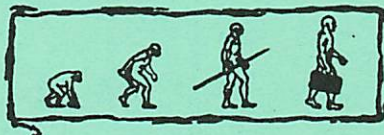


Actes du 14^e colloque de l'AQPC

ÉVALUATION ! ÉVOLUTION ?



Où s'en va le collégial ?

6D8

L'évaluation en stage : dure réalité !

par
CANTIN, Gilles, DELISLE, Suzanne
professeurs
Collège de Saint-Jérôme



Association québécoise
de pédagogie collégiale

L'évaluation en stage: dure réalité!

Suzanne Delisle
Gilles Cantin
Professeurs au Cégep de Saint-Jérôme

INTRODUCTION

Les stages de formation offrent aux étudiants des occasions uniques d'initiation pratique à un domaine professionnel. Parallèlement, ils sont des occasions pour les professeurs-superviseurs d'évaluer le niveau de maîtrise des habiletés des étudiants avant leur entrée sur le marché du travail.

Paradoxalement, ces éléments essentiels de la formation causent d'énormes problèmes lors de l'évaluation. En effet, quelle est la meilleure façon d'évaluer justement les performances d'étudiants qui sont assignés à des milieux de stage différents? Car, comme le dit si bien Pelpel (1989), «au-delà des objectifs communs, de l'homogénéité théorique des terrains d'accueil, chaque stagiaire va rencontrer des conditions particulières, des personnalités différentes, des événements imprévus». Une étude américaine portant sur une série d'études reliées aux stages en enseignement a fait ressortir cinquante variables possibles dans l'analyse du phénomène allant de la complexité des relations entre les individus impliqués à la conception que se font les intervenants de la réussite d'un stage (Applegate, 1987, dans Zimpher, 1989).

Ainsi, si l'on en juge par les écrits consultés et les témoignages des personnes impliquées, l'évaluation des apprentissages en stage apparaît comme une tâche délicate pour les enseignants responsables de la supervision de stagiaires. Délicate, puisque porter un jugement à la fin d'une formation est toujours difficile surtout lorsqu'il doit être posé à la suite d'une démarche où les partenaires étaient impliqués pédagogiquement et même affectivement.

Dans la réalité actuelle, les superviseurs sont la plupart du temps laissés à leur jugement professionnel, bien souvent sans outils valides pour les seconder dans leur évaluation. Cet état de fait provoque des évaluations et des notations pouvant être très différentes d'un superviseur à l'autre pour un même type de clientèle.

De plus, les superviseurs en stage utilisent des outils d'évaluation qu'ils ont construits eux-mêmes et dont la validité est peu établie. À défaut d'une méthode d'évaluation appropriée, la notation du stage est souvent empreinte de subjectivité. La plupart du temps, on choisit des outils où l'on commente de façon générale la performance des stagiaires et pour laquelle on attribue une note entre l'excellence et l'échec sans trop savoir comment nuancer les notes. Bernard et Fontaine rapportent que les démarches des concepteurs sont très souvent circulaires, «on planifie plus ou moins les transformations : on modifie une grille pour revenir à l'ancienne grille quelques années plus tard » (1989, p. 35). Cette situation est facilement explicable, comme le soulignent De Bal *et al* (1976) par le fait que les professeurs disposent rarement du temps nécessaire pour construire des outils d'évaluation et encore moins de temps pour les valider. Bref, les grilles d'évaluation sont perçues comme «une technique de mesure généralement très dispendieuse sous les aspects du temps, de l'énergie et de l'argent» (Morissette, 1984, p.263).

À LA RECHERCHE D'OUTILS PLUS PRÉCIS POUR ÉVALUER EN STAGE

Notre recherche entreprise en 1990¹ avait pour objectif de concevoir des outils d'évaluation en stage permettant d'apprécier la performance des stagiaires sur des faits observables et précis afin d'éviter le plus possible que les utilisateurs appuient leur jugement sur des impressions ressenties.

¹*Évaluation en stage: un modèle de grilles*, recherche réalisée par Gilles Cantin et Suzanne Delisle du département TESH du Cégep de Saint-Jérôme. Elle a été subventionnée par PAREA.

