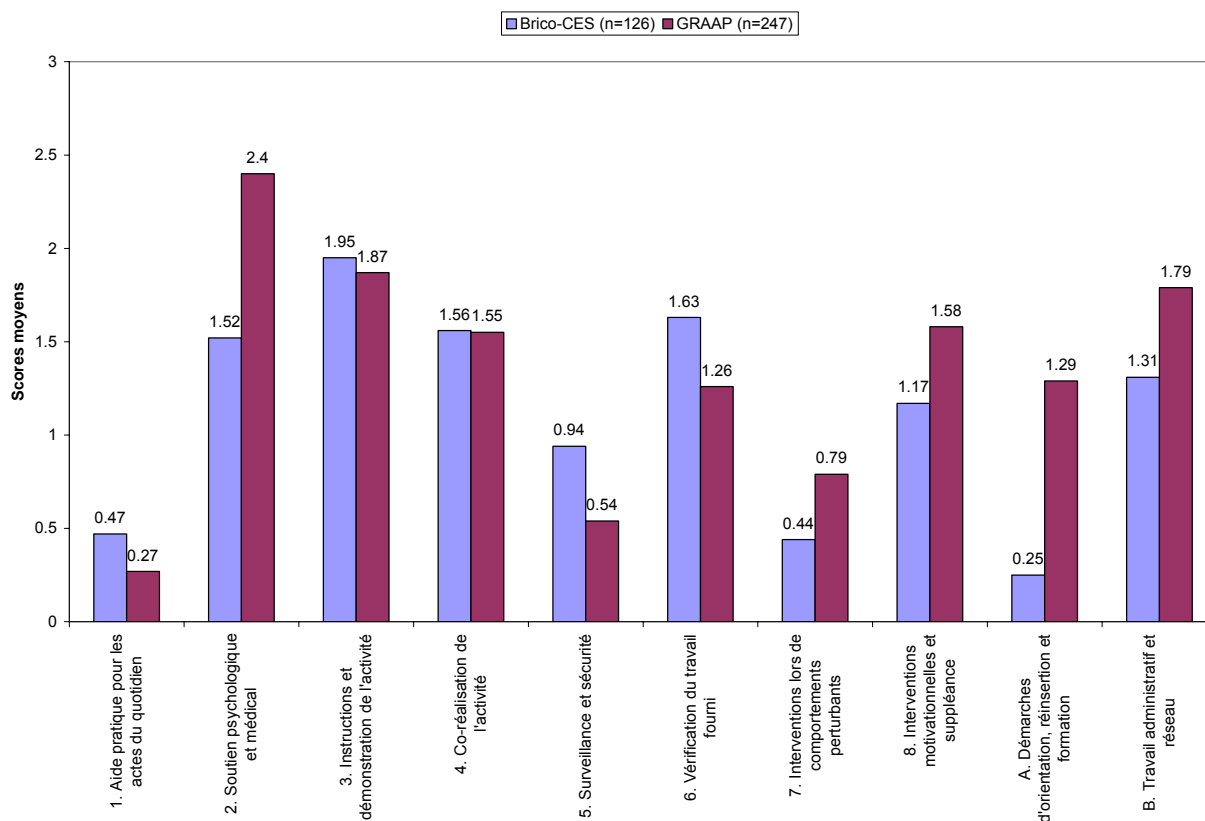


Figure 16 : comparaison entre profils de deux institutions (Brico-CES et GRAAP)

Si les profils spécifiques dressés sur les dix items permettent de comparer entre elles deux à trois institutions au maximum (pour garder une certaine lisibilité des graphiques), l'usage des scores d'encadrement direct (ARBA-ED) ou des prestations administratives, d'orientation et de réseau (ARBA-AOR), permet quant à lui des comparaisons entre un nombre plus important d'institutions. Les comparaisons multi-institutions peuvent aussi être conduites sur un item particulier. A la figure 17 nous pouvons observer une telle comparaison visuelle directe des résultats obtenus à l'item « aide pratique pour les actes du quotidien » dans les quinze institutions ayant participé à notre étude.

