



## **L'évaluation de la formation dans l'entreprise : état de la question et perspectives**

Alain Dunberry  
Céline Péchard

**UQÀM**  
Prenez position

**CIRDEP**

# TABLE DES MATIÈRES

PRÉFACE.....	3
INTRODUCTION .....	4
1. POURQUOI ÉVALUER LA FORMATION EN ENTREPRISE? .....	5
1.1 POURQUOI ÉVALUER? .....	5
1.2 POURQUOI NE PAS ÉVALUER? .....	6
1.3 LA DÉCISION DE FORMER ET L'ÉVALUATION.....	6
2. QUOI ÉVALUER? LES MODÈLES D'ÉVALUATION DE LA FORMATION .....	8
2.1 LE MODÈLE DE KIRKPATRICK (1959) .....	8
2.2 LES PRINCIPALES CRITIQUES DU MODÈLE DE KIRKPATRICK .....	8
2.3 LES ALTERNATIVES PROPOSÉES .....	10
2.3.1 Les variantes du modèle de Kirkpatrick.....	10
2.3.2 Les variantes du modèle CIPP.....	10
2.3.3 Deux autres modèles .....	12
2.4 CONCLUSION .....	13
3. LES PRATIQUES ACTUELLES D'ÉVALUATION DE LA FORMATION DES ENTREPRISES .....	14
4. LES PRINCIPAUX FACTEURS QUI INFLUENCENT LES PRATIQUES D'ÉVALUATION DE LA FORMATION .....	16
4.1 LA TAILLE DE L'ENTREPRISE .....	16
4.2 LA CULTURE DE L'ENTREPRISE .....	16
4.3 LES EXIGENCES MÉTHODOLOGIQUES.....	17
5. L'ÉVALUATION DE LA SATISFACTION .....	18
5.1 EN QUOI CONSISTE L'ÉVALUATION DE LA SATISFACTION .....	18
5.2 POURQUOI ÉVALUER LA SATISFACTION? .....	18
5.3 PRATIQUES ET TENDANCES ACTUELLES D'ÉVALUATION DE LA SATISFACTION .....	19
5.4 LIMITES DES PRATIQUES D'ÉVALUATION DE LA SATISFACTION.....	20
5.4.1 Les limites liées à l'administration de questionnaires de satisfaction. ....	20
5.4.2 La satisfaction telle que pratiquée corrèle mal avec les autres niveaux de l'évaluation.....	20
5.4.3 La satisfaction, un construit multidimensionnel .....	21
6. L'ÉVALUATION DES APPRENTISSAGES.....	22
6.1 EN QUOI CONSISTE L'ÉVALUATION DES APPRENTISSAGES.....	22
6.2 POURQUOI ÉVALUER LES APPRENTISSAGES?.....	22
6.3 PRATIQUES ET TENDANCES ACTUELLES D'ÉVALUATION DES APPRENTISSAGES.....	23
6.3.1 Les pratiques actuelles .....	23
6.3.2 Une tendance nouvelle en évaluation des apprentissages .....	24
6.3.3 La prise en compte des facteurs individuels et organisationnels .....	25
6.4 LIMITES DES PRATIQUES D'ÉVALUATION DES APPRENTISSAGES .....	25
6.4.1 Des exigences de rigueur rarement comblées .....	25
6.4.2 Des exigences mal adaptées à certains apprentissages .....	25
6.4.3 Un pouvoir prédictif limité.....	26
6.4.4 Les risques de l'évaluation des apprentissages .....	26
6.4.5 Un découpage qui ne s'applique pas à toutes les formations.....	26

<b>7. L'ÉVALUATION DES COMPORTEMENTS</b> .....	27
7.1 EN QUOI CONSISTE L'ÉVALUATION DES COMPORTEMENTS .....	27
7.2 POURQUOI ÉVALUER LES COMPORTEMENTS? .....	27
7.3 LES FACTEURS INFLUENÇANT LE TRANSFERT.....	27
7.3.1 Les caractéristiques du stagiaire.....	28
7.3.2 La conception de la formation .....	29
7.3.3 L'environnement de travail.....	29
7.4 PRATIQUES ET TENDANCES ACTUELLES D'ÉVALUATION DES COMPORTEMENTS .....	30
7.4.1 Les pratiques courantes et les conseils généraux .....	30
7.4.2 Déterminer le meilleur moment .....	31
7.4.3 Le choix des acteurs .....	31
7.4.4 La nécessité de faire un diagnostic préalable .....	31
7.4.5 Au-delà de l'évaluation des seuls comportements .....	31
7.5 LIMITES DES PRATIQUES D'ÉVALUATION DES COMPORTEMENTS.....	32
<b>8. L'ÉVALUATION DES RÉSULTATS DE LA FORMATION SUR L'ORGANISATION</b> .....	33
8.1 EN QUOI CONSISTE L'ÉVALUATION DES RÉSULTATS DE LA FORMATION .....	33
8.2 POURQUOI ÉVALUER LES RÉSULTATS DE LA FORMATION .....	33
8.3 PRATIQUES ET TENDANCES ACTUELLES D'ÉVALUATION DES RÉSULTATS DE LA FORMATION .....	33
8.3.1 Les principales difficultés de l'évaluation des résultats de la formation .....	33
8.3.2 Comment dépasser les difficultés de l'évaluation des résultats de la formation .....	34
8.3.3 Le choix des indicateurs et des méthodes.....	35
8.3 LIMITES DES PRATIQUES D'ÉVALUATION DES RÉSULTATS DE LA FORMATION .....	36
<b>9. L'ÉVALUATION DU RENDEMENT FINANCIER DE LA FORMATION</b> .....	37
9.1 EN QUOI CONSISTE L'ÉVALUATION DU RENDEMENT FINANCIER .....	37
9.2 POURQUOI ÉVALUER LE RENDEMENT FINANCIER DE LA FORMATION.....	37
9.3 PRATIQUES ET TENDANCES ACTUELLES D'ÉVALUATION DU RENDEMENT FINANCIER .....	38
9.3.1 Le calcul du ROI selon Phillips (1997).....	38
9.3.2 Deux méthodes simples.....	39
9.3.3 L'approche systémique de Wang et al. (2002).....	40
9.3.4 Quelques critères de sélection de la bonne méthode. ....	41
9.4 LIMITES DES PRATIQUES D'ÉVALUATION DU RENDEMENT FINANCIER.....	41
<b>10. CONCLUSION</b> .....	42
<b>BIBLIOGRAPHIE</b> .....	44

# PRÉFACE

Ce document a été produit dans le cadre du projet de recherche « *Analyse des pratiques d'évaluation de la formation et de son rendement dans des entreprises performantes* » financé par le Programme de subvention à la recherche appliquée (PSRA) du Fonds national de formation de la main d'oeuvre (FNFMO) du Québec. Par la description et l'analyse de pratiques existantes, ce projet souhaite contribuer au développement et au perfectionnement de l'évaluation qui constitue un processus essentiel à la démonstration des effets de la formation dans les entreprises et à son amélioration continue.

Cette recherche a d'abord permis une collecte de données sur les pratiques d'évaluation de 12 grandes et moyennes entreprises considérées comme performantes en formation, également réparties entre les secteurs de la métallurgie, du commerce de détail et des communications graphiques. Ces données ont fait l'objet d'un rapport d'analyse détaillé, et d'un document à caractère plus descriptif pour chacune des 12 entreprises. Un rapport synthèse destiné aux décideurs et praticiens sera aussi produit.

La recension des écrits est un des produits de cette recherche. Elle s'adresse principalement aux décideurs et professionnels concernés par l'évaluation de la formation. Elle vise à faire le point sur l'état des débats et pratiques entourant cette thématique en s'appuyant sur la littérature scientifique internationale.

Cette recherche du Centre interdisciplinaire de recherche sur l'éducation permanente (CIRDEP) a été dirigée par Alain Dunberry avec l'assistance des chercheuses Maryse Larivière et Céline Péchard. Sa réalisation a aussi bénéficié de la collaboration des chercheurs Paul Bélanger, Lucie Morin, Colette Bérubé, Pierre Doray, Louise Gaudreau et Pierre Valois

Nous tenons à remercier vivement la Commission des Partenaires du Marché du Travail et le personnel du PSRA pour leur appui à ce projet; la direction des trois comités sectoriels qui ont contribué à l'identification des entreprises; de même que le personnel des 12 entreprises qui ont généreusement accepté d'y participer.

Février 2007.

# INTRODUCTION

Selon Mark, Henry, et Julnes (2000), l'évaluation de la formation dans l'entreprise vise principalement quatre buts : démontrer la pertinence et la valeur de la formation; améliorer la qualité des programmes et de l'organisation de la formation; vérifier la conformité des programmes aux diverses attentes, normes ou standards; et développer ou tester des connaissances nouvelles. C'est donc dire qu'une entreprise qui évalue sa formation devrait être plus en mesure d'en constater et contrôler les résultats et d'améliorer de manière soutenue ses pratiques.

Il convient ici de bien distinguer entre «*évaluer*» et «*mesurer*». L'évaluation consiste essentiellement à porter un jugement sur la valeur de l'objet à évaluer. Ce jugement peut s'appuyer sur des constats, des impressions ou des mesures. Il est donc possible d'évaluer sans mesurer, ce que nous faisons couramment. Par exemple, nous n'avons pas besoin de prendre des mesures d'espace, de volume ou autres pour évaluer si une chambre d'hôtel nous convient. Par ailleurs, la mesure est essentiellement associée à l'attribution de valeurs numériques et peut s'avérer parfois nécessaire à une bonne évaluation. Ainsi, il est certainement plus sage de s'appuyer sur des mesures de superficie lorsque vient le temps d'évaluer la valeur d'un terrain. Cette recension des écrits couvre un ensemble de pratiques d'évaluation qui peuvent ou non s'appuyer sur des mesures.

On peut regrouper les pratiques d'évaluation de la formation en deux catégories principales : les pratiques formelles et informelles. Les pratiques formelles d'évaluation de la formation peuvent se définir comme étant toutes les pratiques structurées et explicites visant à établir un jugement sur la valeur d'un processus ou produit de la formation. Par exemple, l'utilisation d'un questionnaire en vue d'évaluer la satisfaction des participants à une activité, même si elle est occasionnelle, constitue une pratique formelle d'évaluation de la formation. Il en est de même de l'observation systématique du comportement d'un employé en poste de travail à l'aide d'une grille. Les pratiques informelles d'évaluation de la formation regroupent plutôt les pratiques spontanées, non structurées, implicites ou explicites, visant elles aussi à établir un jugement sur la valeur d'un processus ou produit de la formation. Par exemple, le jugement spontané, sans aucun support technique et non consigné d'un compagnon sur la performance de son apprenti constitue une pratique d'évaluation informelle. Il en est de même des propos recueillis dans le corridor par un responsable sur la qualité d'une activité de formation qui se termine. Jusqu'à maintenant, la littérature scientifique s'est surtout centrée sur les pratiques formelles d'évaluation de la formation. Néanmoins, notre recension des pratiques a permis de constater non seulement l'existence mais aussi

l'importance des pratiques informelles qui dans bien des cas précèdent et suppléent à des pratiques formelles qui seraient trop complexes à mettre en place.

Comme le soulignent Kraiger, Ford, et Salas (1993), l'expression «*évaluation de la formation*» est elle-même porteuse d'ambiguïté. L'évaluation de la formation peut en effet référer à l'évaluation du processus de formation («*training efficiency*» ou efficience de la formation), à comment la formation a été conçue et donnée. Elle peut aussi référer à l'évaluation des résultats de la formation («*training effectiveness*» ou efficacité de la formation), ce que les participants ont appris, ses effets sur leurs comportements et sur l'entreprise. Ces deux sens sont ici couverts.

Après un bref survol des raisons pouvant inciter ou non à évaluer la formation (section 1), ce document présente le modèle de Kirkpatrick (1959) encore fort présent dans le monde de la formation dans l'entreprise, les principales critiques qui lui sont adressées et des alternatives qui y sont proposées (section 2). La section 3 fait état des pratiques d'évaluation des entreprises américaines et québécoises. Subséquemment, la littérature scientifique est présentée selon cinq principaux niveaux ou critères d'évaluation, à savoir la satisfaction, les apprentissages, les comportements, les résultats sur l'entreprise et le rendement financier. Enfin, la conclusion porte sur les développements possibles de l'évaluation de la formation.

# 1. POURQUOI ÉVALUER LA FORMATION EN ENTREPRISE?

Pour quelles raisons une entreprise devrait-elle consacrer une partie de ses ressources de formation à l'évaluation? Le fait que l'on dise souvent qu'il s'en fait relativement peu serait-il justifié? La littérature propose, bien sûr, de bonnes raisons d'évaluer. Mais elle identifie aussi plusieurs raisons qui peuvent expliquer pourquoi certaines entreprises préfèrent ne pas le faire.

## 1.1 POURQUOI ÉVALUER?

Au départ, il importe de bien savoir pourquoi évaluer : la réponse à cette question est capitale pour définir les voies et moyens qui seront mis en place.

S'appuyant essentiellement sur leur expérience, plusieurs auteurs font valoir différentes raisons d'évaluer. Il en est ainsi de deux auteurs majeurs en évaluation de la formation, Kirkpatrick et Phillips. Kirkpatrick (1998) en propose trois : 1) justifier l'existence du département de formation en démontrant sa contribution aux objectifs de l'entreprise; 2) décider si l'on doit poursuivre ou arrêter les programmes de formation; 3) obtenir de l'information sur comment améliorer les futurs programmes de formation. Pour sa part, Phillips (1997) identifie 10 raisons ou fonctions de l'évaluation : 1) déterminer l'atteinte des objectifs de formation; 2) identifier les forces et faiblesses du processus de développement des ressources humaines; 3) comparer les coûts et bénéfices d'un programme; 4) décider qui devrait participer dans les prochains programmes; 5) tester la clarté et la validité de certains tests, études de cas ou exercices; 6) identifier quels participants ont le mieux réussi le programme; 7) rappeler aux participants certains contenus du programme jugés importants; 8) recueillir des données en vue de la mise en marché de futurs programmes; 9) déterminer si le programme offrait une solution adéquate au besoin identifié; 10) constituer une base de données en appui à la prise de décision par les gestionnaires.

Russ-Eft et Preskill (2001) proposent sept raisons d'évaluer : 1) l'évaluation peut contribuer à l'amélioration de la qualité ou la confirmer; 2) elle peut contribuer à accroître les connaissances des membres de l'organisation, soit en comprenant mieux ce qui est évalué, soit en apprenant à bien évaluer; 3) elle peut aider à déterminer les priorités, compte tenu de l'ampleur des besoins et des ressources limitées; 4) l'évaluation, utilisée comme outil diagnostic, peut contribuer à la planification et la livraison de nouvelles activités au sein de l'organisation; 5) elle peut aider à démontrer la valeur ajoutée des efforts de formation; 6) les résultats d'évaluation peuvent aider à

documenter et appuyer certaines demandes de ressources additionnelles; 7) l'expérience en évaluation devient de plus en plus recherchée.

De toute évidence, les raisons évoquées par ces différents auteurs se recourent. Kraiger (2002) propose pour sa part une synthèse. S'appuyant sur quelques auteurs (Kirkpatrick, 1994; Sackett & Mullin, 1993; Twitchell, Holton, & Trott, 2000), il identifie trois raisons majeures : la prise de décision, la rétroaction et le marketing.

- La première est de fournir des informations permettant une prise de décision éclairée sur la formation : par exemple, le niveau de rétention des apprentissages est-il suffisant compte tenu des ressources allouées? Le format du cours correspond-t-il adéquatement au contenu et celui-ci est-il pertinent? Les formateurs sont-ils capables de bien livrer le contenu? Les participants reçoivent-ils suffisamment de formation pour effectuer les tâches qui sont attendues d'eux?
- La deuxième raison principale est la rétroaction. Cette rétroaction peut être donnée tant aux formateurs qu'aux concepteurs ou aux participants à la formation. Elle permet essentiellement des ajustements des activités de formation ou, dans le cas de l'apprenant, de sa démarche d'apprentissage.
- Enfin, la troisième raison concerne le marketing de la formation. Ce marketing peut se faire à l'interne de l'organisation ou auprès de participants potentiels. Sans surprise, l'évaluation de l'impact de la formation sur l'entreprise semble la plus convaincante pour démontrer la valeur de la formation auprès de la haute direction. De plus, la démonstration de ces résultats aurait un effet motivateur sur les futurs participants.