Deux pratiques en évaluation d'étudiants en stage : scientifique et naturaliste

Selon Caruso et Fawcett (1986), les pratiques actuelles d'évaluation dans des milieux de travail découlent de deux principaux points de vue en matière de supervision. L'un est caractérisé par un cadre de référence scientifique et précis, l'autre par un cadre plus qualitatif, plus général. Voyons brièvement ces deux pratiques d'évaluation.

L'approche scientifique met l'accent sur l'objectivité et la précision de la mesure des comportements attendus chez l'individu en formation. Les outils les plus couramment utilisés par les superviseurs sont des échelles d'appréciation (*rating scales*) basées sur des comportements spécifiques. L'objectif premier de ces concepteurs d'outils d'évaluation est de déterminer les critères précis à partir desquels les stagiaires seront évalués. Ces critères sont généralement issus des recherches qui décrivent les comportements et les qualités d'un professionnel efficace. On assume alors que les critères ainsi identifiés permettront d'indiquer le niveau de performance des individus. Cette pratique, selon Acheson et Gall (1992) a l'avantage de faire connaître au tout début du stage les critères précis sur lesquels les stagiaires sont évalués, ce qui contribue à faire diminuer sensiblement le stress relié à l'évaluation.

L'approche naturaliste (*naturalistic*), quant à elle, laisse beaucoup plus de place aux perceptions générales des évaluateurs. Au lieu de faire référence aux compétences et aux comportements précis pour mesurer la performance d'un individu, le superviseur se fie aux perceptions qui se dégagent des pratiques pédagogiques du stagiaire. Les tenants de cette approche utilisent des outils d'évaluation qui permettent une appréciation narrative plus globale des situations qu'ils observent, comme le résumé d'évaluation par exemple. Cette approche possède par contre un inconvénient : comme les outils ne sont pas aussi descriptifs et détaillés que ceux utilisés dans l'approche scientifique, les utilisateurs doivent posséder une grande capacité à décrire correctement les faits et à consigner des anecdotes qui soient significatives. De plus, ils doivent posséder une habileté à synthétiser les faits et gestes afin de dégager une appréciation globale et juste.

Selon les auteurs, l'une et l'autre de ces pratiques possèdent des points forts et des points faibles. Il est donc avantageux, comme ils le proposent, de combiner les meilleurs éléments des deux approches pour établir les bases de l'évaluation en stage. Même si nous avons retenu l'approche scientifique pour élaborer nos outils d'évaluation, nous avons aussi prévu utiliser l'approche naturaliste au moment de la notation pour compléter les informations provenant des grilles plus analytiques.

Deux méthodes en évaluation de performances complexes : analytique et globale

De ces deux points de vue en matière de supervision et d'évaluation découlent des méthodes pour apprécier et évaluer les performances.

La méthode d'appréciation globale consiste à percevoir une production ou une performance dans son ensemble, de manière globale. La méthode analytique consiste à évaluer un candidat à partir de plusieurs points de vue ou critères indépendants autant que possible les uns des autres. Cette dernière méthode possède l'avantage d'offrir à l'étudiant toutes les indications possibles sur les qualités et les défauts de sa performance, ce qui constitue en quelque sorte un moyen de rétroaction (Scallon, 1988).

Pour cette raison, nous préconisons une méthode analytique lors de la construction des grilles d'évaluation où «les indicateurs ou les descripteurs choisis pour juger d'une performance constituent une rétroaction informative pouvant guider l'étudiant dans les améliorations à apporter ou les démarches correctives à entreprendre pour lui faire atteindre un certain degré de maîtrise» (Scallon, 1988, p. 160). Une méthode plus globale d'évaluation pourra suivre l'exercice d'évaluation analytique, mais elle ne viendra que soutenir à ce moment l'évaluateur dans le jugement qu'il doit prendre quant à la performance du stagiaire.

Conditions de base d'une évaluation plus objective

Bien que nous sachions que l'unanimité entre des évaluateurs est rarement observée dans la réalité (De Bal et al, 1976), il importe de respecter des conditions de base pour favoriser un degré d'accord élevé entre les évaluateurs. La consultation des ouvrages sur l'évaluation et l'expérimentation nous a permis de dégager les grandes lignes de ces conditions.

- *Définition des critères*

Afin de laisser le moins de place possible à l'interprétation et aux désaccords entre les évaluateurs, il importe que les évaluateurs s'entendent sur les critères qui serviront à inférer telle ou telle compétence et l'importance qu'on leur accorde. Ces critères doivent alors être définis très précisément pour éviter les ambiguïtés.