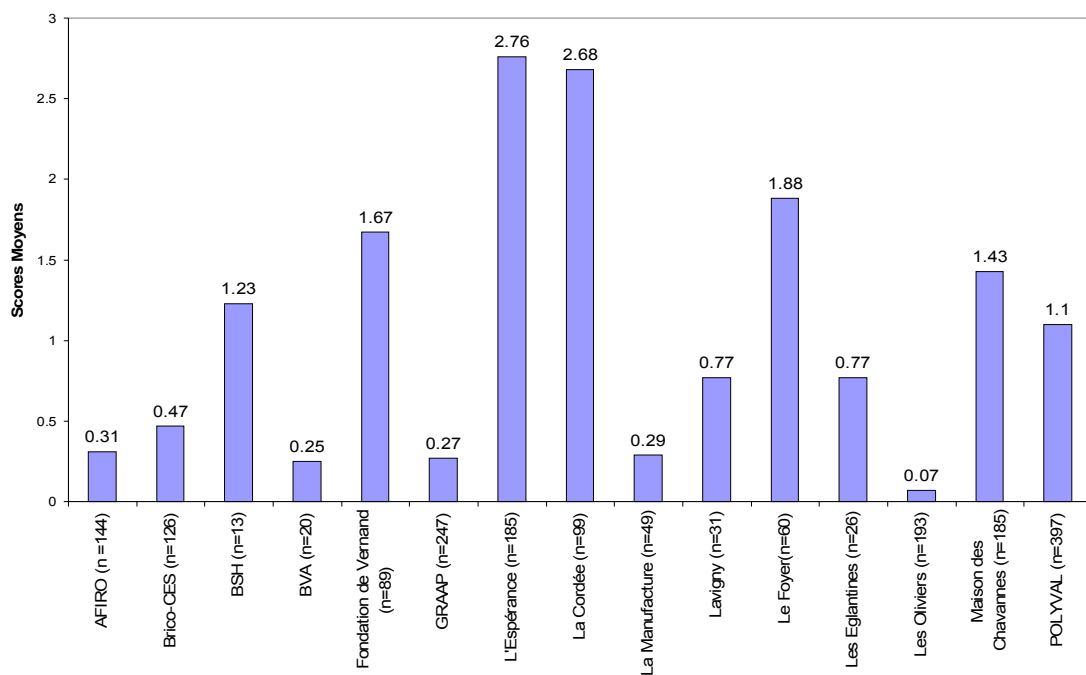


Figure 17 : comparaisons des quinze institutions de l'étude sur l'item « aide pratique pour les actes du quotidien »

Responsables d'institutions et décideurs politiques peuvent aussi avoir l'intérêt de documenter l'évolution temporelle des prestations fournies par un ou plusieurs ateliers donnés.

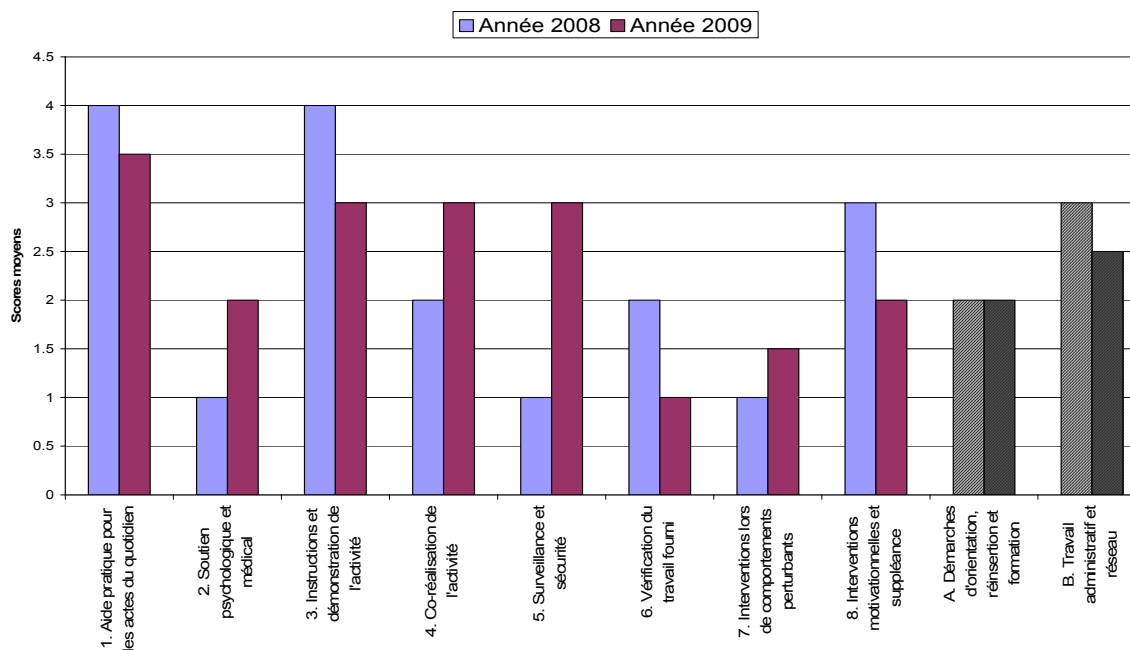


Figure 18 : comparaison temporelle des scores moyens obtenus sur les dix items de la grille ARBA-2 d'une institution (données fictives)

La figure 18 présente à titre illustratif des données fictives permettant de retracer les changements à un an au niveau des scores de chacun des dix items de la grille. Ici aussi, l'usage des profils complets ou l'appui sur les scores ED ou AOR peuvent être envisagés comme sources des mesures.