Cette synthèse de Kraiger offre l'intérêt de situer trois grandes fonctions au sein des unités de formation : la gestion de la formation et les processus de décision qui l'accompagnent; la fonction technique, particulièrement préoccupée par les apprentissages et la qualité des processus; et la promotion de la formation tant à l'intérieur qu'à l'extérieur de l'organisation. C'est déjà indirectement poser la question à savoir dans quelle mesure ces trois fonctions sont actuellement bien servies par les pratiques d'évaluation de la formation.

Enfin, l'étude de Twitchell et al. (2000) a montré que la fréquence des évaluations de la satisfaction, des apprentissages et des comportements était corrélée à la perception par les responsables de la valeur de l'évaluation pour l'amélioration de la formation.

## 1.2 POURQUOI NE PAS ÉVALUER?

Plusieurs auteurs (Phillips, 1997; Kirkpatrick, 1998; Grove & Ostroff, 1991 in Goldstein & Ford, 2001; Lewis & Thornhill, 1994; Twitchell et al., 2000; Russ-Eft & Preskill, 2001; Kraiger, 2002) ont aussi identifié des raisons qui peuvent expliquer pourquoi les pratiques évaluatives des entreprises sont limitées. Encore ici, plusieurs de ces raisons se recoupent. On peut les regrouper en cinq catégories principales :

- personne n'exprime la volonté que la formation soit évaluée; cette attitude peut découler d'une croyance très ferme en la formation, ou de la conviction de savoir déjà, par une évaluation informelle, ce qui fonctionne ou pas; cette absence de volonté peut se manifester au niveau de la haute direction, bien que des auteurs tels que Phillips (1997) soulignent l'importance croissante de démontrer à celle-ci l'apport de la formation à l'entreprise; cette raison vient en tête de liste dans l'enquête de Twitchell et al. (2000) auprès des entreprises américaines;
- le fait de ne pas savoir ce que l'on veut et peut évaluer; il est évident que les modèles de Kirkpatrick et de Phillips, que nous verrons au point suivant, aident les entreprises à identifier quoi évaluer; néanmoins, une mise en application bien contextualisée de ces modèles suppose que l'on a bien compris les buts et rôles de l'évaluation et comment celle-ci peut contribuer aux besoins de l'entreprise;
- l'absence de compétences techniques en évaluation : ce ne sont pas toutes les entreprises qui ont accès à des ressources techniques compétentes, capables de prendre en main le dossier de l'évaluation de la formation; et ce, surtout si l'on souhaite dépasser les seules évaluations de la satisfaction ou des apprentissages, pour aborder l'évaluation des comportements ou des impacts ou encore pour concevoir et mettre en place un système qui réponde de manière adéquate aux besoins; cette raison vient en second lieu dans l'enquête de Twitchell et al. (2000);
- le risque : il est possible que l'on redoute l'impact des constats de l'évaluation sur la direction de la formation, par exemple, si elle démontre l'effet de mauvaises décisions de gestion; ce risque, lié à l'absence de volonté exprimée par la haute direction pour une telle démarche, peut inciter à se faire oublier plutôt que d'aller de l'avant;
- le coût : l'évaluation est une opération qui peut sembler fastidieuse et il se peut que l'on ne soit pas certain que les résultats de celle-ci en justifient les coûts, surtout si l'on a eu une expérience passée décevante; cette raison vient en troisième lieu dans l'enquête de Twitchell et al. (2000).

La plupart de ces raisons sont donc susceptibles d'être remises en cause avec une formation adéquate des décideurs et techniciens responsables de l'évaluation de la formation.

## 1.3 LA DÉCISION DE FORMER ET L'ÉVALUATION

On peut se demander s'il existe un lien entre la décision de former et l'évaluation de la formation. En d'autres termes, l'évaluation du rendement de la formation est-elle susceptible de contribuer à accroître son importance dans l'entreprise?

Selon Betcherman, Leckie, et McMullen (1997), sur un plan macro, la décision de former semble beaucoup associée à des facteurs tels que la taille de l'entreprise, le recours à la technologie et la stratégie de gestion des ressources humaines. Au niveau de l'entreprise elle-même, elle est souvent déterminée par un événement particulier, comme un changement technologique ou organisationnel, ou elle est une activité de nature épisodique. Les employeurs peuvent former pour adapter leur main d'œuvre à des changements technologiques, pour favoriser la mobilité intra-entreprise, mais aussi de manière relativement permanente, en vue d'instaurer et de maintenir une organisation apprenante. Cette dernière tendance est confirmée par Leckie, Léonard, Turcotte, et Wallace (2001) qui arrivent au constat qu'une proportion importante d'entreprises offrent de la formation sans pour autant qu'elle soit associée à un changement dans les modes de production.

Plus récemment, au Québec, Lapierre, Méthot, et Tétrault (2005) constatent que les principales raisons qui amènent les employeurs à investir dans la formation sont le maintien d'une pratique courante (59 %), l'adaptation à des changements (46 %), l'augmentation de la compétitivité (38 %) des problèmes à recruter ou maintenir une main-d'œuvre qualifiée (29 %), l'obligation due à un contrat ou à une autre loi (24 %), l'obligation due à la «Loi du 1 %» (23 %) et la volonté d'accroître la motivation (19 %).

Figgis et al. (2001) ont procédé à une étude visant à identifier comment faire valoir la pertinence d'investir plus dans la formation et l'apprentissage auprès des entreprises australiennes. Selon ces auteurs, le calcul du retour sur l'investissement de formation demeure une préoccupation pour les entreprises, mais celles-ci se satisfont d'indicateurs tels que la façon dont les employés conçoivent leur travail et pensent à l'améliorer suite à la formation. En ce sens, elles auraient plus besoin d'appuis dans la définition et la mise en application de ces indicateurs que de calculs de rendement financier. L'étude constate aussi l'importance de la culture de l'entreprise, et parti-

culièrement celle de la teneur des relations interpersonnelles, qui détermine pour une bonne part le retour sur l'investissement en formation.

Au-delà des considérations d'ordre technique, l'ensemble de ces recherches explique en partie pourquoi il se fait peu d'évaluation formelle du rendement de la formation, même dans les entreprises qui évaluent leur formation.



## 2. QUOI ÉVALUER? LES MODÈLES D'ÉVALUATION DE LA FORMATION

Encore aujourd'hui, le modèle de Kirkpatrick (1959) occupe une place prépondérante dans la littérature sur l'évaluation de la formation en entreprise. Comme le soulignent Salas et Cannon-Bowers (2001), bien qu'il ait fait l'objet de multiples interventions<sup>1</sup>, il continue à être le cadre de référence le plus populaire en matière d'évaluation. Ceci est certainement attribuable à sa pertinence et à sa simplicité, bien que des aspects importants de ce modèle soient périodiquement remis en cause. Cette section propose donc une présentation sommaire de ce modèle, les principales critiques qui lui sont adressées et les alternatives proposées.

### 2.1 LE MODÈLE DE KIRKPATRICK (1959)

Le modèle de Kirkpatrick a pris naissance en 1959, par une série de quatre articles sur l'évaluation des programmes de formation dans la revue « *Training and Development* ». Ces quatre articles définissaient les quatre niveaux de l'évaluation qui devaient par la suite avoir une influence marquante sur les pratiques des entreprises. Ces quatre niveaux sont les suivants :

- Niveau 1 : la réaction. Ce niveau vise ce que les participants pensent du programme ou de l'activité. Kirkpatrick (1998) l'associe à une évaluation de la satisfaction du consommateur. Selon lui, si la formation se veut efficace, elle se doit de provoquer des réactions favorables de la part des participants. Dans le cas contraire, ils risquent de ne pas être motivés et, par surcroît, de compromettre l'existence du programme par des commentaires négatifs. La réaction se mesure généralement à l'aide de formulaires dans lequel le participant fait part de son niveau de satisfaction à l'égard de différents aspects de l'activité.
- Niveau 2 : les apprentissages. Ce niveau vise ce que les participants ont acquis par la formation en termes de connaissances, d'habiletés ou d'attitudes en fonction des objectifs visés. L'importance d'évaluer ce niveau découle du simple fait qu'aucune modification du comportement ne peut survenir si les participants n'ont pas appris. De plus, cette évaluation permet d'éviter de conclure trop rapidement que s'il n'y a pas eu de modification de comportement en poste de travail, c'est parce qu'il n'y a pas eu d'apprentissage. Cette évaluation se fait par l'utilisation de tests (de performance) ou d'examens.
- Niveau 3 : les comportements. Ce niveau évalue le transfert en poste de travail des connaissances, habiletés et attitudes apprises lors de la formation. En somme, il s'agit de savoir ici en quoi la formation

a changé les comportements en situation de travail. L'évaluation, qui devrait se situer environ deux à trois mois après la formation, fait appel ici à des méthodes qui incluent l'enquête ou l'entrevue auprès du formé et de différents acteurs en mesure d'observer son comportement.

- Niveau 4 : les résultats. Il s'agit ici de l'évaluation la plus importante, à savoir l'impact de la formation sur l'entreprise. Il s'évalue entre autres en termes de productivité, de qualité, de taux de roulement du personnel, de qualité de vie au travail, de baisse des coûts, de retour sur l'investissement de formation, etc. Selon l'auteur, l'évaluation à ce niveau se fait rarement, d'une part parce que les formateurs ne savent pas comment procéder, et d'autre part parce qu'il est difficile d'associer les résultats aux seuls effets de la formation. Les méthodes pour évaluer ce dernier niveau varient selon l'objet à évaluer.

Le terme de « *niveau* » souligne la hiérarchie existant entre ces quatre critères d'évaluation. Plus le niveau est élevé, plus il fournit de l'information qui revêt de l'importance pour l'entreprise. De plus, il existe selon l'auteur des liens de causalité entre ces niveaux. C'est parce qu'il y aura eu satisfaction qu'il y aura apprentissage; c'est parce qu'il y aura eu apprentissage qu'il y aura changement de comportement; et c'est parce qu'il y aura eu changement de comportement qu'il y aura résultats sur l'entreprise.

Soulignons enfin que Kirkpatrick insiste sur le fait que les quatre niveaux d'évaluation constituent des parties importantes de l'évaluation. Il ne propose donc pas un choix de quatre critères, mais plutôt qu'une évaluation complète de la formation devrait couvrir ces quatre critères.

### 2.2 LES PRINCIPALES CRITIQUES DU MODÈLE DE KIRKPATRICK

Alliger et Janak (1989), dans un article très souvent cité, remettent en cause le caractère hiérarchique et causal du modèle de Kirkpatrick. En ce qui concerne le caractère hiérarchique, les auteurs prétendent que ce ne sont pas toutes les formations qui visent, par exemple, des modifications de comportement au travail. Dans certains cas, comme des formations d'intégration à l'entreprise, l'évaluation des apprentissages ou de la satisfaction sera plus importante.

Le modèle suppose aussi une relation causale entre les niveaux qui devrait donner lieu à des corrélations entre ceux-ci. En ce qui concerne les relations causales entre les niveaux, les auteurs, s'appuyant sur des études de

<sup>1</sup> "It has been used, criticized, misused, expanded, refined, adapted, and extended" (op.cit., p.487)

Remmers, Martin et Elliott (1949) et Rodin et Rodin (1972) in Alliger et Janak (1989) évoquent entre autres le fait que parfois une réaction négative due à une activité exigeante peut donner lieu à beaucoup d'apprentissage. Par ailleurs, selon Kaplan et Pascoe (1977) in Alliger et Janak (1989), des formations agréables peuvent s'avérer sans résultats. Donc, la relation causale du niveau 1 (réaction) sur le niveau 2 ne se confirme pas toujours.

De plus, Alliger et Janak (1989) présentent les résultats d'une première revue de 12 articles ayant évalué les effets de la formation à différents niveaux. Celle-ci vient infirmer l'hypothèse selon laquelle les changements à chaque niveau sont corrélés entre eux : les auteurs constatent en effet que la satisfaction tend à corrélérer seulement légèrement avec les trois autres niveaux : la corrélation entre la satisfaction et l'apprentissage est de ,07; entre la satisfaction et le comportement, de ,05; et entre la satisfaction et les résultats, de ,48, mais cette dernière est cependant basée sur une seule étude. Par ailleurs, les niveaux 2, 3 et 4 tendent à corrélérer entre eux de manière légèrement plus prononcée : la corrélation entre les apprentissages et les comportements est de ,13; entre les apprentissages et les résultats, de ,40; entre les comportements et les résultats, de ,19. Ceci semble suggérer que la satisfaction devrait être considérée à part des autres critères. Mais dans l'ensemble, force est de constater que la corrélation entre les niveaux est faible.

Dans un article plus récent, Alliger, Tannenbaum, Bennett, Traver, et Shortland (1997) reprennent le même exercice, mais cette fois en subdivisant les niveaux du modèle de Kirkpatrick. Ainsi, la réaction est subdivisée en trois dimensions : la dimension de l'utilité (« *cette formation me sera utile dans mon travail* »), la dimension affective (« *cette formation m'a beaucoup plu* ») et une dimension mixte. L'apprentissage est subdivisé en trois dimensions : connaissances immédiatement en fin de formation, rétention des connaissances un certain temps après la formation, et comportements dans le cadre de la formation. Le terme « *comportement* » est remplacé par « *transfert* » afin de souligner qu'il s'agit là d'un comportement en situation de travail, et mesuré un certain temps après la formation. Compte tenu du peu d'évaluations d'impact disponibles, l'étude n'a pas pris en compte le niveau 4. S'appuyant cette fois-ci sur 34 études, les auteurs font les principaux constats suivants. La corrélation entre l'ensemble des réactions et l'apprentissage immédiat est de ,07, ce qui confirme le résultat de l'étude antérieure. Il apparaît pertinent de distinguer les réactions reliées à l'utilité des réactions purement affectives. En effet, si les réactions affectives corrélerent très peu avec les autres niveaux, les réactions reliées à l'utilité corrélerent autant ( $r = ,18$ ) avec le transfert que les comportements en fin de formation et plus que les connaissances en fin de formation ( $r = ,11$ ) et un certain temps après la formation ( $r = ,08$ ). Mais dans l'ensemble, on constate des corrélations relativement faibles, nettement

insuffisantes pour véritablement prédire les résultats à un niveau supérieur.

La critique de Holton (1996) est certainement la plus virulente. Celui-ci considère que l'approche de Kirkpatrick est en fait une simple taxonomie des résultats de la formation qui se limite à identifier ce qui doit être évalué et non un modèle et encore moins une théorie. En effet, un véritable modèle d'évaluation de la formation devrait mieux préciser les résultats de la formation, tenir compte des variables intermédiaires, telles que la motivation à apprendre ou les caractéristiques des apprenants, qui influencent ces résultats et identifier des relations de cause à effet, ce que ne réussit pas à faire le modèle en question. Dans une réaction à l'article de Holton, Kirkpatrick (1996) conclut en disant que ce qu'il souhaite d'abord et avant tout, c'est que son « *modèle – ou taxonomie – continue à être utile aux professionnels du développement des ressources humaines qui sont plus intéressés par les idées pratiques que les recherches académiques.* » (p. 24).

Kraiger et al. (1993) avaient déjà émis une critique du modèle de Kirkpatrick concernant surtout le manque de clarté quant aux changements pouvant être attendus de l'apprentissage et, en conséquence, la difficulté à identifier des techniques d'évaluation appropriées. Plus récemment, Kraiger (2002) énonce cinq critiques majeures à l'encontre du modèle de Kirkpatrick : 1) le modèle a peu de bases théoriques et ignore tous les acquis des théories cognitivistes des années 70 et 80; 2) les construits utilisés, tels que la satisfaction ou l'apprentissage, ne sont pas clairement définis, se révélant plus complexes qu'il n'apparaît; 3) une application stricte de l'approche amène des comportements illogiques, comme ne pas évaluer les apprentissages avant d'avoir confirmé une réaction positive alors que les participants ne sont peut-être plus disponibles; 4) l'approche présume de relations entre les résultats de la formation qui ne se confirment pas, comme c'est le cas entre la satisfaction et l'apprentissage; et 5) l'approche ne tient pas compte du but de l'évaluation, en incitant à évaluer les quatre niveaux, alors que dans certains cas l'évaluation d'autres critères serait plus pertinente en regard des besoins de l'entreprise.