- *Définition des degrés de l'échelle*

Il en est de même pour les échelons de l'appréciation : chacun d'eux doit être défini précisément de façon à ce que chacun des évaluateurs les utilise de la même façon.

- *Consignes précises d'utilisation des grilles*

Les grilles d'évaluation doivent toujours être accompagnées d'un mode d'utilisation qui clarifie les consignes en définissant systématiquement la façon de remplir les grilles et les formulaires d'évaluation finale s'il y a lieu.

Bien que nous soyons conscients qu'il est utopique d'espérer une unanimité parfaite entre des évaluateurs, une recherche de De Bal *et al* (1976) nous confirmait dans notre hypothèse: l'adoption des mêmes critères d'évaluation par des évaluateurs permet d'accroître sensiblement le degré de concordance entre eux.

Des outils précis pour observer

Pour pouvoir se servir de l'observation comme support à l'évaluation en stage, il faut avoir recours à des outils qui permettent de recueillir l'information le plus justement possible. On s'accorde à dire que plus les outils sont précis, plus l'évaluation a de chances d'être précise (Morissette, 1984; Morissette et Gingras, 1989; De Bal *et al*, 1976). Mais la construction d'instruments efficaces ne va pas sans causer d'énormes difficultés aux concepteurs, qui sont bien souvent les superviseurs eux-mêmes réunis en équipe de travail.

En effet, l'identification des gestes significatifs qui permettent d'inférer l'habileté du stagiaire n'est pas simple dans un domaine de performances complexes. L'identification des savoir-faire et des savoir-être est un exercice très difficile quand on sait qu'il est presque impossible de les décrire spontanément. Il faut plutôt se fier à toute une série d'inférences qui sont à leur tour difficiles à établir. Pensons par exemple à la difficulté d'identifier les manifestations de l'attitude d'ouverture face au changement! Combien de gestes et de comportements nous permettent d'inférer cette attitude? Comment les décrire, comment les nommer? Comment apporter des précisions en évitant de créer un grille interminable de nombreux comportements qu'il faut démontrer pour prouver sa compétence, dans un domaine de relations interpersonnelles par exemple? Fortin dit justement que «la compétence professionnelle peut revêtir tellement de formes qu'il est impossible d'en faire une liste exhaustive» (1984, p. 85).

Cela est d'autant plus difficile qu'il semble impossible de prévoir tous les gestes associés aux situations. Intervenir auprès des enfants, par exemple, est une acte global, intuitif et complexe et, en conséquence, exige de l'éducateur qu'il réinvente constamment les gestes à poser (Paré et Laferrière, 1985). L'éducateur doit s'ajuster continuellement à la réalité changeante, il doit créer. Ces particularités font ressortir la complexité du travail de l'éducateur et la difficulté de s'entendre sur les caractéristiques d'un bon enseignant ou d'un bon éducateur lorsque vient le temps de construire des outils d'évaluation.

Malgré toutes ces difficultés, nous avons quand même tenté d'identifier les indices et les critères à partir desquels les superviseurs pourraient donner une appréciation plus objective de la performance des stagiaires.

Les compétences: base de la construction des outils

L'approche par compétences a été retenue comme principe organisateur des stages et de l'évaluation car elle s'avérait la base la plus adéquate pour cerner la complexité des habiletés attendues au terme des stages. L'approche par compétences présente l'avantage d'être globale et d'intégrer simultanément des connaissances, des attitudes et des habiletés. À l'intérieur d'une telle approche, les objectifs de formation sont remplacés par des résultats d'apprentissage qui sont directement reliés à la profession (Tremblay, 1990). Cette précision des résultats permet aux évaluateurs et aux stagiaires d'avoir une idée plus concrète des exigences de la profession et par conséquent de clarifier les attentes qui conduisent vers la maîtrise de ces compétences.

À l'intérieur d'une telle approche les compétences peuvent être décomposées en «un éventail d'habiletés, d'attitudes ou de rôles qui se présentent comme des indices de compétence» (Chené-Williams, 1980, dans Barbès, Colloque 1990, p. 39). «C'est à partir de ces critères qu'on pourra juger du degré de performance ou de la maîtrise plus ou moins grande d'un savoir-faire chez un individu donné» (Tremblay, 1990, p.17).