6.2 Emploi d'ARBA-2 en milieu naturel

L'utilisation de toute grille à une large échelle implique une logistique adéquate. La compilation d'un nombre important de grilles nécessite une méthode et des outils adaptés. L'expérience accumulée lors du traitement des grilles de la présente étude nous a montré que la centralisation des données réalisée à partir de bases de données indépendantes issues des institutions impliquées simplifie considérablement le travail de mise en commun des résultats. Par ailleurs l'implémentation de certaines procédures de contrôle automatisées nous a aidés à détecter les éventuelles omissions ou erreurs. Elles peuvent se révéler extrêmement précieuses, et ce dès la saisie des données.

Une réflexion doit également être menée concernant le protocole d'utilisation régulière de la grille à l'échelle du canton. Si l'on désire évaluer tous les usagers fréquentant un atelier protégé, il s'agit alors d'une procédure de recensement. Si l'on ne désire évaluer qu'une partie représentative des usagers, il s'agit d'un échantillonnage. Un recensement complet nécessite l'évaluation de chaque usager au moins une fois (par exemple à l'entrée ou à la sortie) mais le plus souvent plusieurs fois (1x/an, entrée, sortie). Le recensement exige la tenue disciplinée d'un calendrier d'évaluation et des choix clairs quant à la fréquence des évaluations.

Si le protocole adopté est un échantillonnage, il faut déterminer les modalités de sélection de l'échantillon qui sera représentatif de la population dans l'atelier. Divers choix existent : (1) échantillons de temps (toute ou une partie de la population d'un atelier est évaluée à plusieurs reprises dans l'année), (2) échantillons aléatoires simples (sélection aléatoire des usagers dans les ateliers), (3) échantillons aléatoires stratifiés (choix d'échantillons qui respectent certaines proportions de la population d'usagers fréquentant l'atelier), (4) échantillons sélectionnés sur la base de critères prédéfinis (échantillons non aléatoires).

A priori le recensement exige davantage de ressources et paraît plus lourd à conduire. Il a par contre le mérite de l'exhaustivité et intègre pleinement la procédure d'évaluation dans le travail routinier des moniteurs. L'échantillonnage est peut-être moins exigeant, mais il se révèle souvent plus complexe à programmer et nécessite en général un coordinateur qui veille au respect des stratégies de sélection adoptées. Selon le type d'échantillonnage choisi, il peut conduire à des périodes creuses suivies de périodes plus intenses sur le plan des évaluations (et donc du travail à accomplir). Ces déséquilibres conduisent parfois à des surcharges dont l'issue est souvent une moins bonne qualité des évaluations fournies.

6.3 Modèle quadrifactoriel du besoin d'aide en atelier protégé

L'étude des grilles ARBA nous a poussés à conduire une réflexion théorique sur la question de l'évaluation des besoins d'encadrement. La figure 19 en retrace schématiquement les grandes lignes. Les usagers d'ateliers protégés présentent chacun un profil particulier de compétences et déficits. Les tâches qui sont confiées à ces personnes sont avant tout liées aux spécificités de l'atelier dans lequel elles évoluent. Ces activités peuvent s'avérer plus ou moins complexes et difficiles ; elles exigent de la part de ceux qui les accomplissent des habiletés plus ou moins grandes à différents niveaux (aptitudes manuelles et intellectuelles, degré d'autonomie, capacités de résolution de problèmes, etc.). De l'interaction de ces deux composantes (compétences personnelles - difficulté des tâches) résulte une performance (cf. ①). Dans leur travail d'encadrement, les moniteurs comparent le degré de réussite des usagers dans les tâches qui leur ont été confiées à la performance souhaitée (cf. ②). Lorsqu'un usager n'est pas en mesure d'atteindre la performance attendue, les moniteurs peuvent faire le choix de renforcer l'aide fournie (cf. ③), d'adapter la complexité du travail, c'est-à-dire de modifier les tâches proposées (cf. ④), ou même de rediriger l'utilisateur vers un autre atelier. Ces modifications affectent et modifient naturellement les conditions antérieures, la nature et la difficulté des tâches (cf. ⑤). Dès lors les moniteurs repartent vers une nouvelle évaluation suivie d'éventuels réajustements, recommençant en quelque sorte le cycle d'évaluation et de soutien décrit ci-dessus.