Enfin, Geertshuis, Holmes, Geertshuis, Clancy, et Bristol (2002) apportent une critique fort intéressante qui se distingue de celles évoquées précédemment. Les auteurs soulignent que l'éducation tout au long de la vie implique une nouvelle vision du travailleur qui devient un participant actif à son développement personnel. En conséquence, les modèles d'évaluation de la formation doivent reconnaître cet « *empowerment* » (ou « *pouvoir d'agir* ») et les professionnels de la formation ne peuvent plus se contenter des seuls modèles où les critères sont définis par les voies hiérarchiques. Il faudrait désormais avoir une vision de l'apprenant comme étant aussi un être indépendant et autonome, et que les méthodologies d'évaluation utilisées lui donnent l'occasion de mieux

participer à ce processus, entre autres en lui permettant de réfléchir, donner son point de vue personnel, débattre et argumenter ou encore questionner l'information qui lui est transmise.

## 2.3 LES ALTERNATIVES PROPOSÉES

De multiples modèles d'évaluation de la formation en entreprise ont été proposés au cours des dernières années. Il convient de préciser que ces modèles se situent dans le plus vaste ensemble des modèles d'évaluation de programme, ces modèles pouvant très bien être appliqués à l'évaluation de la formation en entreprise. Or, une recension récente de Stufflebeam (2001) fait état de 22 approches d'évaluation, ces approches pouvant regrouper plusieurs modèles. C'est dire que le choix d'alternatives proposées ici est volontairement restreint. Le lecteur souhaitant aller plus en profondeur sur cette question pourra se référer soit au document de Tamkin, Yarnall, et Kerrin (2002) pour ce qui est des modèles plus directement associés à l'évaluation de la formation en entreprise, ou au document de Stufflebeam (2001) qui permet d'avoir une vision élargie de ce que pourrait être l'évaluation de la formation dans l'entreprise au-delà des pratiques actuellement reconnues dans ce domaine précis.

Les modèles que nous avons retenus ici peuvent être classés en trois catégories : les variantes du modèle de Kirkpatrick, les variantes du modèle CIPP et deux autres modèles qui se distinguent des précédents, à savoir ceux de Holton (1996) et de Kraiger (2002). Cette section du rapport se limite essentiellement à faire état de ce que ces modèles proposent d'évaluer, le lecteur pouvant se référer aux textes originaux s'il souhaite approfondir l'un d'entre eux.

### 2.3.1 Les variantes du modèle de Kirkpatrick

Certaines critiques ont donné lieu à plusieurs variantes du modèle de Kirkpatrick. Parmi celles-ci, notons celles de Hamblin (1974), Molenda, Pershing, et Reigeluth (1996), Phillips (1997) et Kearns et Miller (1997). Ces variantes apparaissent particulièrement près du modèle original, comme en témoigne le tableau qui suit. En effet, elles se centrent exclusivement sur l'évaluation des produits et résultats de la formation. Toutes conviennent de la pertinence d'évaluer la réaction, l'apprentissage et le comportement qui dans certains cas devient le transfert des apprentissages en poste de travail. Les distinctions se situent principalement au niveau des résultats. En effet, Hamblin et Phillips distinguent les résultats ou impacts de la formation et les résultats à caractère plus financier, ce qui est justifiable du fait que leur calcul réfère à d'autres indicateurs et méthodes d'évaluation. Par ailleurs, Molenda et al. (1996) ajoutent, à l'impact d'affaires, l'impact social. Notons aussi que ces derniers auteurs intègrent à l'évaluation un niveau premier de comptabilité des activités (« *activity accounting* ») qui s'apparente à ce que l'on pourrait appeler la mesure des extrants opérationnels (nombre de personnes formées).

### 2.3.2 Les variantes du modèle CIPP

Le modèle CIPP est largement utilisé en évaluation de programme. Il est associé à Stufflebeam et al. (1971) qui proposent que l'évaluation de programme porte sur les quatre éléments suivants :

- ↳ le Contexte : cette évaluation porte sur l'environnement qui conditionne les besoins de formation et les facteurs ou problèmes sous-jacents à ses besoins; celle-ci, semblable à une étude de faisabilité, permettra de déterminer les finalités et objectifs de la formation;

**TABLEAU 1**  
**Analyse comparative de variantes du modèle de Kirkpatrick**

Kirkpatrick (1959)	Hamblin (1974)	Molenda et al. (1996)	Phillips (1997)	Kearns et Miller (1997)
		Comptabilité des activités		
Réaction	Réaction	Réaction	Réaction	Réaction
Apprentissage	Apprentissage	Apprentissage	Apprentissage	Apprentissage
Comportement	Comportement au travail	Transfert de l'apprentissage	Mis en application au travail	Transfert au poste de travail/comportement
Résultats	Organisation	Impact d'affaires	Résultats d'affaires	Valeur ajoutée
	Valeur ultime	Impact social	Retour sur l'investissement	

- les Intrants : cette évaluation permet de déterminer les ressources humaines, matérielles et financières de même que les stratégies qui seront mises en place afin d'atteindre les objectifs;
- le Processus : cette évaluation porte sur la façon selon laquelle les intrants sont utilisés pour atteindre les objectifs, donc principalement comment la formation est effectivement menée; elle rejoint en quelque sorte l'évaluation de la réaction et permet de comprendre, du moins en partie, le niveau d'atteinte des objectifs;
- les Produits : cette évaluation porte sur les différents extrants et résultats du programme.

Le diagramme qui suit permet de visualiser l'ensemble du modèle. Le contexte agit sur l'ensemble des éléments. Les intrants sont investis dans le processus qui génère les produits. Le recours au CIPP amène donc à décloisonner l'évaluation de la formation des seuls produits et résultats pour prendre en compte les processus, les intrants et le contexte de manière explicite.

Trois modèles d'évaluation de la formation présentent des analogies à la fois avec le CIPP et le modèle de Kirkpatrick. Il s'agit du CIRO de Warr et al. (1970 in Tamkin et al., 2002), du IPO de Bushnell (1990) et du modèle de Kaufman et al. (1995).

Le modèle CIRO de Warr et al. (1996) est assez semblable au modèle de Stufflebeam et al. (1971). Il propose l'évaluation du contexte, des intrants, des réactions et des produits («outcomes»), l'évaluation des réactions rejoignant celle des processus. Les résultats peuvent être évalués à trois niveaux (immédiat, intermédiaire et terminaux). Ces niveaux présentent donc une analogie avec les apprentissages, les comportements et les résultats de Kirkpatrick.

Le modèle IPO de Bushnell (1990) propose d'évaluer : les intrants, dont l'expérience du formateur et les caractéristiques des apprenants; les processus, à savoir le design de l'activité et son mode de livraison; les extrants, qui comprennent les réactions des apprenants, les connaissances et habiletés et l'amélioration de la performance en situation de travail; et les produits («outcomes») qui concernent les profits, la satisfaction du client et la productivité.

Enfin, le modèle de Kaufman et al. (1995) définit six niveaux : les intrants, les processus, les acquisitions (ou apprentissages), la performance (ou comportement), les résultats organisationnels et les résultats sociétaux. Pour les auteurs, les acquisitions et la performance se situent au niveau micro, les résultats au niveau macro et les résultats sociétaux, au niveau «méga».

L'intérêt de ces derniers modèles consiste d'abord à distinguer clairement le processus des extrants et résultats. En effet, la satisfaction ou réaction, bien qu'elle puisse être considérée comme un «résultat» de la formation, porte essentiellement sur le processus de formation, puisqu'on demande aux participants de poser des jugements sur un ensemble d'éléments de celui-ci. De plus, ils ajoutent une évaluation des intrants en amont du processus, qui peut aider à mieux expliquer les produits et résultats.

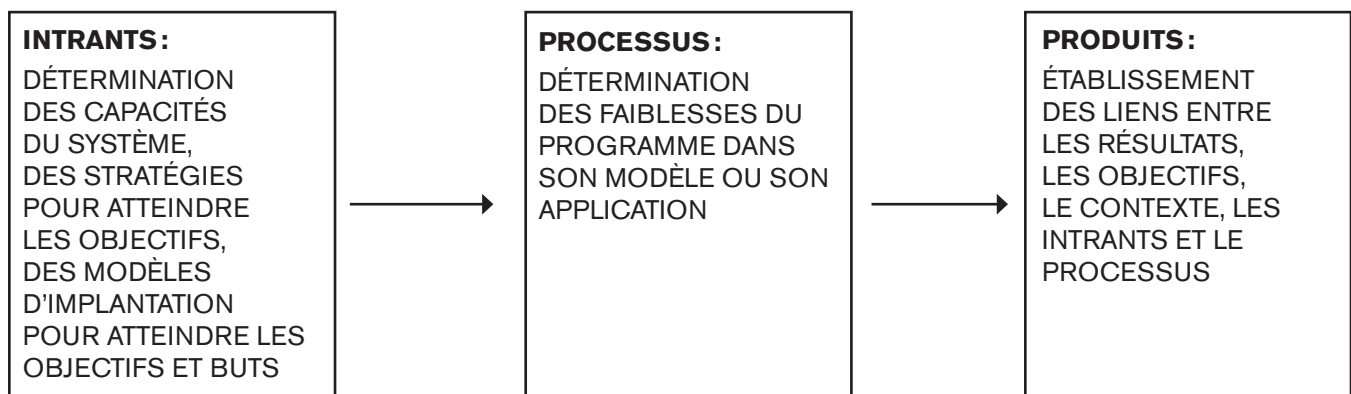
Un seul modèle ajoute de manière explicite le contexte comme étant un élément déterminant à prendre en considération, ce qui est un peu étonnant. En effet, c'est du contexte que proviennent les besoins de formation, celui-ci exerce une influence constante sur le processus de formation et les résultats prendront toute leur valeur de par leur apport final (ou impact) sur le contexte. Peut-

## DIAGRAMME 1

**Le modèle CIPP de Stufflebeam** (Selon Nadeau, 1988)

### CONTEXTE :

L'ENVIRONNEMENT, LES BESOINS,  
LES PROBLÈMES SOUS-JACENTS À CES BESOINS  
DÉTERMINATION DES BUTS ET OBJECTIFS



**TABLEAU 2**

**Analyse comparative de modèles présentant des analogies avec le CIPP de Stufflebeam**

CIPP de Stufflebeam et al. (1971)	CIRO de Warr et al. (1970)	IPO de Bushnell (1990)	Kaufman et al. (1995)
Contexte	Contexte		
Intrants	Intrants	Intrants	Intrants
Processus	Réactions	Processus	Processus
Produit	Produits immédiats	Extrants	Acquisitions
	Produits intermédiaires	Produits	Performance
	Produits terminaux		Résultats
			Résultats sociétaux

on, en effet, procéder à une véritable évaluation d'impact sans en quelque sorte procéder, du moins en partie, à une évaluation de contexte? L'importance de plus en plus grande apportée à l'environnement de travail dans le transfert et la production des résultats, comme on le verra plus loin, devrait amener à terme une meilleure prise en compte de ce critère.

**2.3.3 Deux autres modèles**

Deux autres modèles semblent se distinguer de l'ensemble par leur appui plus prononcé sur la littérature scientifique. Il s'agit des modèles de Holton (1996) et de Kraiger (2002).

Holton (1996) propose une esquisse de modèle d'évaluation s'appuyant sur les résultats de recherches des dernières années et comportant des variables intermédiaires susceptibles d'expliquer l'apprentissage et le transfert. L'évaluation porte essentiellement sur trois niveaux de résultats définis par Kirkpatrick : l'apprentissage, la performance individuelle et les résultats organisationnels. Cependant, l'auteur identifie des éléments susceptibles d'influencer ces résultats. Ceux-ci sont regroupés en trois catégories principales : les éléments facilitateurs (p. ex., l'influence des aptitudes individuelles sur l'apprentissage, l'influence du design du transfert sur la performance); les éléments environnementaux (p. ex., l'influence de la réaction sur l'apprentissage, ou du climat de transfert sur la performance); et les éléments motivationnels (p. ex., la motivation à apprendre ou à transférer). À ces éléments s'ajoute la prise en compte de facteurs pouvant avoir une influence secondaire, tels que les caractéristiques de la personnalité ou les attitudes au travail. L'utilisation de ce modèle peut ainsi mener à ne pas se limiter à la seule évaluation des résultats de la formation, mais aussi à celle de certains facteurs explicatifs susceptibles d'influer sur le niveau de résultat obtenu.

L'intérêt de ce modèle réside dans le fait qu'il attire l'attention du professionnel sur un ensemble de variables souvent négligées qui permettent de comprendre les résultats générés par la formation. Par ailleurs, sa complexité convient mieux, au quotidien, aux chercheurs qu'aux praticiens des entreprises qui sont fortement

susceptibles de retourner rapidement au modèle de Kirkpatrick.

Selon Kraiger (2002), de nombreux praticiens interrogés sur le pourquoi de la réalisation d'une évaluation de la formation expliquent que c'est principalement à cause de l'habitude ou la volonté de poursuivre une démarche mise en place dans l'entreprise par leurs prédécesseurs. Or, la première chose qui doit guider l'évaluation de la formation, c'est d'abord et avant tout les raisons qui l'animent. Il faut savoir par exemple si les résultats vont être utilisés pour prendre des décisions concernant le contenu des formations ou concernant la fonction formation de l'entreprise car des objectifs de l'évaluation vont dépendre le choix des méthodes et des objets à évaluer.

Ceci l'amène à proposer un modèle alternatif d'évaluation de la formation sommairement décrit au tableau 3. Ce modèle identifie trois cibles d'évaluation : le contenu de la formation et son design, les changements chez l'apprenant, et le rendement pour l'organisation. Ces cibles sont en lien direct avec les trois principales raisons d'évaluer qu'il identifie, à savoir la prise de décision, le

**TABLEAU 3**  
**Les cibles de l'évaluation (Kraiger, 2002)**

Cibles	Objets	Méthodes	Raisons
Contenu et design de la formation	Design Livraison Validité	Méthodologies établies d'élaboration de formation  Jugement d'expert Comité conseil Évaluation de la satisfaction	Feedback au formateur
Changements chez l'apprenant	Changements cognitifs Changements comportementaux Changements affectifs	Tests écrits Enquêtes Entrevues Échantillons de travaux	Feedback à l'apprenant Prise de décision
Rendement pour l'organisation	Transfert Résultats Performance	Estimations Enquêtes Analyses bénéfices-coûts	Prise de décision Marketing



feedback et le marketing de la formation. Pour chacune de ces cibles, il précise des objets à évaluer de même que des méthodes appropriées. Par exemple, le rendement peut être évalué par le transfert des apprentissages dans les comportements, les résultats de la formation sur l'organisation ou la performance globale de l'apprenant. Les méthodes proposées à cet effet sont les estimations, les enquêtes ou les analyses bénéfico-coûts.

Notons que ce modèle se centre seulement sur les processus et les résultats de la formation et que différemment du modèle de Holton, il ne suggère pas de variables intermédiaires susceptibles d'expliquer les résultats obtenus. Par ailleurs, il détaille de manière assez explicite les objets à évaluer de même que les méthodologies appropriées. Par exemple, les changements cognitifs font l'objet de 25 produits répartis en 7 catégories et le transfert comporte cinq «*sous-objets*» d'évaluation.

## 2.4 CONCLUSION

Malgré les critiques, le modèle de Kirkpatrick est fortement susceptible de continuer à garder le haut du pavé dans le domaine de l'évaluation de la formation. Les raisons reposent principalement sur sa simplicité et le fait que la chaîne causale qu'il propose allant de la satisfaction à l'impact sur l'organisation rejoint les préoccupations des gestionnaires de la formation dans l'entreprise. Cependant, comme le préconise Phillips (1997), l'ajout d'un cinquième niveau pour le rendement financier s'avère ici bien fondé, ce type de résultats étant de nature différente et commandant une approche méthodologique particulière et exclusivement quantitative.