ÉTAPES DE CONSTRUCTION DES GRILLES D'ÉVALUATION

C'est à partir de cette approche que nous avons construit les outils d'évaluation des stagiaires. Le rapport de recherche présente chacune des étapes en détails. De façon générale, on distingue sept étapes principales :

1. Détermination du type de résultat à mesurer : identification des compétences professionnelles.
2. Priorisation des compétences professionnelles en fonction des niveaux de stage.
3. Détermination des critères qui permettront d'inférer un degré d'atteinte de la compétence : identification des habiletés reliées aux compétences.
4. Détermination des critères qui permettront d'inférer les habiletés : identification des indicateurs.
5. Détermination d'une échelle de mesure graduée permettant d'apprécier le niveau de performance atteint pour chaque comportement : construction de l'échelle d'appréciation.
6. Détermination de la formule de notation : construction du formulaire de notation.
7. Détermination d'un mode d'évaluation des grilles.

1-Identification des compétences professionnelles

En tout premier lieu, c'est à l'identification des résultats attendus qu'il faut s'attarder. À partir de l'analyse de listes de compétences et d'habiletés des éducatrices en services de garde nous avons effectué notre propre sélection des compétences professionnelles. Ces listes de compétences et d'habiletés proviennent de différentes maisons de formation et d'associations canadiennes ou américaines qui utilisent déjà l'approche par compétences pour former leurs étudiants ou pour accréditer le personnel en services de garde.

Nous avons analysé ces listes en fonction de la philosophie éducative du département en matière de garde d'enfants² et de l'exhaustivité des fonctions spécifiques remplies par les éducateurs pour s'acquitter de leur rôle auprès des enfants.

2-Priorisation des compétences selon les niveaux de stage

Après avoir déterminé les compétences d'une éducatrice professionnelle, nous avons sélectionné les compétences à retenir pour chacun des trois stages. En effet, certaines compétences se transposaient bien dans la formation pratique à cause de leur caractère expérimental, tandis que d'autres pouvaient plus facilement être abordées et développées durant certains cours théoriques.

Les compétences retenues pour un stage sont celles qui réfèrent aux habiletés jugées importantes pour le niveau de stage. Ces habiletés deviennent de plus en plus complexes d'un stage à l'autre et leur acquisition est l'objet d'une exigence également croissante.

3-Identification des habiletés

Après avoir déterminé les compétences spécifiques à chacun des stages, nous avons transposé ces compétences en habiletés de façon à permettre aux stagiaires et aux superviseurs de connaître précisément ce qui devait être travaillé en cours de stage.

Les habiletés professionnelles ont été, par la suite, identifiées à partir des grilles déjà élaborées au département. Nous nous sommes aussi référés aux recherches et aux documents sur le sujet ainsi qu'aux habiletés déjà identifiées par d'autres collèges ou associations. Une première liste d'habiletés était ainsi constituée et validée avec la collaboration d'enseignants du département, considérés comme experts dans l'une ou l'autre des disciplines connexes

² Cf. Projet éducatif pour les enfants de 0 à 6 ans dans les services de garde. Ce document est disponible à la Librairie coopérative du Cégep de Saint-Jérôme.

à la compétence. Cet exercice nous a aussi permis de tenir compte des différentes clientèles en services de garde : poupons, enfants de 2 à 5 ans, enfants de 6 à 12 ans.

• *Identification des critères éliminatoires*

L'exercice de priorisation des habiletés nous a permis de nous rendre compte qu'il y avait des comportements et des attitudes de base essentiels au travail auprès des enfants. À l'inverse, l'incapacité d'un étudiant à démontrer ces attitudes de base pouvait porter atteinte au développement sain et harmonieux des enfants. De sorte que nous en sommes venus à préciser des critères, désignés sous le nom de critères éliminatoires. Ils décrivent des comportements ou des attitudes qui mettent en danger, d'une façon ou d'une autre, la santé physique et/ou psychologique des enfants ou qui rendent impossible le déroulement normal d'un stage. Ces critères servent alors comme points de repère aux superviseurs pour retirer une stagiaire du milieu de garde en cours de stage.