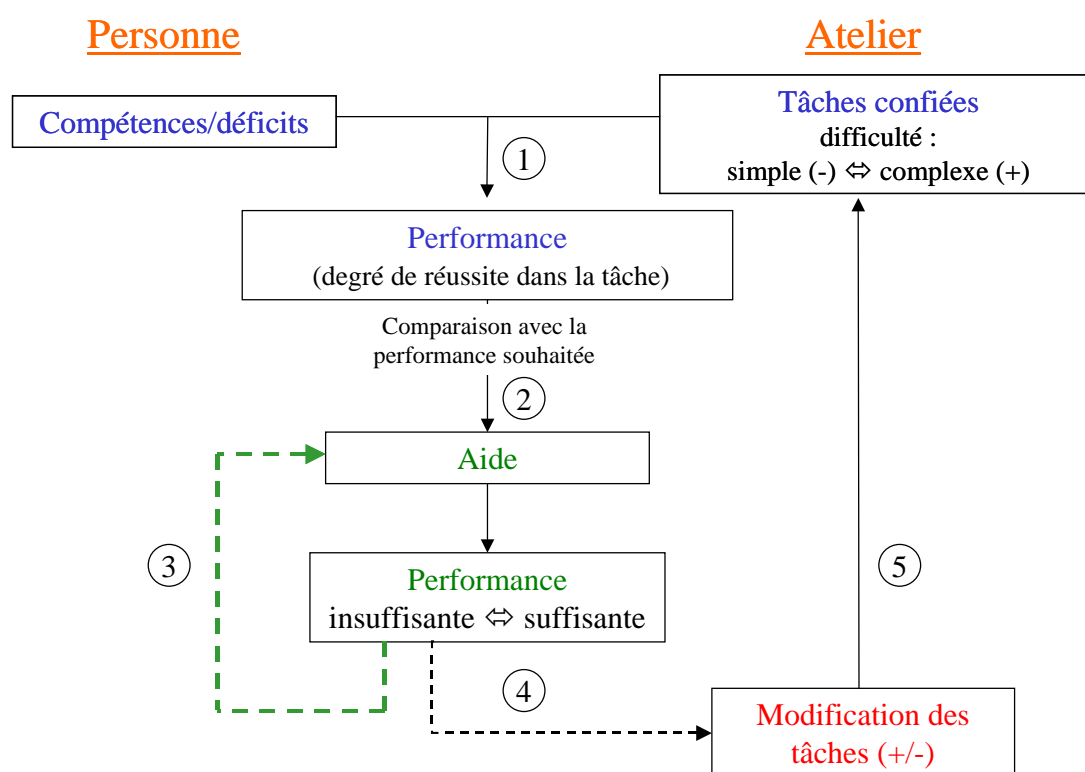


Figure 19 : modèle synthétique du besoin d'aide en atelier protégé ; les chiffres se rapportent à ceux cités dans le texte

Il est important de noter que des usagers disposant de faibles ressources peuvent manifester une certaine autonomie si on leur attribue des tâches relativement simples. Des usagers bénéficiant au contraire d'un niveau de fonctionnement plus élevé peuvent se voir confier des tâches plus complexes et exigeantes ; ceci peut les conduire à avoir besoin d'un encadrement plus soutenu. Il est important de noter qu'un décalage entre performance souhaitée et performance réelle ne débouche pas nécessairement sur un renforcement de l'aide : il peut conduire à un réaménagement des tâches confiées qui réduit le besoin d'encadrement. Dans un constant exercice de réévaluation, les moniteurs d'atelier ont pour objectif de permettre à chacun de fonctionner au niveau optimal de ses possibilités, compte tenu des activités à exécuter et des possibilités réelles d'encadrement. L'évaluation de l'aide fournie est donc un processus dynamique dont le résultat varie en fonction de paramètres extérieurs à l'utilisateur lui-même.

La principale conséquence est que l'évaluation des seules compétences ou déficits ne permet pas de déduire de façon directe l'aide à apporter. Il nous semble donc nécessaire de réinscrire le besoin d'aide au sein d'un modèle relatif plus complet intégrant quatre dimensions : (1) le niveau de compétences de l'utilisateur, (2) la complexité de la tâche, (3) la performance observée, (4) la performance souhaitée (v. figure 20).