Néanmoins, les critiques adressées à ce modèle sont pertinentes. Comme le soulignent Alliger et Janak (1989) :

*« Le pouvoir du modèle de Kirkpatrick repose sur sa simplicité et sa capacité à aider les gens à réfléchir sur les critères d'évaluation. En d'autres mots, il fournit un vocabulaire et une ébauche de taxonomie de critères. En même temps, le modèle de Kirkpatrick, de par son vocabulaire facile à adopter et de nombreuses suppositions (souvent implicites), peut mener à des malentendus et des généralisations abusives »* (traduction libre, p.331-332).

Si les praticiens peuvent facilement s'accommoder du caractère sommaire du modèle de Kirkpatrick, il est certain que les chercheurs ou praticiens plus soucieux de comprendre les phénomènes sous-jacents à l'apprentissage et intéressés, par exemple, à l'identification de facteurs pouvant prédire l'impact de la formation, ont avantage à ne pas s'y limiter. L'évaluation de la formation n'est pas une opération en soi : elle s'intègre à l'ensemble de l'ingénierie de la formation. À ce titre, elle devrait donc être une occasion privilégiée non seulement de

constater et mesurer les résultats de la formation, mais aussi fournir des éléments qui alimentent la réflexion des professionnels et décideurs sur les processus de formation en relation avec un contexte donné.

De plus, la logique de ce modèle est limitative. En effet, l'entreprise forme pour obtenir des résultats, mais aussi pour développer une main d'œuvre de plus en plus mobile, et qu'elle souhaite de plus en plus autonome et créative. L'empowerment des travailleurs est une thématique qui gagne en importance pour les entreprises qui se veulent compétitives. Ceci implique donc un rôle de plus en plus actif du travailleur dans la définition de ses objectifs et résultats d'apprentissage. Or, le modèle de Kirkpatrick est principalement issu d'un paradigme où l'entreprise, en fonction de résultats recherchés, décide des apprentissages et comportements que le travailleur devra acquérir, en s'assurant que cela se passe de manière satisfaisante pour lui. Il y a donc place à des modèles d'évaluation plus ouverts où, tout en reconnaissant la nécessité d'arrimer l'apprentissage des travailleurs aux visées corporatives, il y ait aussi place à des apprentissages qui contribuent au développement professionnel des individus.

Quoi qu'il en soit, un des débats actuels importants dans le monde de l'évaluation de programme concerne son utilité pour ceux qui sont concernés par le programme et sa gestion (Sridharan, 2003; Leviton, 2003). En effet, il n'y aura de développement de l'évaluation que dans la mesure où celle-ci contribuera de manière significative au développement de la formation. Patton (1997) propose même un modèle d'évaluation de programme essentiellement centré sur l'utilité. Les quelques recherches menées sur l'utilité de l'évaluation recensées par Fitzpatrick et al. (2004) font ressortir les cinq facteurs suivants contribuant à l'utilisation des résultats de l'évaluation :

- la pertinence de l'évaluation pour les décideurs ou autres parties prenantes;
- l'implication des usagers dans la planification et la phase de rédaction de l'évaluation;
- la réputation ou crédibilité de l'évaluateur;
- la qualité de la communication des résultats de l'évaluation;
- le développement de procédures pour faciliter la mise en œuvre des recommandations.

La solution n'est donc pas seulement dans la sélection du modèle d'évaluation mais aussi dans les processus d'évaluation mis en place et leur organisation assurant que ces conditions seront respectées.

### 3. LES PRATIQUES ACTUELLES D'ÉVALUATION DE LA FORMATION DES ENTREPRISES

Au Canada, l'étude de Betcherman et al. (1997), s'appuyant sur un sondage téléphonique auprès de 2500 entreprises de toutes tailles, constate en un premier temps que 88,7 % d'entre elles évaluent d'une manière ou d'une autre leur formation. Plus de la moitié (54,3 %) auraient recours à une évaluation informelle par le formateur ou le superviseur. Environ le tiers (35,5 %) procéderait à une analyse formelle de l'impact de la formation. Par ailleurs, les auteurs conviennent, sur la base d'études de cas menées parallèlement, que leurs données surévaluent la rigueur des pratiques d'évaluation des entreprises.

La dernière enquête auprès des employeurs assujettis à la loi 90 du Québec (Lapierre et al., 2005) arrive aux constats suivants : en 2002, 49 % des employeurs qui financent la formation évaluent toutes ou certaines des activités de formation, comparativement à 59 % en 1998. Entre 11 et 13 % des employeurs assujettis évaluent fréquemment leurs activités à un ou plusieurs niveaux du modèle de Kirkpatrick; en 1998, c'était le cas de 20 % des employeurs. La baisse des pratiques évaluatives serait, selon les auteurs, attribuable à la hausse des activités d'entraînement à la tâche qui seraient peu souvent évaluées. Notons toutefois que Dunberry (2006) fait le constat auprès d'entreprises québécoises performantes que c'est surtout l'entraînement à la tâche qui fait l'objet de leurs pratiques évaluatives.

Le tableau 4 présente les pratiques d'évaluation des entreprises américaines selon les rapports annuels de l'« *American Society for Training and Development* » (ASTD). Ce sondage distingue trois catégories d'entreprises : celles faisant appel à son service de « *benchmarking* » ou « BMS » et qui représentent l'échantillon le plus important de grandes entreprises (p. ex, en 2004 : n = 281; nombre moyen d'employés : 8 213); celles qui participent au forum de « *benchmarking* » ou « BMF », qui sont moins nombreuses et comptent parmi les 500 plus

grandes entreprises (p. ex, en 2004 : n = 24; nombre moyen d'employés : 57 868); et les meilleures entreprises ou « BEST » qui ont été officiellement honorées pour leurs efforts de promotion de la formation (p. ex, en 2004 : n = 29; nombre moyen d'employés : 45 870). Les données du tableau reflètent les pratiques du groupe le plus représentatif, à savoir les entreprises « BMS ». Ce tableau permet de constater que pour l'ensemble des études, ces pratiques diminuent systématiquement en passant aux niveaux supérieurs d'évaluation. On évalue surtout la satisfaction, puis à des degrés divers les apprentissages, mais peu les comportements en poste de travail et les résultats sur l'organisation. Et cette tendance s'avère relativement stable au fil des ans.

De fait, Twitchell et al. (2000) constatent que sur une période de 40 ans, le nombre d'évaluations menées dans les entreprises américaines a somme toute peu évolué : quelques gains ont été réalisés en termes d'entreprises ayant des pratiques d'évaluation en particulier aux niveaux 1 et 2 entre 1968 et 1989, mais peu depuis; et très peu de gains aux niveaux 3 et 4 qui sont pourtant les niveaux cruciaux pour démontrer l'apport de la formation à l'entreprise.

Les rapports de l'ASTD permettent aussi de constater qu'il existe un certain écart entre les grandes entreprises « BMS » et les très grandes entreprises « BMF » aux trois premiers niveaux. Ainsi, Sugrue et Rivera (2005) constatent que 91,3 % de ces dernières (n=18) évaluent la satisfaction, 53,9 % les apprentissages, 22,9 % les comportements et 7,6 % les résultats des programmes. L'évaluation du retour sur l'investissement a été calculée par 2 % de ces entreprises.

Quant aux meilleures entreprises, dont la taille est celle de très grandes entreprises, Sugrue et Rivera (2005) font le constat suivant :

**TABLEAU 4**  
**Pourcentage des entreprises américaines utilisant des méthodes d'évaluation selon le niveau**

États-Unis	Niveau 1 Satisfaction	Niveau 2 Apprentissage	Niveau 3 Comportement	Niveau 4 Résultats
1999 (Selon Sugrue, 2003)	77 %	38 %	14 %	7 %
2000 (Selon Sugrue, 2003)	78 %	36 %	9 %	7 %
2001 (Selon Sugrue, 2003)	91 %	32 %	17 %	9 %
2002 (Selon Sugrue, 2003)	75 %	41 %	21 %	11 %
2004 (Selon Sugrue et Kim, 2004)	74 %	31 %	14 %	8 %

TABLEAU 5

**Les pratiques d'évaluation de la formation des entreprises québécoises finançant la formation et assujetties ou non à la loi 90** (Selon Lapierre et al., 2005)

Québec	Niveau 1 Satisfaction	Niveau 2 Compétences en fin de formation	Niveau 3 Comportement	Autres façons
1998	69 %	78 %	68 %	
2002	61 %	71 %	59 %	20 %

« Toutes les meilleures entreprises ont en place des systèmes pour mesurer et faire rapport sur les activités et l'impact de la fonction formation dans l'entreprise. Des données des systèmes de gestion de l'apprentissage sont mises en relation avec des données recueillies auprès des unités d'affaires et présentées selon des cartes de pointage ou tableaux de bord, et mettent l'accent sur l'investissement et les aspects opérationnels de la fonction formation. Certaines des meilleures entreprises ont créé des cadres pour recueillir et mettre en évidence la valeur de l'apprentissage, incluant des indicateurs tels que les niveaux de compétence et la rétention. Plusieurs des meilleures entreprises ont inclus des indicateurs convenus de succès pour des initiatives particulières d'apprentissage, tels que l'augmentation des flux d'information au sein de l'organisation, ou la vitesse de prise de décision. Elles ont aussi inclus dans leurs rapports des indications sur l'alignement des activités d'apprentissage avec les besoins d'affaires. » (p. 15).

Différemment de ces études, Lapierre et al. (2005) constatent, suite à leur enquête par questionnaire que, parmi les employeurs qui évaluent la formation en 2002<sup>2</sup>, 71 % d'entre eux disent évaluer les compétences à la fin de la formation, 61 % la satisfaction des employés formés et 59 % le transfert des compétences acquises en poste de travail. En 1998, ces pourcentages étaient respectivement de 78 % pour les compétences en fin de formation, 69 % pour la satisfaction et 68 % pour les comportements. Ces pratiques formelles sont complétées d'autres pratiques formelles (p. ex : examens ou tests par des organismes formateurs) ou informelles (p. ex : « évaluation verbale », « contrôle du travail effectué », etc.). Si on applique ces pourcentages à l'ensemble des employeurs de 2002 et non seulement aux 49 % qui disent évaluer la formation, on aurait alors 35 % des employeurs qui évalueraient les compétences en fin de formation, 30 % la satisfaction, et 29 % le transfert en poste de travail.

Il faut retenir l'écart entre les constats de Lapierre et al. (2005) et les études américaines en terme d'importance relative des pratiques d'évaluation. En effet, selon les données de cette enquête, l'évaluation la plus fréquente serait celle des compétences en fin de formation (ou

« apprentissages » selon le modèle de Kirkpatrick) et non pas de la satisfaction. De plus, un pourcentage nettement plus élevé des employeurs (29 % vs 9 à 14 %) évaluerait le transfert en poste de travail. Dans la mesure où ces résultats ne prêtent pas aux mêmes réserves que ceux de Betcherman et al. (1997), à savoir qu'il y a surestimation des pratiques, ils témoignent d'un meilleur équilibre dans le choix de ce qui est évalué, qui devrait permettre aux employeurs québécois de mieux juger des résultats de la formation.

Dunberry (2006), dans sa recension des pratiques d'évaluation de la formation dans 12 entreprises québécoises performantes, arrive à des constats qui vont dans le même sens. Les entreprises québécoises visitées « évaluent leur formation, principalement lorsque celle-ci est reliée de près aux activités productives. Différemment de ce que laissent croire les recensions américaines, les efforts d'évaluation d'une forte proportion d'entre elles se situent principalement sur les apprentissages et les comportements des formations à l'emploi. Par ailleurs, ces dernières accordent aussi une importance certaine à l'évaluation de la satisfaction. » (p. 50).

Notons que selon Lapierre et al. (2005), une forte proportion d'employeurs québécois (56 % des employeurs qui ont financé la formation, donc 52 % de l'ensemble des employeurs assujettis à la loi) ont décerné une attestation de compétences acquises suite à la formation, bien que la loi ne les y oblige pas; 23 % des employeurs qui financent la formation l'ont même fait fréquemment. Cette pratique répandue au Québec, et qui dépasse les exigences de la loi, pourrait être un incitatif à instaurer des pratiques d'évaluation formelles des compétences en fin de formation.

À ces pratiques d'évaluation formelles s'ajoutent les pratiques informelles d'évaluation de la formation. Dunberry (2006) recense de telles pratiques dans 11 des 12 entreprises québécoises performantes visitées. Celles-ci compléteraient les pratiques formelles, mais aussi les précéderaient, pouvant souvent être mises à profit dans l'élaboration de ces dernières. Malheureusement, la littérature scientifique semble bien discrète sur ces pratiques.

<sup>2</sup> Qui, rappelons-le, comptent pour 49 % des répondants.



## 4. LES PRINCIPAUX FACTEURS QUI INFLUENCENT LES PRATIQUES D'ÉVALUATION DE LA FORMATION

### 4.1 LA TAILLE DE L'ENTREPRISE

De nombreuses études (Sarnin, 1998; Garand & Fabi, 1996; Paradas, 1993a) font état des multiples contraintes qui confèrent une spécificité au développement de la formation dans les PME. De ces études, il faut certainement retenir que le degré de formalisation des pratiques de formation, bien qu'il constitue un enjeu stratégique, est souvent sommaire et qu'il croît avec la taille de la PME (Paradas, 1993b). Garand et Fabi (1996) précisent que moins du tiers des PME ont une politique de formation, la très grande majorité n'a aucune procédure de formation, et moins du quart possèdent un plan de formation. Ceci va de pair avec une fonction formation souvent inexistante dans la PME, surtout celles de moins de 150 employés (Mahe de Boislandelle, 1988 in Paradas, 1993).

Cette situation se répercute dans les pratiques d'évaluation de la formation que Paradas (1993a) a analysées dans un échantillon de 39 PME françaises. L'auteure constate que «54 % des PME n'ont aucune méthode d'évaluation. Seulement 5 entreprises ont élaboré des questionnaires, 1 une grille et 2 procèdent par entretiens «ouverts». Les autres «méthodes» évoquées sont le «nez», les réunions avec les chefs de service, la définition des postes, les rapports de stages.» (p.93) Par ailleurs, elle constate aussi un très fort pourcentage de responsables désirant améliorer la qualité de leurs pratiques d'évaluation de la formation.

Pour leur part, Garand et Fabi (1996) sur la base de 45 études empiriques regroupant plus de 5000 PME constatent que :

*«...l'évaluation des retombées de formation demeure presque totalement subjective(...) La plupart des démarches d'évaluation des résultats consistent en de simples constats(...) Ce processus demeure surtout informel avec peu de procédures et quelques observations ponctuelles en TPE(...). En MGE, cette évaluation devient plus formelle, comparable à l'appréciation du personnel avec des modes quantitatifs et qualitatifs, des formulaires et des rencontres structurées. Cependant, la distinction se fait difficilement en PME entre l'évaluation de l'individu (rendement) et l'évaluation des résultats de formation.» (p. 34).*

Enfin, les données récentes de Lapierre et al. (2005) viennent confirmer l'effet de la taille de l'entreprise sur les pratiques d'évaluation de la formation. En effet, les chercheurs constatent que les entreprises qui évaluent toutes ou certaines activités de formation se retrouvent en plus

forte proportion chez celles qui comptent 100 employés et plus versus celles qui en comptent 9 ou moins (68 % contre 40 %) ; et chez celles qui comptent plus de 50 employés versus celles qui en ont moins (60 % contre 43 %). Les entreprises qui évaluent les compétences acquises en fin de formation sont en plus forte proportion celles qui ont une masse salariale de plus de 1 million \$ (77 %, versus 64 % pour les entreprises dont la masse salariale se situe entre 250 000 \$ et 499 999 \$; et 69 % pour les entreprises dont la masse salariale se situe entre 500 000 \$ et 999 999 \$). Et il existe un lien direct entre le niveau de masse salariale et la proportion des entreprises qui évaluent la satisfaction des employés formés : respectivement 52 %, 61 % et 68 % par ordre croissant de masse salariale.