Identification des indicateurs

Les indicateurs correspondent à des éléments ou à des comportements observables qui devraient amener plusieurs correcteurs à apprécier une même activité de façon semblable. Il a donc fallu identifier les comportements qui permettraient d'inférer le développement des habiletés reliées à chacune des compétences. Cette étape a exigé plusieurs exercices de validation pour assurer que les éléments choisis soient les plus pertinents, les plus représentatifs des habiletés visées.

Pour identifier ces comportements, Morissette (1984) propose deux questions pour guider l'identification des indicateurs au niveau des tâches et des attitudes.

SAVOIR-FAIRE

Quels sont les principaux comportements des personnes habiles à réaliser l'activité ou la tâche que l'on souhaite mesurer?

SAVOIR-ÊTRE

Quels comportements ont habituellement les personnes qui ont acquis l'attitude que l'on souhaite mesurer?

Nous avons ajouté une troisième question en lien avec les qualités essentielles, de façon à s'assurer de faire ressortir tous les éléments pertinents.

QUALITÉS

Quelles sont les caractéristiques personnelles ou les qualités essentielles à la tâche ou l'attitude que l'on souhaite mesurer?

Tout en référant à ces questions pour identifier les indicateurs, il a fallu tenir compte simultanément de plusieurs règles qui nous assuraient de sélectionner les indicateurs les plus pertinents et ainsi éviter de produire une longue liste de comportements. Les documents sur l'évaluation qui ont été consultés à l'intérieur de cette recherche nous ont permis de faire une synthèse de ces règles à suivre pour choisir les indicateurs d'une grille d'évaluation de performances complexes.

En résumé, les indicateurs doivent être **pertinents et significatifs** (en lien direct avec les habiletés visées); **discriminants** (qui comportent un certain degré de difficulté); **englobants** (qui désignent une série d'autres comportements associés ou sous-entendus); **réalistes** par rapport aux capacités des étudiants; **univoques** (définis en termes simples et non interprétables); et **observables** (qui décrivent concrètement les habiletés). Ajoutons que l'ensemble des indicateurs d'un même champ de compétence doivent constituer un **échantillon représentatif** des habiletés que l'on veut mesurer.

Construction de l'échelle d'appréciation

La grille d'évaluation ainsi constituée présente les champs de compétences sur lesquels devait porter l'évaluation. Pour chaque compétence, les habiletés reliées sont listées et pour chaque habileté, une série d'indicateurs est proposée. Pour apprécier le niveau de performance des stagiaires, nous avons inclus une échelle d'appréciation qui sert à nuancer l'opinion de l'observateur en regard du comportement indicateur décrit.

Notre choix s'est plutôt porté vers une échelle graphique de type *rating scale*, de façon à minimiser les chances d'interprétation puisque les évaluateurs devaient se prononcer sur la fréquence d'apparition des gestes attendus, ce qui nous semblait inciter à l'objectivité. Précisons qu'il ne s'agissait pas de compter le nombre de fois où la stagiaire exécutait tel ou tel geste, mais plutôt d'observer la fréquence à laquelle elle accomplissait l'action attendue et appropriée au moment où la situation le requerrait. Ainsi, on a adopté le principe suivant : plus la stagiaire pose le ou les gestes appropriés au moment opportun, plus on peut inférer qu'elle maîtrise l'habileté reliée.

Construction du formulaire de notation

Ce type de construction de grilles d'évaluation où chaque indicateur apparaît l'un après l'autre donne un portrait plutôt analytique des performances de la stagiaire. Nous avons voulu compléter cette première forme d'évaluation au moyen d'une appréciation plus générale où le superviseur pouvait tenir compte de l'ensemble des habiletés attendues.

Cette méthode dite de l'appréciation globale consiste à percevoir une production, un produit ou un processus dans son ensemble. Dans un esprit de synthèse, nous avons élaboré une grille de notation où les habiletés visées dans le stage concerné sont regroupées sur un même formulaire accompagné d'une échelle d'appréciation critériée pour chaque champ de compétence. Dans ce formulaire, l'échelle d'appréciation est numérique et chaque échelon correspond à un niveau de performance décrit dans les consignes d'utilisation. Le superviseur doit référer à ce barème de notation pour apprécier chaque champ de compétence.