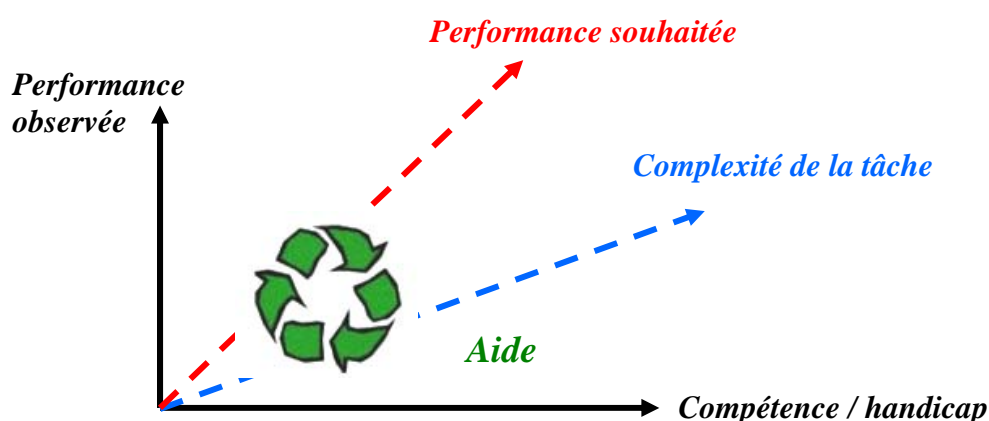


Figure 20 : modèle quadrifactoriel du besoin d'aide en atelier protégé

Comme l'illustre la figure 20, l'aide fournie correspond à une mesure résultant des quatre facteurs précités. ARBA-2 cherche à objectiver cette mesure, notamment à travers son score d'encadrement direct, mais ne constitue en rien une évaluation de ces quatre facteurs. Les résultats obtenus via ARBA-ED représentent donc une mesure de type secondaire. On pourrait imaginer la calculer si l'on disposait d'une équation mathématique valide articulant les quatre paramètres du modèle ainsi que des mesures fiables de chacun d'eux, chose aujourd'hui évidemment impossible.

6.4 Limites de la grille

Malgré les indubitables progrès réalisés par rapport à la version originale, ARBA-2 n'est pas exempte de limites ou de défauts, lesquels sont toutefois pour l'essentiel inhérents à la méthode choisie. Les évaluations reflètent de façon indirecte le fonctionnement des usagers à un moment donné et dans un atelier particulier. Divers facteurs comme l'état de santé des usagers ou la charge de travail dans l'atelier peuvent influencer et faire varier les résultats. Comme nous venons de l'évoquer, ARBA-2 ne mesure pas ces aspects. La grille ne peut donc pas fournir les moyens de son interprétation, en particulier les raisons ou les causes des scores obtenus.

Formation des moniteurs

En leur qualité d'évaluateurs, les moniteurs d'atelier forment le chaînon central dans le déploiement et l'utilisation de la grille ARBA à large échelle. Il donc est primordial que ces professionnels puissent bénéficier d'une formation de base à l'utilisation de la grille. La question du maintien des connaissances et de la formation des nouveaux collaborateurs doit également être posée. Malgré sa simplicité apparente, ARBA nécessite selon nous une formation initiale (1-2h) qui peut être complétée par une assistance sur le terrain en cas de questions spécifiques.

Capacité de discrimination

Pour compléter les grilles, les évaluateurs vont inévitablement devoir faire appel à diverses capacités cognitives, en particulier leur mémoire. Ces évaluations peuvent logiquement être affectées par différents biais, comme une perception altérée des durées passées au contact des usagers ou une remémoration à caractère sélectif des prestations dispensées. Nous n'avons pas conduit d'évaluation de la fidélité inter-juge d'ARBA-2, partant du principe qu'il vaut sans doute mieux que la grille fasse d'emblée l'objet d'un consensus entre moniteurs fréquentant le même atelier. La recherche d'un consensus entre moniteurs pouvant être au contact des mêmes usagers nous a paru être non seulement une solution plus simple à mettre en œuvre, mais aussi une formule d'emblée plus fiable. En effet plusieurs moniteurs peuvent lors d'une journée type être au contact d'un usager donné et participer à son encadrement. Les tendances d'un moniteur à surestimer ou sous-estimer le temps dédié à l'encadrement d'un usager sur tel ou tel aspect mesuré par ARBA-2 peuvent être ainsi compensées par le regard d'un second ou d'un troisième moniteur ayant également affaire avec le même usager.