De toute évidence, l'effet de la taille ne permet pas une simple transposition des pratiques d'évaluation des GE vers les PME, mais nécessite la prise en compte d'une réalité essentiellement différente.

### 4.2 LA CULTURE DE L'ENTREPRISE

On sait que la décision de former s'appuie souvent sur des pratiques plus ou moins formelles d'évaluation et des critères et indicateurs plus larges que le rendement sur l'investissement ou l'analyse bénéfices-coûts qui tentent de mesurer l'apport financier de la formation (Betcherman et al., 1997; Betcherman et al., 1998). On attribue souvent ce recours à d'autres critères et indicateurs à la «culture de la formation» de l'entreprise sans bien connaître pour autant ceux qui sont alors utilisés par les décideurs (Billet & Cooper, 1997; Leckie et al., 2001; Figgis et al., 2001). Ceci suggère que les pratiques d'évaluation sont multiples, formelles, mais aussi informelles, et qu'une certaine connaissance de la philosophie de gestion et de la culture de l'entreprise est essentielle à leur compréhension.

Lewis et Thornhill (1994) insistent sur l'effet d'une culture organisationnelle défensive sur les pratiques d'évaluation, où l'on préfère ne pas poser certaines questions dont on craint la réponse. Ils soulignent que l'évaluation, et surtout l'évaluation du rendement, comporte en elle-même le risque de remettre en cause des décisions prises par les niveaux hiérarchiques supérieurs ou par ceux-là mêmes qui sont en mesure d'initier cette opération. Ceci rejoint aussi un constat de McIntyre (1994).

En fait, la culture de l'organisation est certainement un des facteurs majeurs qui explique pourquoi il se fait ou non de l'évaluation. En effet, les pratiques évaluatives sont aussi liées, entre autres, à la pension de l'or-

ganisation à admettre facilement des erreurs ou des divergences d'opinions, ou simplement à reconnaître la pertinence d'évaluer la formation. Ainsi, Betcherman et al (1997) notent que les entreprises de type organisation apprenante sont les plus susceptibles de procéder à l'évaluation du rendement de la formation, tout comme d'avoir effectué une analyse formelle des besoins de formation. Comme le proposent Lewis et Thornhill (1994), il devient donc pertinent d'accompagner la promotion de l'évaluation d'un changement de culture organisationnelle pour lequel ils donnent quelques indications. Et réciproquement, il est aussi important de prendre en compte la culture organisationnelle qui prévaut dans les organisations qui pratiquent l'évaluation de la formation.

#### 4.3 LES EXIGENCES MÉTHODOLOGIQUES

Comme le démontrent certains auteurs (Arvey & Cole, 1989; Phillips, 1997), une évaluation rigoureuse visant à démontrer des effets attribuables à la formation comporte des exigences en termes de design que peu d'entreprises peuvent rencontrer dans leurs pratiques courantes. En effet, ces auteurs préconisent à tout le moins l'utilisation d'un groupe contrôle, les deux groupes étant évalués avant et après la formation. De plus, comme le souligne Phillips (1997), le groupe contrôle devrait être constitué au hasard de sujets semblables à ceux du groupe qui recevra la formation. À cela s'ajoute évidemment un ensemble d'exigences portant sur l'élaboration et la validation des instruments de mesure qui seront utilisés.

Et même avec une certaine rigueur, certains problèmes sont susceptibles d'apparaître. Selon Goldstein et Ford (2002), il devient de plus en plus évident que l'utilisation de design expérimentaux, dont l'utilisation de groupes contrôle, peut influencer sur le comportement des participants au point de compromettre les résultats des programmes de formation. Les auteurs proposent une revue de ces menaces à la validité interne et externe des évaluations de la formation qui doivent être présentes à l'esprit de l'évaluateur. Notons, entre autres :

- les événements extérieurs à la formation qui peuvent survenir entre le pré-test et le post-test et qui peuvent, par exemple, accroître la sensibilité des apprenants au contenu; par exemple, un accident grave pendant un cours sur la sécurité;
- l'effet du contenu du pré-test, qui attire l'attention des participants sur certains aspects de la formation et peut les prédisposer à une bonne performance au post-test;
- le mode d'assignation des participants au groupe expérimental et au groupe contrôle; idéalement, les sujets devraient être assignés à la fois au hasard et en les pairant par rapport à certains paramètres jugés importants;

- la mortalité expérimentale, qui peut faire abandonner la formation à des sujets ayant eu un score faible au pré-test;
- la tendance des gestionnaires ou formateurs à compenser les sujets du groupe contrôle lorsque l'effet du programme s'avère très bénéfique aux sujets du groupe expérimental;
- l'effet parfois démoralisant sur certains sujets d'être assignés à un groupe contrôle qui peut les amener à sous performer lorsqu'ils sont évalués;
- et ainsi de suite.

Il n'est donc pas étonnant que depuis plusieurs années, on constate que les évaluations de la formation, de quel que niveau qu'elles soient, ne répondent pas à toutes les exigences de la rigueur scientifique (Carnevale & Schulz, 1990; Tannenbaum & Yukl, 1992; Dionne, 1996). On attribue cette situation à diverses raisons, dont l'expertise limitée des évaluateurs et les contraintes inhérentes au contexte dans lequel se situe l'évaluation.

En ce sens, il semble se dégager de la littérature (Carnevale & Schulz, 1990; Benabou, 1997; Kirkpatrick, 1998; Moy & McDonald, 2000; Goldstein & Ford, 2002) une certaine tendance à promouvoir des pratiques d'évaluation simples et adaptées qui bien que perfectibles, permettent néanmoins un meilleur éclairage en vue d'accroître l'efficacité et l'efficacé des pratiques de formation. Par exemple, on reconnaît que même sans groupe contrôle, la validité de l'évaluation présente pour certains chercheurs des risques acceptables dans la mesure où «*les programmes de formation sont de courte durée, lorsqu'il est peu probable que les connaissances transmises dans le cadre du cours soient accessibles autrement et lorsque les questions posées aux stagiaires sont bien ciblées*» (Haccoun, Jeanrie, & Saks, 1997, p. 109).

Bien que l'évaluateur doive rester tout à fait conscient des limites de l'évaluation de la formation qu'il sera en mesure de réaliser et interpréter avec prudence les résultats obtenus en conséquence, il n'en demeure pas moins qu'il vaut mieux évaluer de manière imparfaite que de ne pas évaluer du tout.

## 5. L'ÉVALUATION DE LA SATISFACTION

L'évaluation de la satisfaction est une des, sinon la pratique d'évaluation de la formation la plus répandue dans les entreprises. Il existe donc un souci de savoir si les participants ont ou non réagi favorablement à la formation qui leur a été proposée. Implicitement, et à des degrés variables, on considère qu'une bonne formation générant des apprentissages devrait aussi provoquer des réactions favorables, et qu'une formation provoquant des réactions défavorables devrait être remise en cause, ou à tout le moins améliorée.

Mais en quoi consiste la satisfaction? Quelles sont les raisons qui peuvent inciter à l'évaluer? Comment peut-on être sûr de bien l'évaluer? Et quelles sont les limites inhérentes à ce type de pratiques?

### 5.1 EN QUOI CONSISTE L'ÉVALUATION DE LA SATISFACTION

« Évaluation de la satisfaction », « évaluation des réactions » et « évaluation à chaud » sont les trois expressions les plus courantes pour désigner une évaluation faite à la fin d'une formation et visant à recueillir ce que ressentent des stagiaires suite à la formation qu'ils ont reçue. Ce type d'évaluation se fait généralement par l'utilisation d'un questionnaire.

À la différence du terme « réaction », le terme « satisfaction » appuie le point de vue de Kirkpatrick (1998) selon lequel il est important que les stagiaires réagissent favorablement à une formation. L'expression « évaluation à chaud » insiste sur le fait de laisser s'exprimer les stagiaires en toute fin de formation et ce de manière plus ou moins guidée (Barzucchetti & Claude, 1995).

On peut ajouter en complément qu'une autre définition de la satisfaction se place, elle, du point de vue de l'entreprise cliente d'un prestataire de formation (Mulder, 2001). Elle est employée dans l'expression « *satisfaction totale* » et désigne le vécu d'une entreprise cliente vis-à-vis du service fourni par un organisme de formation auquel elle a fait appel. Elle couvre « *la satisfaction vis-à-vis du management du projet de formation dans sa totalité, de la préparation du projet, à sa mise en place et jusqu'aux activités de suivies* » (traduction libre, Mulder, 2001, p. 325).

### 5.2 POURQUOI ÉVALUER LA SATISFACTION?

Quelles sont les raisons qui peuvent expliquer l'intérêt pour l'évaluation de la satisfaction?

Selon Kirkpatrick (1998) la satisfaction peut être considérée comme un critère important de l'évaluation à deux égards. D'une part on peut penser que l'opinion que les stagiaires vont communiquer une fois la formation terminée va en partie influencer l'opinion des décideurs quant à l'intérêt de maintenir ou pas le programme. D'autre part, on peut croire que si les stagiaires ne réagissent pas positivement à la formation, ils ne seront pas motivés pour apprendre. Si la première de ces raisons apparaît toujours valable, on sait par ailleurs que l'effet causal de la satisfaction sur l'apprentissage n'a pas été démontré (Alliger & Janak, 1989; Alliger et al., 1997). Néanmoins, comme le soulignent Goldstein et Ford (2002), il serait difficile de défendre le maintien d'un programme de formation qui provoque systématiquement des réactions négatives de la part des participants.

Selon Phillips (1997), c'est la satisfaction des « clients » qui est recherchée à travers tout processus d'évaluation, ceux-ci pouvant être séparés en deux groupes : d'un côté, les participants directs à la formation; de l'autre, les gestionnaires et managers qui les encadrent et les dirigeants de l'entreprise. Ainsi, les niveaux 1 (satisfaction) et 2 (apprentissage) de l'évaluation fournissent de l'information permettant de juger si les stagiaires sont satisfaits tandis que les niveaux 3 (comportements), 4 (résultats) et 5 (retour sur l'investissement) fournissent des informations attendues par les supérieurs hiérarchiques afin de leur permettre de décider si le programme doit être poursuivi et de justifier de futurs investissements de formation.

Selon Kraiger (2002), les réactions des participants peuvent être utilisées comme « *méthode primaire de mesure* » par les concepteurs pour évaluer le contenu de la formation, son design et son mode de livraison. Pour ce faire, il faudra procéder à une sélection de questions judicieuses en fonction de dimensions que l'on souhaite évaluer et des objectifs poursuivis. Il faut bien noter que le recueil des réactions des participants est une méthode parmi d'autres pour prendre des décisions concernant la formation et ses modalités de livraison. Par exemple, il est aussi possible d'avoir recours à des panels d'expert pour juger de la qualité du matériel.

Selon Barzucchetti et Claude (1995), le fait de donner aux stagiaires un rôle d'acteur et de leur permettre de s'exprimer sont des raisons de plus d'effectuer ce type d'évaluation « à chaud ». Elle permet de donner des indications sur le degré d'adhésion des stagiaires même s'il est important d'avoir bien conscience des limites de ce type d'indicateur.

À la question du début, on peut donc répondre que l'évaluation de la satisfaction permet de manière très facile et peu coûteuse d'avoir un feed-back des apprenants sur le processus de formation qui leur est proposé. Il témoigne aussi d'un souci des apprenants et de leur point de vue, ce qui est tout à fait de bon aloi. Et ce feed-back profite d'une part aux concepteurs et prestataires de formation qui peuvent ainsi avoir des indications sur comment améliorer la formation, et d'autre part aux décideurs qui peuvent savoir si la formation correspondait à leurs attentes.

### 5.3 PRATIQUES ET TENDANCES ACTUELLES D'ÉVALUATION DE LA SATISFACTION

Les pratiques les plus courantes concernant l'évaluation de la satisfaction des stagiaires consistent à leur fournir un questionnaire immédiatement après la formation. Les aspects pouvant être abordés dans ces questionnaires sont nombreux et la difficulté consiste à les sélectionner. Cette sélection doit s'opérer en fonction de deux éléments essentiels : l'activité à évaluer et les buts recherchés par l'évaluation (Lee & Pershing, 1999).

Ces derniers auteurs proposent 10 dimensions que peuvent couvrir les questionnaires de satisfaction : les objectifs et le contenu du programme; le matériel du programme; les méthodes et technologies de livraison; le formateur ou facilitateur; les activités de formation; la durée du programme; l'environnement de formation; les attentes à l'égard du transfert; la logistique; et l'évaluation d'ensemble de la formation. Les auteurs suggèrent de terminer le questionnaire en demandant aux participants de formuler des recommandations pour l'amélioration du programme. Cet article précise le contenu de ces dimensions et fournit des exemples de questions qui peuvent être posées. Il est évident que le questionnaire n'a pas nécessairement à couvrir toutes ces dimensions, mais bien celles qui s'avèrent pertinentes aux yeux de ceux qui souhaitent l'évaluation.

Pour leur part, Morgan et Casper (2000), s'appuyant sur une analyse factorielle des réponses de 9000 sujets à 32 questions de satisfaction, identifient 6 dimensions composant la satisfaction : la satisfaction à l'égard de l'instructeur, de la gestion de la formation, du processus d'évaluation des apprentissages, de l'utilité du programme de formation, du matériel du cours et de la structure du cours. On peut donc aussi avoir recours à ces dimensions afin de bâtir un questionnaire qui se veut complet.

Kirkpatrick (1998) donne quelques conseils pratiques pour la réalisation de ces questionnaires :

- déterminer ce que l'on veut savoir (les sujets sur lesquels on veut avoir un retour, par exemple le calendrier de la formation, les repas, les supports audiovisuels, les cas pratiques, les exercices, etc...);

- concevoir un questionnaire qui va quantifier la satisfaction;
- encourager les commentaires écrits et les suggestions;
- obtenir 100% de réponses immédiates;
- obtenir des réponses honnêtes (en permettant les réponses anonymes par exemple);
- développer des standards acceptables (ils peuvent varier selon l'importance accordée à ce qui est évalué);
- mesurer la satisfaction par rapport à ces standards et réaliser des actions appropriées;
- communiquer la satisfaction de façon adéquate.

Pour que ces questionnaires soient le plus fiables possible, il faudrait, selon certains auteurs (Covert, 1984; Swisher, 1980 in Pershing & Pershing, 2001), qu'ils comportent :

- un texte d'introduction comprenant les buts de l'évaluation, un appel à la coopération et les procédures assurant la confidentialité et l'anonymat;
- un texte de conclusion qui explique quoi faire une fois le questionnaire complété et qui remercie les participants;
- des directives pour répondre aux questions, selon la nature des questions (ouvertes ou fermées) et l'échelle employée;
- une présentation claire et professionnelle incluant la qualité de la langue et de la reprographie et l'espacement entre les questions;
- un texte concernant les règles de confidentialité et d'anonymat déjà annoncées dans l'introduction.

Pershing et Pershing (2001) ont listé quelques problèmes de conception des questionnaires de satisfaction les plus courants, outre l'absence des éléments cités ci-dessus. En voici quelques-uns, pour lesquels il faut faire preuve de vigilance :

- questions visant à évaluer deux choses en même temps;
- questions subjectives et biaisées dans leur formulation, incitant les répondants à fournir une certaine réponse;
- questions trop longues ou complexes;
- questions portant sur ce que les autres savent, pensent ou ressentent et pour lesquelles le stagiaire ne peut avoir la réponse;
- catégories de réponses inadéquates en regard de la question;



- échelle de réponse mal balancée, favorisant des réponses négatives ou positives;
- échelles mélangeant des chiffres et des mots, plutôt que de se limiter à l'un ou à l'autre;
- échelle de réponse non conventionnelle, avec laquelle les répondants ont peu de chances d'être familiers.

#### 5.4 LIMITES DES PRATIQUES D'ÉVALUATION DE LA SATISFACTION

Les pratiques d'évaluation de la satisfaction ont certaines limites qu'il est essentiel de connaître si l'on souhaite faire un bon usage des données qu'elles génèrent.