Évaluation des grilles

Le processus d'élaboration des grilles se termine par une évaluation complète des grilles elles-mêmes effectuée par les superviseurs et les chercheurs à la fin de chaque stage d'expérimentation. Il importe à ce stade de vérifier la pertinence et la compréhensibilité des éléments des grilles ainsi que le format de celles-ci et l'utilisation de l'échelle d'appréciation. L'objectif de cette étape est d'amasser le plus d'informations sur l'utilisation des grilles d'évaluation de façon à apporter des corrections qui pourront les améliorer.

CONCLUSION

Cette recherche a permis d'élaborer des grilles où les habiletés et les comportements étaient énoncés de façon à permettre aux superviseurs de référer aux mêmes critères pour juger de la performance des stagiaires. Mais parallèlement, et c'est là une des grandes qualités d'une approche analytique en évaluation, ce type d'outil permet aux étudiants de connaître précisément les points sur lesquels ils sont évalués, réduisant sensiblement le stress relié à l'évaluation.

Ajoutons que ce type d'outils sert aussi bien l'aspect formatif de l'évaluation. Au moyen des indicateurs regroupés autour de compétences, l'étudiant peut obtenir un portrait précis de ses forces et de ses faiblesses lui permettant d'apporter les correctifs nécessaires à une meilleure maîtrise des habiletés attendues.

La conception d'outils d'évaluation en stage a permis à l'équipe de professeurs du département de clarifier ses attentes par rapport aux compétences à développer chez les stagiaires. Cette démarche a favorisé également l'élaboration d'une philosophie éducative auprès des enfants de 0 à 12 ans en services de garde. Cette recherche aura permis de perfectionner la pratique des stages bien qu'on ne puisse prétendre avoir réglé le problème dans sa totalité. Nous poursuivons notre réflexion en vue d'instaurer une pratique des stages qui soit la plus bénéfique pour le développement des habiletés professionnelles des étudiants.

RÉFÉRENCES:

ACHESON, Keith et Meredith GALL. *Techniques in the Clinical Supervision of Teachers*, New-York, Longman, 1992, 238 p.

BARBES, Pierre. «Vers une définition du concept de compétence» dans *Cahier de participation, Colloque AQPC*, 1990, p. 36 à 45.

- BERNARD, Huguette et France FONTAINE. «L'utilisation des objectifs pour évaluer l'apprentissage lors d'un stage» dans *Dossier sur l'évaluation*, Montréal, Université de Montréal, Service pédagogique, 1989, p. 35 à 49.
- CARUSO, Joseph J. et Temple FAWCETT. *Supervision in Early Childhood Education, a Developmental Perspective*, New-York, Teachers College Press, 1986, 242 p.
- DE BAL, R., DE LANDSHEERE, V. G. et J. PAQUAY-BECKERS. *Pédagogie et recherche; Construire des échelles d'évaluation descriptives*, Bruxelles, Ministère de l'Éducation nationale et de la culture française, 1976, 243 p.
- FORTIN, Nicole. *Penser les stages*, Gouvernement du Québec, Ministère de l'Éducation, 1984, 135 p.
- MORISSETTE, Dominique. *La mesure et l'évaluation en enseignement*, Sainte-Foy, Les Presses de l'Université Laval, 1984, 421 p.
- MORISSETTE, Dominique et Maurice GINGRAS. *Enseigner des attitudes?*, Bruxelles, De Boeck-Wesmael, 1989, 193 p.
- PARÉ, André et Thérèse LAFERRIERE. *Inventaire des habiletés nécessaires dans l'enseignement au primaire*, Québec, Centre d'intégration de la personne, 1985, 65 p.
- PELPEL, Patrice. *Les stages de formation : objectifs et stratégies pédagogiques*, Paris, Bordas, 1989, 199 p.
- SCALLON, Gérard. *L'évaluation formative des apprentissages, L'instrumentation*, Québec, Les presses de l'Université Laval, 1988, 263 p.
- TREMBLAY, Gilles. «A propos des compétences comme principe d'organisation d'une formation : éléments de réflexion théorique et perspectives historiques» dans *Bulletin D'information*, vol. 6, n° 9, 1990, 31 p.
- ZIMPER, Nancy. «Evaluating Field-Based Experiences in Teacher Education» in Ayers, Jerry B. et Berney, Mary F., *A Practical Guide to Teacher Education Evaluation*, Boston, Kluwer Academic Publishers, 1989, p. 69 à 84.