Risques de surévaluation systématique

Les évaluateurs sont dans une très large mesure des salariés au service d'une institution publique ou privée. Ils pourraient manifester une certaine inquiétude quant à l'exploitation des scores ARBA. Cela

pourrait occasionnellement les amener à surévaluer les aides fournies. Ce risque augmente quand les moniteurs réalisent seuls des évaluations sans faire appel à un consensus inter-juges, ou quand des choix de cotation sont décidés a priori au sein d'une institution, puis imposés aux évaluateurs sans tenir compte des observations réelles du terrain. Par exemple l'activité de vérification du travail fourni pourrait être systématiquement cotée à 7 pour tous les usagers d'un atelier donné, parce que cet atelier nécessite en principe tel degré de contrôle de la production, sans vérifier si ce principe est effectivement appliqué ou si certains usagers mériteraient d'être cotés avec des scores inférieurs ou supérieurs. Il est évidemment impossible avec ce type de grille d'empêcher ou de prévenir des dérapages de cette nature. Néanmoins les diverses possibilités de comparaison de profils entre institutions ou avec des moyennes choisies (p. ex. cantonales) devrait permettre aux récipiendaires des résultats de repérer plus facilement des scores inattendus et de s'en étonner.

Erreur de mesure et interprétation des scores

Le score de chaque item, du fait du découpage du temps en intervalles, sera associé d'une marge d'erreur pouvant aller de zéro à quatre minutes pour les items 1 à 8, et de zéro à vingt-neuf minutes pour les items A et B. Un évaluateur cotera par exemple quinze minutes (3 points) une intervention qui en aura coûté en réalité dix-huit. Un autre évaluateur (ou le même évaluateur sur un autre item) pourrait par contre coter 4 points (soit vingt minutes d'encadrement) les mêmes dix-huit minutes effectives. La capacité de discrimination de l'outil étant d'un point, soit un intervalle de cinq minutes par item, une erreur se situant à cinq minutes est tout à fait tolérable, et sans doute inévitable aussi, compte tenu des capacités attentionnelle et mnésiques des moniteurs eux-mêmes. Mais si l'on considère à présent le score d'encadrement direct d'ARBA-2, on se rend immédiatement compte que ce score correspond à une somme de huit items. Ce score peut donc être entaché au final d'un erreur beaucoup plus importante, car l'on additionne les erreurs de mesure qui sont associées à chaque item. Si l'on poursuit ce raisonnement, le score ED pourrait être entaché d'une erreur de mesure totale maximale de quarante minutes (par addition des huit erreurs de cinq minutes possibles sur chaque item), ce qui est là une erreur de mesure trop importante et peu acceptable.

En multipliant ainsi les points, on augmente la marge d'erreur liée au mode d'interprétation des scores, c'est-à-dire à la transformation immédiate des points ARBA-2 en minutes d'intervention. Dans certains cas particuliers ou spécifiques, nous n'excluons pas la possibilité d'avoir un décalage entre un minutage ARBA et une estimation personnelle d'un moniteur. Seul le recours à un véritable chronométrage de l'activité des moniteurs pourrait éliminer un tant soit peu ce risque.

Nous avons heureusement pu démontrer que les estimations du temps total passé à encadrer un patient obtenues via les scores d'ARBA-2 correspondent bien aux estimations personnelles subjectives des moniteurs. Ceci peut s'expliquer par le fait que très souvent les erreurs ont statistiquement tendance à s'annuler, c'est-à-dire que les erreurs de surestimation compensent à la longue les erreurs de sous-estimation. C'est une observation classique en statistique : les erreurs

dites aléatoires ont tendance à s'annuler pour parvenir à une moyenne proche de zéro. On peut en déduire qu'ARBA-2 est une mesure plus fiable du degré d'encadrement d'un atelier ou d'une institution que d'un usager individuel, puisque lors de mesures individuelles le phénomène de compensation statistique des erreurs aléatoires ne peut se produire (contrairement à la situation où on applique l'outil à un nombre important d'utilisateurs). Pour terminer, nous parlons ici d'erreurs aléatoires de mesure et non pas d'erreurs systématiques qui pourraient être dues par exemple à l'interprétation erronée d'une consigne. Le seul moyen de se préserver de ce type d'erreur est d'assurer une bonne formation des évaluateurs, de faire appel à des consensus entre moniteurs accompagnant le même usager, ou de réaliser régulièrement des mesures de fidélité inter-juges pour s'assurer de la cohérence dans l'utilisation de l'outil par l'ensemble des moniteurs impliqués.