##### 5.4.1 Les limites liées à l'administration de questionnaires de satisfaction.

Reitz (1997) formule quatre critiques à l'égard de cette pratique :

- les réponses sont essentiellement subjectives et deux personnes peuvent avoir une appréciation très différente tout en ayant toutes les deux raison; il faut donc faire preuve de prudence lorsque l'on souhaite généraliser à partir de ces résultats;
- les données s'appuient sur des faits passés dont le répondant peut avoir un souvenir inégal, surtout si la formation s'étend sur une longue période;
- les questionnaires sont souvent répondus à la hâte en fin de formation, ce qui peut contribuer à l'erreur de mesure;
- malgré l'anonymat, les répondants peuvent appréhender des représailles et donner des réponses faussement positives.

Kraiger (2002) nous rappelle les limites de ces réactions émises par les participants. Par exemple, les stagiaires novices ne sont pas forcément les mieux indiqués pour juger du niveau de connaissance d'un formateur ou du temps alloué pour la pratique. Pour ce faire, le recours à l'avis d'experts internes à l'entreprise peut s'avérer être une meilleure solution.

Deux études, celles de Lee (1998 in Pershing & Pershing, 2001) et de Pershing et Pershing (2001), portant sur la qualité de ces questionnaires sont loin d'être encourageantes. La première a porté sur 77 questionnaires de satisfaction d'entreprises privées. Elle conclut que seulement la moitié des questionnaires peut être jugée de qualité professionnelle. La seconde a porté sur 50 questionnaires régulièrement administrés par une école de médecine. Les résultats sont encore plus décevants : les auteurs notent plusieurs problèmes dans 60 % des cas et estiment que 92 % de ces questionnaires n'étaient pas de qualité professionnelle. En

conséquence, Pershing et Pershing (2001) suggèrent fortement que des évaluations approfondies des questionnaires de satisfaction en usage soient réalisées afin de pouvoir en corriger les faiblesses. Cela serait utile à la fois aux formateurs, aux stagiaires, aux évaluateurs et aux concepteurs des formations. Ils suggèrent que des règles de conception homogènes soient respectées.

Enfin, est-il nécessaire de dire que le questionnaire écrit n'est pas la seule façon d'évaluer la satisfaction et que cette forme comporte ses propres limites? En effet, toute formation a ses particularités qui sont souvent nivelées par cette formule, surtout si on utilise de manière extensive un seul questionnaire maison. Et il n'est pas toujours facile d'exprimer son véritable point de vue par le biais de questions préétablies et quelques lignes ouvertes de commentaires. Les échanges libres entre participants et formateurs en fin de formation, par exemple, ont certainement encore toute leur pertinence et risquent parfois moins de masquer certaines réalités à cause de limitations techniques. Une fois les buts de l'évaluation clarifiés et les dimensions à évaluer choisies, il est certainement indiqué de se questionner sur les meilleures façons d'aller chercher l'information requise.

##### 5.4.2 La satisfaction telle que pratiquée corrèle mal avec les autres niveaux de l'évaluation.

L'évaluation de la satisfaction est peu coûteuse et facile à obtenir de tous les participants immédiatement à la fin de la formation. Compte tenu des moyens limités dont disposent les services de formation, on peut être tenté de croire que c'est là la façon la plus efficace d'évaluer la formation dans son ensemble. Autrement dit, qu'il serait possible de prédire si le stagiaire a appris quelque chose, si son comportement au travail va changer, ou si des résultats vont être obtenus au niveau de l'entreprise en ne mesurant que le niveau de la satisfaction. Si c'était le cas, c'est-à-dire si des corrélations positives étaient observées entre ces niveaux, l'application pratique deviendrait très intéressante pour l'entreprise, car des économies de moyens pourraient être faites.

Cette corrélation entre la satisfaction des apprenants et les apprentissages, comportements et résultats attendus de la formation est un sujet qui revient périodiquement. Plusieurs chercheurs ont démontré que cette corrélation était, somme toute, inexistante (Alliger & Janak, 1989; Warr, Allan, & Birdi, 1999; Holton, 1996; Boyle & Crosby, 2004; Rough, 1994 in Tamkin et al., 2002). Rappelons que Alliger et Janak (1989), sur la base d'une méta-analyse de plusieurs articles n'ont trouvé qu'une corrélation très faible de la satisfaction avec l'apprentissage ( $r = ,07$ ) et le comportement ( $r = ,05$ ). Des recherches menées dans des écoles ont par exemple montré que c'est seulement dans la mesure où les apprenants ont une expérience désagréable de l'apprentissage qu'ils commencent à apprendre (Remmers, Martin & Elliott,

1949; Rodin & Rodin, 1972 in Alliger & Janak, 1989). Rough (1994 in Tamkin et al., 2002) indique de son côté qu'avec un apprentissage qui touche à de dimensions plus profondes de la personne (ou « transformationnel ») par opposition un apprentissage plus ordinaire (ou « linéaire »), il y a souvent une période de frustration et de résistance, signe que les personnes commencent justement à apprendre, et qui est susceptible de générer de l'insatisfaction. Un plus récent article de Alliger et al. (1997), dont il est fait état précédemment, en arrive à un certain niveau de corrélation des réactions reliées à l'utilité de la formation avec les connaissances en fin de formation ( $r = ,26$ ) et le transfert en poste de travail ( $r = ,18$ ), mais qui demeurent trop modestes pour parler d'un lien évident.

Un récent article de Tan, Hall, et Boyce (2003) en arrive à la conclusion que ce sont les réactions affectives négatives (ne pas avoir apprécié la formation) qui ont mieux prédit l'apprentissage. Les auteurs font l'hypothèse que les participants qui ont déjà des acquis avant la formation ont évalué celle-ci plus durement parce qu'elle n'a pas rencontré leurs attentes très élevées. Ces mêmes participants sont ceux qui auraient les scores les plus élevés au test d'acquisition de connaissances déclaratives qui mesure l'apprentissage. Notons que l'on peut ici se demander si les connaissances qui leur ont permis d'avoir un score élevé au test d'apprentissage avaient été acquises lors de la formation ou avant. En d'autres termes, les mêmes résultats pourraient simplement signifier qu'une formation qui nous apprend peu provoque des réactions négatives.

Selon Mathieu, Martineau, et Tannenbaum (1993), la satisfaction agit plus comme un modérateur et un médiateur entre les variables influençant l'apprentissage. En ce sens, il est normal de ne pas trouver de corrélation entre celle-ci et l'apprentissage, mais elle demeure néanmoins importante et doit être prise en compte.

En somme, l'évaluation de la satisfaction a sa place, mais ne saurait être utilisée comme substitut à l'évaluation des apprentissages, des modifications de comportement et des résultats sur l'organisation attribuables à la formation.

#### 5.4.3 La satisfaction, un construit multidimensionnel

Selon Warr et Bunce (1995 in Kraiger, 2002), la satisfaction des participants comporte le plaisir, l'utilité perçue et la difficulté perçue. Alliger et al. (1997) distinguent la dimension affective de la dimension utilitaire de la satisfaction. On a vu précédemment que les travaux de Morgan et Casper (2000) avaient permis à ces auteurs d'identifier six facteurs constitutifs de la satisfaction, notamment des éléments liés à la personne et au contexte de la formation. Ces derniers auteurs considèrent que

l'utilisation des deux seuls facteurs d'affectivité et d'utilité est insuffisante à traduire la complexité de l'information recueillie dans les questionnaires de satisfaction. Tan et al. (2003) proposent pour leur part de distinguer les dimensions affectives et cognitives de la satisfaction.

Ceci revient donc à dire que la satisfaction ne constitue pas, dans le cas de la plupart des questionnaires de satisfaction, une seule variable, mais un amalgame de plusieurs variables et qu'elle doit donc être considérée dans toute cette complexité.

Malgré les très faibles corrélations constatées jusqu'à maintenant entre la satisfaction et les apprentissages, comportements et résultats de la formation, certains chercheurs (Tannenbaum, Mathieu, Salas, & Cannon-Bowers, 1991; Mathieu et al., 1993; Morgan & Casper, 2000; Tan et al., 2003) défendent le point de vue selon lequel des corrélations plus substantielles pourraient être observées à partir du moment où la satisfaction serait considérée comme multidimensionnelle. Mais il est évident que, dans l'état actuel des recherches sur cette question, on est encore loin du moment où les praticiens pourront miser sur l'évaluation de la satisfaction pour prédire les autres résultats de la formation.

\* \* \*

Si elle ne permet pas de récolter des informations fiables sur le niveau d'apprentissage et de transfert des apprentissages à cause de toutes les critiques que nous venons de voir, l'évaluation de la satisfaction est toutefois un bon moyen de comprendre « *ce qui s'est passé* » pendant la formation (Barzucchetti & Claude, 1995). Elle peut permettre de « *proposer des éléments de diagnostic quant à l'efficacité du dispositif de formation et donc fournir des idées d'amélioration de celui-ci* » (ibid, p. 72). L'évaluation de la satisfaction, tel qu'elle est souvent pratiquée actuellement, est donc davantage une évaluation des processus que des effets de la formation. C'est une démarche qui reste nécessaire dans un objectif d'amélioration de la qualité de la formation.

Cependant, pour en tirer plein bénéfices, les organisations devront certainement apprendre à mieux préciser à quelles fins elles font ces évaluations et, conséquemment, sur quelles dimensions précises ces évaluations doivent porter. Elles devront aussi veiller à la qualité technique de leurs questionnaires. Et comme le souligne Tamkin et al. (2002), ce type d'évaluation, à lui seul, est certainement peu approprié pour juger de la valeur intrinsèque d'une formation.

## 6. L'ÉVALUATION DES APPRENTISSAGES

### 6.1 EN QUOI CONSISTE L'ÉVALUATION DES APPRENTISSAGES

Le terme d'apprentissage désigne le processus à l'origine de l'ensemble des modifications chez le participant attribuables à la formation.

En règle générale, la formation s'appuie sur une analyse préalable des besoins de formation qui servira de base à la formulation des objectifs de formation. Ceux-ci peuvent être regroupés en différentes catégories, ce qui a donné lieu, au cours des cinquante dernières années, à l'élaboration de taxonomies des objectifs. Celles de Bloom (1956) pour le domaine cognitif et de Krathwohl et al. (1964) pour le domaine affectif sont particulièrement connues. Ce qu'elles font entre autres ressortir, c'est que, comme la satisfaction, les produits de l'apprentissage sont de nature multidimensionnelle.

De manière assez sommaire, Kirkpatrick (1998) propose trois catégories de produits de l'apprentissage : les connaissances, les habiletés et les attitudes.

*«Les connaissances concernent ce que l'employé doit savoir au terme de la formation. Ceci fait référence aux faits, principes et règles relatifs au contenu de formation. Quant aux habiletés, on réfère au savoir-faire relié à un poste de travail. Les attitudes, elles, font référence au savoir-être de l'employé dans son environnement de travail.»* (traduction libre, Reitz, 1997, p. 6).

Dans un article souvent cité, Kraiger et al. (1993) critiqueront cette approche trop sommaire de l'apprentissage et proposeront une classification des produits de l'apprentissage, en s'appuyant entre autres sur les taxonomies de Bloom (1956) et Gagné (1984). Nous reviendrons plus loin sur cette question.

### 6.2 POURQUOI ÉVALUER LES APPRENTISSAGES?

Selon Kirkpatrick (1998), l'évaluation des apprentissages est importante parce que sans elle, il ne saurait y avoir de changement de comportement. Et si l'on devait constater qu'il n'y a pas de changement de comportement sans avoir préalablement mesuré les apprentissages, on serait alors tenté d'attribuer cette absence de changement au fait qu'il n'y a pas eu d'apprentissage, ce qui dans certains cas peut être faux.

Tracey, Hinkin, Tannenbaum, et Mathieu (2001) vont dans ce sens :

*«La satisfaction et l'apprentissage jouent un rôle important dans le processus de formation. Une réaction positive à une formation peut influencer la volonté d'un individu à utiliser la connaissance récemment acquise et à suivre d'autres programmes de formation. L'apprentissage est un élément fondamental du transfert et de l'application sur le lieu de travail.»* (traduction libre, p. 20).

La littérature (Alliger & Horowitz, 1989; Erickson, 1990) fait état de plusieurs raisons d'évaluer les apprentissages. En somme, il existe trois raisons principales :

- évaluer l'efficacité de la formation : la formation a été élaborée pour répondre à des besoins spécifiques traduits en objectifs d'apprentissage : il est donc normal que l'on souhaite savoir dans quelle mesure ces objectifs ont été atteints;
- savoir si le processus de formation s'avère adéquat pour l'atteinte des objectifs de formation : l'évaluation du degré d'atteinte de chacun des objectifs de formation contribue, avec l'évaluation de la satisfaction, à évaluer si la formation a été conçue, instrumentée et livrée de manière adéquate; en effet, les participants ne sauraient être considérés comme les seuls responsables de leur niveau d'atteinte des objectifs d'apprentissage;
- être en mesure de fournir aux apprenants une rétroaction sur leur niveau d'apprentissage : cette rétroaction revêt un caractère d'évaluation formative qui fait partie intégrante du processus de formation.

Notons cependant que l'ensemble de ces objectifs d'évaluation des apprentissages, si on souhaite bien les atteindre, commande des moyens qui ne sauraient se limiter au seul test de performance en fin de formation, qui constitue toutefois la pratique la plus courante d'évaluation des apprentissages. De plus, l'évaluation du processus de formation est plus directement abordée par les questionnaires de satisfaction que par l'évaluation des apprentissages bien que le fait de savoir si et jusqu'à quel point les participants ont appris représente en soi une donnée fort importante, malheureusement parfois négligée.

L'évaluation des apprentissages, également appelée «*évaluation pédagogique*», peut aussi être vue comme une occasion, pour un organisme de formation, de prendre un engagement sur le niveau de connaissances qui sera atteint par les personnes formées (Meignant, 2003). Ainsi, l'organisme peut engager sa responsabilité sur ce niveau d'acquis comme gage de sérieux ou à des fins de marketing.

### 6.3 PRATIQUES ET TENDANCES ACTUELLES D'ÉVALUATION DES APPRENTISSAGES

#### 6.3.1 Les pratiques actuelles

Selon Tamkin et al. (2002), les organisations font de plus en plus de l'évaluation des apprentissages une partie intégrante du programme de formation, à cause des difficultés d'accès et de la résistance des organisations à une évaluation post-formation. L'évaluation pendant la formation est vue à la fois comme une occasion de mesurer les apprentissages et une façon de faciliter la rétention des apprenants. Le fait que les apprenants sachent qu'ils vont être évalués a de fortes chances d'influer sur la qualité de leur participation à l'activité.

Il y a un consensus de la littérature sur le fait que l'évaluation des apprentissages ne peut se mettre en œuvre que dans la mesure où des objectifs ont été clairement définis au préalable (Phillips, 1991; Newby, 1992 in Tamkin et al., 2002; Meignant, 1997; Aubret & Gilbert, 2003; Le Boterf, 2002; Barzucchetti & Claude, 1995). En effet, c'est la nature des objectifs visés par la formation qui déterminera comment les apprentissages seront évalués.

Une fois l'objectif clairement défini et situé, il est possible de choisir une ou plusieurs modalités d'évaluation qui apparaîtront les plus appropriées. Barzucchetti et Claude (1995) proposent la synthèse reproduite au tableau 6. Notons que l'on peut s'interroger sur la pertinence d'utiliser des tests d'intelligence pour évaluer des apprentissages cognitifs. Certains outils ou techniques, tels que les tests de maîtrise procédurale ou la résolution de problèmes, pourront aussi être utilisés pour l'évaluation du domaine psychomoteur, selon la manière dont ils auront été conçus.

Kraiger et al. (1993) et Kraiger (2002) ont particulièrement exploré la question de la taxonomie des objectifs d'apprentissage. Ainsi, Kraiger (2002) propose une taxonomie très élaborée des différents produits de la formation selon trois grandes catégories : les produits cognitifs, les produits affectifs et les produits comportementaux. Pour chaque grande catégorie de produits, l'auteur identifie les produits et sous-produits et propose pour chacun d'eux une ou des méthodes de mesure. Par exemple, les produits affectifs se subdivisent en produits ayant trait à la motivation ou ayant trait à l'attitude. Les

produits ayant trait à l'attitude concernent soit l'efficacité personnelle, sur quoi porte l'attitude ou la force de l'attitude. Ces trois sous-produits se mesurent à l'aide de questionnaires ou d'autoévaluation.