Les erreurs dont nous avons parlé ici sont intimement liées au fait que l'on transforme des scores en minutes, chaque score correspondant à un intervalle de minutes et non à une mesure précise du temps. On ne doit donc jamais oublier cela dans la manipulation et l'interprétation des scores ARBA, même si un réel effort a été accompli pour favoriser la meilleure correspondance possible entre un score et la signification de celui-ci en termes de durée effective d'encadrement. Dans le même ordre d'idées, il ne faut pas oublier enfin que le point ARBA d'encadrement direct n'équivaut pas au point ARBA des prestations administratives, d'orientation et de réseau, ce dernier étant plus lourd puisque couvrant un intervalle de trente minutes (et non pas cinq). C'est la raison pour laquelle ARBA-2 fournit deux scores distincts et non plus un seul.

7. CONCLUSION

Les modifications majeures résultant du passage de la grille ARBA initiale à la grille ARBA-2 finalisée sont résumées au tableau 17. L'adoption d'un modèle théorique du besoin d'aide permet de situer de façon claire et intelligible les concepts visés par la grille ARBA-2. Les évaluations fournies par cet instrument sont clairement dissociées d'une mesure des déficits / habiletés des usagers, évitant ainsi un raccourci erroné qui ferait correspondre un certain degré de déficit avec un niveau plus ou moins élevé d'encadrement effectif. Comme le modèle du besoin d'aide met en relation dynamique au moins quatre facteurs, il serait possible dans l'idéal de déduire le besoin d'aide à partir de la mesure de ces quatre paramètres. Mais aujourd'hui l'option retenue avec ARBA-2 de mesurer directement l'aide fournie (mesure secondaire) est sans doute la plus réaliste car elle permet d'échapper à cette multiplicité de mesures primaires dont nous n'avons pas le premier outil à disposition.

	ARBA	ARBA-2
Modèle théorique	-	modèle quadrifactoriel du besoin d'aide
Concepts mesurés	autonomie + aide fournie	aide fournie
Découpage théorique de la grille	a priori	empiriquement fondé (analyse factorielle)
Nombre d'items	28	10
Système de cotation	fréquence des interventions	durée cumulée des interventions
Différenciation des prestations dispensées en l'absence des usagers	non	oui
Scores	1 total + 4 sous-totaux	2 totaux + possibilité d'établir des profils institutionnels détaillés (par item)
Informatisation de l'outil	simple	simple et permettant le regroupement d'importants volumes de données sous forme de tableaux

Tableau 17 : éléments de comparaison entre ARBA et ARBA-2

Le découpage théorique retenu pour la grille ARBA-2 s'appuie sur des dimensions empiriquement fondées via des analyses factorielles en composantes principales, qui sont des analyses exploratoires. En effet, le découpage *a priori* de la grille ARBA originale ne correspondait pas aux dimensions issues de l'analyse factorielle. La nouvelle grille propose un découpage compatible avec

les résultats de l'analyse factorielle. Une analyse factorielle confirmatoire pourrait maintenant être conduite qui certainement vérifierait les résultats obtenus auparavant. Le manque de temps et des logiciels statistiques permettant la réalisation de ce type d'analyse nous ont conduit à renoncer à cette option. Elle pourrait éventuellement être reprise ultérieurement.

Nous avons réduit le nombre d'items de vingt-huit à dix. La perte d'information a toutefois été minimisée puisque nous avons repris dans une très large mesure le contenu de la grille initiale grâce à un système de regroupement des items fortement corrélés entre eux. Ainsi, plutôt que d'élimination de l'information par réduction du nombre d'items, nous pouvons parler d'une meilleure synthèse de cette information à travers un travail de reformulation des items qui a permis de mieux cibler les dimensions d'intérêt. ARBA-2 a perdu une grande partie du détail des vingt-huit items d'ARBA ce qu'elle a gagné en simplicité et en synthèse. De plus, par rapport à l'ancienne formule, la structure factorielle empirique de la nouvelle grille colle davantage à la répartition des items qu'elle propose. Elle pourrait cependant être améliorée. Quatre dimensions se dessinent : le soutien psychosocial, l'encadrement au poste de travail, l'aide pour les problèmes d'autonomie pratique, et les tâches administratives, d'orientation professionnelle et de réseau. L'instrument présente encore un certain déséquilibre sur les deux dernières dimensions, car elles sont composées d'un nombre trop faible d'items. L'élaboration de trois ou quatre nouveaux items permettrait peut-être de rétablir un certain équilibre à ce niveau. Mais cet allongement risque de sacrifier la concision actuelle de la grille à une meilleure qualité structurelle.

Le système de cotation de la grille ARBA ne mesurait que la fréquence des interventions, ce qui apparaissait comme une solution peu satisfaisante en raison de son caractère trop simplificateur. Le système de cotation adopté pour ARBA-2 prend en compte la durée cumulée des aides fournies. Les scores ARBA-2 peuvent être convertis en durées d'encadrement exprimées en minutes. Cette démarche mérite néanmoins d'être conduite avec une certaine précaution, du fait des marges d'erreur qui peuvent s'additionner dans le calcul des scores. Ceci dit, les scores globaux ARBA d'encadrement direct, même transformés en minutes, se sont révélés en moyenne tout à fait similaires aux estimations personnelles subjectives des moniteurs quant au temps d'encadrement total dédié à leurs usagers. On peut donc penser que plus une institution dispose de mesures ARBA-2 de ses usagers, plus les résultats ont une bonne probabilité d'être fidèles à la réalité ; pour les profils plus individuels ou dans des ateliers très petits, les risques d'erreurs apparaissent par contre plus grands, du fait que les compensations statistiques des éventuelles erreurs de mesure ne peuvent pas se produire.

ARBA-2 propose très clairement deux scores indépendants : un score d'encadrement direct (ARBA-ED) et un score de prestations administratives, d'orientation et de réseau (ARBA-AOR). Les aides au poste de travail ou dans l'atelier sont comptabilisées de façon indépendante des prestations dispensées en l'absence des usagers ou en dehors de l'atelier. Cette façon de faire est plus naturelle compte tenu du mode de fonctionnement des ateliers. Elle facilite l'usage de la grille car s'adapte mieux aux repères temporels habituels qui varient en fonction du type d'encadrement fourni. En effet

si des mesures par tranches de cinq minutes au cours d'une journée type conviennent pour évaluer la durée d'encadrement dédié à la surveillance, au soutien psychologique ou à des instructions concernant l'activité à réaliser, des mesures par intervalles d'une demi-heure sur un mois type sont plus pertinentes pour coter les prestations administratives, d'orientation et de réseau consacrées à un usager.

Les scores issus de la grille ARBA initiale étaient difficilement interprétables. En effet, le score total réunissait deux types de mesures différentes au sein d'un seul et même indice. Le nouveau découpage en dix items permet d'établir des profils institutionnels ou personnels détaillés permettant une lecture et une interprétation plus fine des résultats. Grâce à cette simplification du nombre d'items la nouvelle grille semble paradoxalement offrir des possibilités plus importantes encore dans l'usage de ses scores et de ses profils, notamment à des fins comparatives : la comparaison de profils sur vingt-huit items s'avère beaucoup plus difficile à manier que celle sur dix items.

Enfin, si l'informatisation simple dont bénéficiait la grille ARBA satisfaisait déjà ses utilisateurs, il a été possible d'y ajouter des procédures permettant le regroupement automatisé d'importants volumes de données, minimisant également les données manquantes ou erronées.

En résumé, nous pensons avoir atteint avec ARBA-2 les objectifs initialement fixés. Cet instrument paraît satisfaire au cahier des charges que nous lui avons fixé et est aujourd'hui prêt à être employé à plus large échelle dans une routine institutionnelle qui mérite cependant d'être précisément définie.

8. Références

Direction Générale de l'Action Sociale du Département de la Solidarité et de l'Emploi (DSE-DGAS) (2006). Analyse des ressources et besoins d'aide dans les établissements accueillant des personnes handicapées (EPH) (ARBA-Atelier) - mode d'emploi détaillé. Récupéré le 5 août 2008 à l'adresse http://www.geneve.ch/handicap/doc/arba_me_atelier_05.pdf.

Pomini, V., Golay, P., Blanc, M., & Pillonel, D. (2008, juin) *Analyse des Ressources et Besoins d'Aide dans les établissements accueillant des personnes handicapées*. Communication affichée présentée aux XIe Journées de recherche des Départements de Psychiatrie de Lausanne et de Genève, 12 juin 2008.

Pomini, V., & Fernandez, S. (2006). *Grilles ARBA d'analyse des ressources et besoins d'aide. Rapport d'évaluation psychométrique* (Rapport). Prilly : Service de Psychiatrie Communautaire, Unité de Réhabilitation.

Fédération des Ateliers pour personnes Handicapées (FAH) (2006). *Rapport d'activités 2005*. Récupéré le 5 août 2008 à l'adresse http://www.insos-fah-vd.ch/telechargements/rapport_activites_2005.pdf.

Fédération des Ateliers pour personnes Handicapées (FAH) (2007). *Rapport d'activités 2005*. Récupéré le 5 août 2008 à l'adresse http://www.insos-fah-vd.ch/telechargements/FAH_2006.pdf.

Références en ligne :

Site officiel de l'Etat de Genève / Politique du handicap / ARBA

Adresse : <http://www.geneve.ch/handicap/politique/arba.htm> (consulté le 04.07.08)