Il est évident que le recours à cette taxonomie oblige le formateur à bien préciser la nature de ses objectifs de formation, ce qui constitue un apport en soi positif. Ceci étant fait, il trouvera des suggestions de modes d'évaluation appropriés. Par ailleurs, cette taxonomie est fort élaborée et pourra en rebuter quelques-uns. Elle propose des modes qui sont souvent déjà connus. Ainsi, les produits comportementaux sont essentiellement évalués par des observations systématiques à l'aide de grilles («*checklists*») ou des évaluations de performance. Enfin, elle ne résout pas pour autant tous les défis techniques assurant une certaine rigueur dans l'élaboration des outils, leur mise en application et le traitement et l'interprétation des données qui en découlent. En d'autres termes, on risque de retirer peu d'avantages à être très rigoureux au niveau de l'identification des objectifs et des méthodes de mesure, si cette rigueur ne s'étend pas à l'ensemble du processus d'évaluation.

Kirkpatrick (1998) fournit quelques conseils pratiques pour réaliser une évaluation des apprentissages :

- utilisez un groupe contrôle dans la mesure du possible (le «*groupe de contrôle*» s'oppose au «*groupe expérimental*» qui reçoit la formation);

**TABLEAU 6**  
**Techniques et outils d'évaluation des apprentissages selon le domaine**  
(Barzucchetti & Claude, 1995)

Domaine	Techniques et outils d'évaluation
Cognitif	Questionnaire écrit « réponses à fournir » ou « à choix multiples » Interrogation orale Test de raisonnement logique Test d'intelligence Test de maîtrise procédurale Problème à résoudre Étude de cas Dossier projet à créer Étude ou recherche à effectuer Rapport ou synthèse à rédiger Simulation de situation réelle
Psychomoteur	Test psychomoteur Travaux pratiques, manipulations Simulation de situation réelle
Affectif	Jeu de rôle Simulation de situation réelle



- évaluez les savoirs, les habiletés et/ou attitudes à la fois avant et après le programme;
- utilisez un test de type «*papier-stylo*» pour évaluer les connaissances et les attitudes;
- utilisez un test de performance pour mesurer les habiletés;
- obtenez 100% de réponses;
- utilisez les résultats de cette évaluation pour réaliser des actions appropriées.

Kraiger (2002) propose pour sa part les lignes directrices suivantes :

- *«lorsque vous choisissez les instruments de mesure des apprentissages, soyez très clairs sur les buts pour lesquels une évaluation est souhaitée;*
- *rédigez les objectifs d'apprentissage dans le langage le plus aisé possible pour les apprenants, les formateurs ou les décideurs de l'organisation, plutôt que de les forcer dans une description comportementale pour la seule fin de pouvoir les mesurer;*
- *rappelez-vous que la formation peut résulter en changement de comportements mais sans amélioration de la performance, ou que la formation peut entraîner un changement de performance sans modification de comportement: sachez exactement ce que vous attendez;*
- *utilisez une taxonomie telle que proposée pour identifier les objectifs dans différentes catégories de produits, puis choisissez des méthodes de mesure qui conviennent à ces produits;*
- *profitez de votre conscience accrue des produits d'apprentissage visés afin de guider vos décisions pendant tout le processus d'élaboration de la formation;*
- *concevez des outils de mesure qui permettent aux apprenants de démontrer leurs acquis cognitifs, affectifs et comportementaux en lien avec les objectifs de formation; parce que le but est de mesurer l'apprentissage et non l'apprenant, plus la mesure est directe, mieux c'est.»* (traduction libre, p. 358).

La littérature scientifique fait souvent valoir la nécessité d'un groupe contrôle et même de mesures avant et après la formation, si on souhaite démontrer de manière rigoureuse qu'il y a eu apprentissage (Arvey & Cole, 1989). Par ailleurs, comme on l'a vu en #4.3, il existe un vaste consensus sur le fait que de telles exigences sont souvent irréalistes dans le contexte de l'organisation. En somme, de telles exigences méthodologiques sont surtout pertinentes lorsque, par exemple, la démonstration des effets d'une formation est requise pour appuyer des décisions lourdes de conséquences. Elles ne doivent surtout pas saper à la base la volonté de mettre en place des pratiques soutenues d'évaluation des apprentissa-

ges qui ne seront peut-être pas complètement rigoureuses mais qui seront, si elles respectent certaines règles, bien mieux que rien du tout.

### 6.3.2 Une tendance nouvelle en évaluation des apprentissages

La plupart des pratiques actuelles d'évaluation des apprentissages visent essentiellement le contrôle d'une conformité avec des objectifs définis par l'organisation. Ces pratiques sont appelées à évoluer dans une perspective où la place de l'apprenant est plus importante et où l'on souhaite développer des travailleurs autonomes et faisant preuve d'un certain «*empowerment*».

Précédemment, on a déjà fait allusion à la critique que font Geertshuis et al. (2002) du modèle de Kirkpatrick (1959). Cette critique amène les auteurs à expérimenter des nouvelles modalités d'évaluation des apprentissages qui s'appuient sur les principes suivants :

- le stagiaire est considéré comme un client final et non pas comme la ressource interchangeable d'un employeur;
- le stagiaire est soutenu pour devenir autonome dans la gestion de son parcours, capable de développer des plans, devis et stratégies, et en mesure de procéder à l'évaluation de sa propre performance de même que de juger de la qualité et des produits d'une expérience d'apprentissage.

Ainsi, dans la pratique, il est demandé aux stagiaires d'évaluer à la fois la formation qu'ils ont suivie et leur propre performance. Pour ce dernier point, on leur demande, préalablement à la formation :

- d'auto-évaluer l'impact qu'ils attendent de la formation sur leurs performances de travail;
- de penser à des façons d'augmenter leur motivation;
- de réfléchir aux manières d'apprendre qui leur conviendraient le mieux;
- de réfléchir aux éléments qui leur semblent être des freins à leur apprentissage.

Cette démarche implique donc une participation beaucoup plus active de l'apprenant à son processus d'apprentissage et à son évaluation. C'est donc là un premier pas vers une formation et une évaluation qui non seulement permet l'atteinte d'objectifs, mais vise aussi le développement de l'autonomie de l'apprenant.

Cette vision d'un stagiaire participant actif du processus d'apprentissage et de son évaluation, qui ne peut plus se limiter à recevoir passivement et exécuter ce que les autres décident pour lui, mais qui gère plutôt de manière proactive sa performance rejoint entre autres la posi-

tion de Martocchio et Hertenstein (2003). Ces auteurs reconnaissent toutefois que tous les employés ne sont pas prédisposés à s'engager de cette façon dans des activités de formation et que le contexte de la formation joue un rôle important à cet égard. Elle rejoint aussi celle de Holton (1996) qui prétend que la promptitude d'une personne à recevoir une formation sera entre autres influencée par son degré d'implication dans l'évaluation des besoins et dans la planification de la formation, son degré de clarification des attentes et son niveau de choix.

### 6.3.3 La prise en compte des facteurs individuels et organisationnels

Plusieurs recherches (Geertshuis et al., 2002; Tracey et al., 2001; Holton, 1996; Martocchio & Hertenstein, 2003) ont permis de montrer que, contrairement à de nombreuses pratiques, le formateur et le contenu de la formation ne sont pas les seuls éléments à prendre en compte lors d'une évaluation des apprentissages. Dans un cadre de recherche qui se veut rigoureux, les facteurs individuels et organisationnels doivent aussi être considérés si l'on souhaite mieux prédire et expliquer à partir des résultats d'une évaluation des apprentissages.

Mais la prise en compte de ces facteurs par l'entreprise dans un processus d'évaluation courant apparaît peu réaliste. En effet, il paraît peu pensable qu'une entreprise puisse concrètement et systématiquement évaluer à quel point, par exemple, une personne est prête à recevoir la formation, quelle est son attitude au travail, les caractéristiques de sa personnalité ou encore sa motivation à transférer les apprentissages sur le lieu de travail, bien que la recherche ait prouvé que ces facteurs ont une influence sur l'apprentissage via la motivation (Holton, 1996; Tracey et al., 2001; Hicks & Klimoski, 1987; Alliger & Horowitz, 1989; Tannenbaum et al., 1991).

Contrairement aux facteurs individuels qui nécessitent un recueil et une analyse complexes, il semble cependant que certains facteurs organisationnels peuvent, même recensés partiellement, permettre de mieux comprendre les effets de la formation sur le niveau d'apprentissage. Ainsi, selon Tracey et al. (2001), le temps dont disposent les stagiaires pour se préparer pour la formation et pour utiliser les connaissances et les habiletés nouvellement acquises lorsqu'ils reviennent en poste aurait un impact important sur le niveau d'apprentissage. Les auteurs concluent leur article en insistant sur le fait que « *les variables liées à la tâche, au groupe et à l'organisation ont probablement un impact profond sur les types de contextes d'apprentissage. À ce titre, ils devraient être complètement pris en compte et compris avant toute mise en œuvre de programmes de formation.* » (traduction libre, Tracey et al., 2001. p. 20).

## 6.4 LIMITES DES PRATIQUES D'ÉVALUATION DES APPRENTISSAGES

### 6.4.1 Des exigences de rigueur rarement comblées

Comme on l'a vu précédemment, peu de pratiques d'évaluation des apprentissages risquent de rencontrer toutes les exigences de rigueur scientifique énoncées dans la littérature. Parmi celles-ci, rappelons entre autres la définition préalable d'objectifs mesurables et observables et l'utilisation de pré-test et post-test et de groupe contrôle. À cela devrait s'ajouter l'élaboration rigoureuse des outils et un traitement statistique des données lorsque cela s'y prête. Ces exigences sont certainement incontournables dans un contexte de recherche ou de démonstration de l'effet d'une formation. Par ailleurs, elles ne peuvent que difficilement être toutes transposées dans le quotidien de la formation en entreprise.

Le fait de ne pas appréhender correctement la nature multiple des apprentissages est une autre des critiques. Des auteurs ont exprimé la nécessité de distinguer les diverses dimensions des apprentissages pour ne pas les limiter aux domaines de la connaissance déclarative et du comportement. Parmi eux, Kraiger, et al (1993) critiquent le fait que les apprentissages sont souvent mesurés en donnant aux stagiaires des tests qui ne couvrent que les connaissances déclaratives. Une définition incomplète des apprentissages à évaluer ne peut qu'aboutir à une évaluation partielle des apprentissages.

### 6.4.2 Des exigences mal adaptées à certains apprentissages

L'évaluation des apprentissages avant et après la formation (ou l'utilisation de pré-test et post-test) ne s'applique pas à tous les apprentissages. En effet, cette pratique est davantage appropriée pour mesurer les connaissances déclaratives (par exemple, la connaissance de faits) que l'acquisition de « *soft skills* » (Bee & Bee, 1994 in Tamkin et al., 2002).

Par ailleurs, certaines compétences humaines et valeurs (par exemple l'intégrité) ne s'apprennent pas de façon linéaire mais d'une façon « *transformationnelle* » (Mezirow, 1997). C'est-à-dire qu'il est nécessaire que les stagiaires découvrent les choses par eux-mêmes avant que les apprentissages ne prennent vraiment place et puissent faire l'objet d'une évaluation. Dans de tels cas, les tests peuvent non seulement s'avérer peu performants, mais ils peuvent aller jusqu'à entraver les changements transformationnels en cours. (Rough, 1994 in Tamkin et al., 2002).

### 6.4.3 Un pouvoir prédictif limité

L'évaluation des apprentissages ne permet pas nécessairement de prédire le comportement qui s'en suivra. Un stagiaire peut savoir répéter ce qu'il a appris (une théorie ou des principes, par exemple) sans être capable de le mettre en pratique. Dans ces circonstances, il est nécessaire d'évaluer les apprentissages en les reportant dans le contexte où ils devront être mis en application (Erikson, 1990).

Alliger et al. (1997) ont observé des corrélations relativement peu élevées entre le transfert et les connaissances immédiatement en fin de formation ( $r = ,11$ ), les connaissances après un certain temps ( $r = ,08$ ) et l'évaluation des comportements en fin de formation ( $r = ,18$ ). En d'autres termes, l'évaluation des apprentissages, qu'il s'agisse des connaissances ou des comportements, permet peu de prédire les comportements en situation de travail.

### 6.4.4 Les risques de l'évaluation des apprentissages

L'évaluation des apprentissages soulève parfois quelques résistances de la part des apprenants et des organisations et qui sont certainement en partie justifiées. Les adultes n'aiment pas être évalués et les formes d'évaluation courantes ont souvent un côté scolaire qui peut les amener à sous performer. De plus, la question du choix des destinataires et de l'usage allant être fait des résultats peut être à l'origine d'une certaine frilosité (Meignant, 2003). On comprend bien évidemment que les résultats d'une évaluation personnalisée traitée en termes de « *M. X est capable de...* » ou « *n'est pas capable de ...* » et diffusée à l'encadrement peut susciter des réserves. Il semblerait toutefois que ce type de réserve disparaisse dès lors que tous les partenaires s'accordent sur l'importance de l'enjeu. « *On observe d'ailleurs que lorsque cette évaluation présente des enjeux forts- reconnaissance d'acquis donnant droit à une nouvelle qualification, habilitation à utiliser un matériel particulier, essentiellement pour des raisons de sécurité (habilitation électrique par exemple), ou, bien sûr, formations diplômantes-, la difficulté à évaluer disparaît comme par enchantement.* » (Meignant, 2003, p. 349).

Cette évaluation peut aussi être détournée de ses buts véritables. Ghodsian (1997 in Tamkin et al., 2002) fait remarquer que les formateurs étant souvent évalués par la performance des stagiaires à la fin de la formation, ils peuvent être incités à les faire travailler sur des tâches précises propices au développement de la performance à court terme, plutôt que sur des tâches variées, ce qui facilite le transfert et la performance à long terme.

### 6.4.5 Un découpage qui ne s'applique pas à toutes les formations

L'évolution récente de la formation de la salle de classe vers des formes de plus en plus près des situations

réelles en poste de travail, et l'utilisation d'approches par compétence questionne quelque peu le découpage entre les apprentissages et les comportements propre au modèle de Kirkpatrick. De manière concrète, les entreprises s'assurent de façon croissante que les apprentissages (et en particulier ceux qui ont trait à la tâche, par opposition à d'autres qui peuvent être de nature plus générale) soient faits dans le contexte de leur pratique avec l'aide de pairs.

La limite entre la période de formation et la période de transfert devient donc de plus en plus imprécise. S'il est possible après ce type de formation de procéder à une évaluation des connaissances qui rejoint ce que l'on entend par évaluation des apprentissages dans la littérature scientifique, l'évaluation des apprentissages comportementaux ressemble de plus en plus à celle des comportements en situation de travail que nous allons aborder dans la section suivante.

\* \* \*

Les limites que nous venons d'énoncer ne devraient pas amener les organisations à restreindre l'évaluation des apprentissages. Celle-ci, même quand elle n'est pas une pratique soutenue, garde toute sa pertinence, à tout le moins pour que les concepteurs et décideurs sachent dans quelle mesure les apprenants ont appris de la formation et, de manière plus formative, pour que ces derniers aient une rétroaction claire sur ce qu'ils maîtrisent ou non. D'ailleurs, l'évaluation des apprentissages en fin de formation fait parfois partie intégrante des exigences d'accréditation au poste de travail et est dans certains cas évoquée dans la documentation ISO.

Mais il n'est pas toujours réaliste pour l'entreprise, surtout la petite et moyenne qui ne dispose pas des ressources techniques requises, de recourir à des classifications conceptuelles complexes des produits des apprentissages et à des méthodologies élaborées. On retiendra de ces travaux d'abord la pertinence de bien identifier, dès le départ, et de la manière la plus explicite possible, tous les objectifs d'apprentissage visés par la formation. Cette étape cruciale déterminera le contenu et le design de la formation et devrait aussi permettre déjà d'entrevoir dans quelle mesure et comment se fera le transfert des apprentissages dans la performance.

On retiendra aussi qu'il faut préciser les raisons qui amènent à évaluer les apprentissages. Ce n'est que dans la mesure où ces raisons justifieront les efforts nécessaires à cette opération qu'il sera pertinent d'aller de l'avant. Ces raisons influenceront comment cette évaluation se fera, selon que les fins soient, par exemple, purement pédagogiques, consistent à démontrer l'efficacité d'une activité, ou visent une certification ou accréditation. En ce sens, elle pourra se situer à l'intérieur d'un vaste continuum allant d'un simple test élaboré par le formateur et corrigé par les apprenants eux-mêmes, à une évaluation externe en règle avec pré-test / post-test, groupe contrôle et analyse statistique des données.

## 7. L'ÉVALUATION DES COMPORTEMENTS

### 7.1 EN QUOI CONSISTE L'ÉVALUATION DES COMPORTEMENTS

«*Évaluation des comportements*» et «*évaluation du transfert d'apprentissage*» sont les expressions se référant au troisième des quatre niveaux de Kirkpatrick. Le terme «*transfert des apprentissages*» désigne la mise en application effective, pendant une certaine période de temps, des apprentissages en situation de travail. Pour Baldwin et Ford (1988), deux conditions doivent être remplies pour qu'il y ait transfert, soit le maintien des comportements appris pendant un certain temps et leur généralisation à des conditions de pratique ou contextes différents de ceux de la formation. L'expression «*transfert d'apprentissage positif*» est également utilisée (Newstrom, 1984; Wexley & Latham, 1981 in Baldwin & Ford, 1988). Cette évaluation se fait habituellement par l'observation au poste de travail et le recours à des données chiffrées liées à la productivité. Cette évaluation peut aisément correspondre à une évaluation des compétences puisqu'on évalue alors un savoir agir en situation de travail.

Si les apprentissages sont évalués pendant ou à la fin de la formation, les comportements le sont un certain temps après. À titre d'exemple, Kirkpatrick (1998) situe ce délai entre deux et six mois selon la formation, le temps que l'apprenant «*retourne à son travail, prenne en considération le nouveau comportement proposé et le mette à l'essai.*» (traduction libre, p.50).

### 7.2 POURQUOI ÉVALUER LES COMPORTEMENTS?

On peut considérer que les apprentissages en eux-mêmes n'ont que peu d'intérêt pour les entreprises, et que ce sont davantage les résultats organisationnels qui sont visés. Néanmoins, les activités de formation sont généralement initiées sur la base d'une identification de besoins directement liés à la performance. L'évaluation du transfert des apprentissages s'impose donc logiquement pour savoir si la formation a atteint ses objectifs et pourra contribuer à l'amélioration de la performance, puis des résultats organisationnels (Kozlowski et al., 2000 in Salas & Cannon-Bowers, 2001).

Les approches traditionnelles du transfert considèrent en général le lien entre apprentissage, performance et résultats organisationnels de façon linéaire (Yamnill & McLean, 2001). Le diagramme 2 donne une illustration de l'enchaînement de ces étapes.

Par ailleurs, il est important de distinguer transfert des apprentissages et performance. Ainsi, Gilbert (1996 in

Brethower, 2001) identifie des variables susceptibles d'interférer sur la performance, entre autres : l'efficacité de la production, la disponibilité des outils et du matériel pour supporter les processus de travail, la connaissance partagée et l'adhésion aux buts, standards et moyens de mesure du travail, la disponibilité en nombre suffisant de personnel compétent pour atteindre les buts visés et leur connaissance précise des attentes à leur égard, leurs compétences à atteindre les buts, le feed-back qu'ils reçoivent sur leur performance, etc. On voit donc que plusieurs de ces variables n'ont aucun lien avec la formation et le transfert subséquent des apprentissages dans la situation de travail.

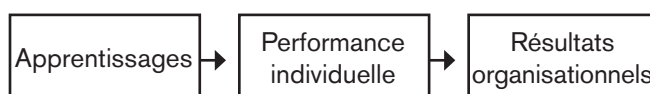
La raison principale d'évaluer les comportements consiste donc à savoir précisément dans quelle mesure les apprentissages acquis pendant la formation ont été transférés au poste de travail et pourront contribuer, en autant que les autres variables évoquées précédemment seront prises en compte, à améliorer la performance. Or, le seul transfert des apprentissages dans la situation de travail dépend aussi d'un ensemble de variables déjà identifiées dans la littérature. En d'autres termes, si on se réfère au diagramme précédent, la flèche entre apprentissages et performance individuelle est soumise à deux ensembles de facteurs : ceux qui sont liés au simple transfert des apprentissages en situation de travail et ceux qui permettront aux apprentissages transférés de se traduire en performance individuelle.

Si les organisations souhaitent que ce transfert s'opère positivement, il importe qu'elles connaissent les facteurs qui influencent le transfert et la performance et, dans la mesure du possible, en contrôlent l'influence si elles souhaitent optimiser l'effet de la formation sur les résultats organisationnels.

### 7.3 LES FACTEURS INFLUENÇANT LE TRANSFERT

Le fait que les apprentissages réalisés dans les formations soient difficilement transférés en situation de travail est un problème souvent évoqué dans la littérature. De ce fait, les chercheurs ont beaucoup travaillé la question des facteurs prédictifs de l'efficacité de l'apprentissage.

**DIAGRAMME 2**  
**Lien entre les apprentissages et les résultats organisationnels** (Holton, 1996)





Ils considèrent que le transfert des apprentissages est soumis à de nombreux facteurs qui interagissent entre eux et que l'on peut qualifier de « *système d'influences* » (Holton, 1996).

Dans le modèle de Baldwin et Ford (1988), ce « *système d'influences* » peut être vu comme le résultat du jeu de trois types de facteurs (les « *entrées* ») :

- les caractéristiques du stagiaire,
- la conception de la formation,
- l'environnement de travail.

### 7.3.1 Les caractéristiques du stagiaire

La recherche est partagée sur l'effet des facteurs liés à la personnalité des stagiaires sur le transfert des apprentissages. Elle doit également faire face à un problème de tri et de hiérarchisation de ces facteurs qui sont extrêmement nombreux. En effet, des éléments à la fois cognitifs, affectifs et liés à l'aptitude physique peuvent entrer en jeu. Des chercheurs ont ainsi développé des outils comprenant plus de cinquante descripteurs pour conduire des analyses sur les aptitudes influençant la performance au travail (Fleishman & Mumford, 1989 in Holton, 1996).

L'efficacité des tests d'aptitude comme éléments prédictifs du transfert des apprentissages en situation de travail par exemple, ou les liens qui pourraient exister entre la personnalité du stagiaire et son aptitude à transférer ce qu'il a appris sont remis en question (Gordon & Kleiman, 1976; Tubiana & Ben-Shakhar, 1982; Ghiselli, 1966 in Baldwin & Ford, 1988).

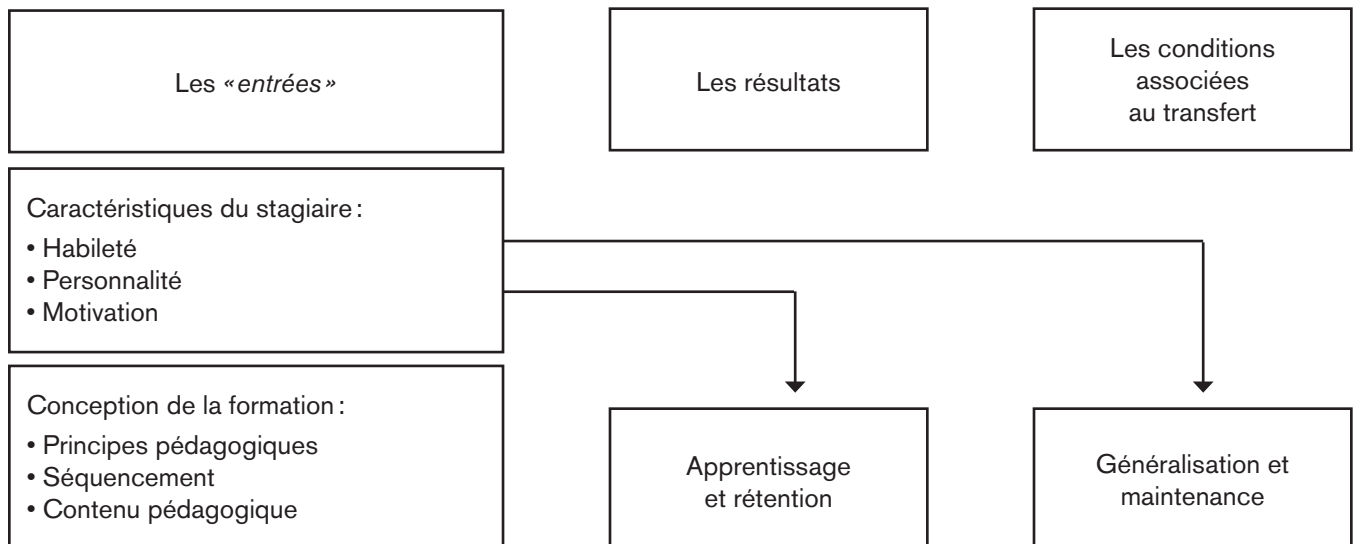
Parmi les facteurs reconnus comme influents sur le transfert, citons toutefois le niveau de motivation et le sentiment d'efficacité personnelle des stagiaires.

La motivation peut être définie comme la volonté du stagiaire d'utiliser à son travail, les connaissances et les habiletés acquises lors de la formation (Noe & Schmitt, 1986 in Yamnill & McLean, 2001; Salas & Cannon-Bowers, 2001). Holton (1996) a proposé de distinguer la motivation selon qu'elle désigne la promptitude à apprendre, le comportement au travail, des caractéristiques de la personnalité ou la motivation à transférer les apprentissages.

Il a été prouvé que des stagiaires profondément impliqués dans leur emploi sont plus motivés pour apprendre et pour transférer leurs savoirs sur leur lieu de travail (Ryman & Biersner, 1975; Noe & Schmitt, 1986 in Baldwin & Ford, 1988). De même, les gestionnaires qui croient en la valeur de la formation sont plus motivés pour appliquer leurs nouveaux acquis au travail (Baumgartel, 1984 in Baldwin & Ford, 1988). Il a été établi également que l'atteinte des objectifs de la formation par les stagiaires jouait un rôle important vis-à-vis de leur motivation à apprendre et à transférer ce qu'ils ont appris (Tannenbaum et al., 1991).

Le sentiment d'efficacité personnelle des stagiaires « *renvoie, dans le contexte de la formation, à la croyance que possède le stagiaire qu'il ou elle peut maîtriser une tâche et un comportement et les appliquer au travail* » (Gaudine & Saks, 2004). Il a été prouvé à maintes reprises que la formation augmentait le sentiment d'efficacité personnelle des stagiaires et que ce dernier était un élément prédictif fiable de leur capacité d'apprentissage

**DIAGRAMME 3**  
**Les facteurs qui influencent le transfert** (Baldwin & Ford, 1988)



et de performance (Salas & Cannon-Bowers, 2001; Tannenbaum et al., 1991; Mathieu et al., 1993; Haccoun et al., 1997; Warr et al., 1999).

### 7.3.2 La conception de la formation

Il a été prouvé que les stratégies d'intervention pouvaient être conçues de façon à améliorer le niveau de transfert des apprentissages (Brinkerhoff & Montesino, 1995; Kraiger et al., 1995 in Salas & Cannon-Bowers, 2001; Brinkerhoff & Apking, 2001) et qu'une des causes d'échec du transfert pouvait résider dans la façon dont la formation est conçue (Holton, 1996).

Quatre principes d'apprentissage de base généraux ont été étudiés et appliqués par la recherche sur l'amélioration du transfert des apprentissages au niveau de la conception de la formation (Baldwin & Ford, 1988). Ces principes peuvent éclairer le travail de conception de formations en entreprise :

- le premier est celui des «*éléments identiques*» et affirme que le transfert est amélioré dans la mesure où on retrouve des éléments identiques entre la situation de formation et le poste de travail (Thorndike & Woodworth, 1991 in Yamnill & McLean, 2001).
- le second est celui des «*principes généraux*». Il défend l'idée selon laquelle le transfert est facilité lorsqu'on apprend aux stagiaires non pas uniquement des habiletés pratiques, mais aussi des règles générales et des principes théoriques (Laker, 1990 in Yamnill & McLean, 2001).
- le troisième est celui de la «*variabilité des stimuli*». Il préconise que le transfert est amélioré lorsqu'une variété de stimuli (utilisation de tâches différentes) est utilisée lors de l'apprentissage (Ghodsian et al., 1997 in Tamkin et al., 2002)
- le quatrième principe est celui des «*conditions de pratique*». Il suggère que les matériaux appris par une pratique étalée dans le temps sont retenus plus longtemps que ceux acquis par une pratique intensive.

Certaines pratiques peuvent aussi contribuer à une augmentation du niveau de transfert. Parmi celles-ci, notons :

- la fixation d'objectifs : la fixation d'objectifs pendant et après la formation améliore le transfert (Locke & Latham, 1990 in Richman-Hirsch, 2001; Werner et al., 1994; Wexley & Baldwin, 1986 in Holton, 1996);
- l'autogestion : selon Richman-Hirsch (2001), cette pratique consiste à former l'apprenant à évaluer les obstacles potentiels à sa performance, exercer un suivi de l'influence de l'environnement à cet égard, planifier des réactions en conséquence et donner des récompenses lorsque les obstacles sont contournés.

- l'organisation de sessions post-formation : la façon dont l'activité se termine est également un élément jugé déterminant vis-à-vis du transfert (Archambault, 1997), car elle permet entre autres aux participants d'avoir «*des échanges concrets sur l'utilité de la formation*». Ces sessions post-formation portent différents noms suivant leur principe d'organisation : «*formation de suivi*» (Archambault, 1997) ou «*session de renforcement*» (Ashby & Wilson, 1977; Kingsley & Wilson, 1977 in Baldwin & Ford, 1988), «*systèmes de binômes*» (Barrett, 1978; Karol & Richards, 1978 in Baldwin & Ford, 1988); etc.

### 7.3.3 L'environnement de travail

De toute évidence, l'environnement de travail a également une influence sur le transfert des apprentissages (Richman-Hirsch, 2001; Quinones, 1997 in Salas & Cannon-Bowers, 2001; Tracey, 1992 in Holton, 1996). Les recherches récentes ont particulièrement porté sur le climat de transfert de l'organisation (Mathieu et al., 1992 in Holton et al., 2000). Il s'agit ici d'une variable intermédiaire entre le contexte organisationnel, et l'attitude et le comportement au travail de l'individu. Selon Holton et al. (2000), si certains auteurs y voient une influence déterminante, et même dans certains cas aussi importante que la formation elle-même, il n'existe pas actuellement de compréhension claire de ce en quoi cela consiste exactement.

La façon dont les personnes sont gérées est un élément déterminant du type d'environnement de travail. Par exemple, une gestion incluant des échanges préalables à la formation et le soutien du responsable hiérarchique, contribue de façon importante à l'augmentation du transfert des apprentissages (Reber & Wallin, 1984 in Baldwin & Ford, 1988; Smith-Jentsch et al., 2000 in Salas & Cannon-Bowers, 2001). Le niveau d'autonomie au travail (absence de contrainte des pairs ou du manager) est aussi un élément important à cet égard.

Kirkpatrick (1998) distingue cinq attitudes du responsable hiérarchique ayant une influence sur le climat et donc sur le transfert : l'attitude préventive (le responsable interdit au stagiaire d'appliquer ce qu'il a appris); l'attitude décourageante; l'attitude neutre; l'attitude encourageante; l'attitude exigeante. Cette dernière désigne le fait que le responsable sait ce qui a été appris. Il a éventuellement lui-même suivi le programme, et il s'assure que les apprentissages vont être transférés en situation de travail. On pourra éventuellement parler de valorisation des actions du stagiaire issues de la formation lorsque l'environnement immédiat du stagiaire cherche à savoir ce qui peut être mis en application. (Facteau et al., 1995; Tracey et al., 2001 in Salas & Cannon-Bowers, 2001; Haccoun et al., 1997).

D'autres éléments entrent en ligne de compte comme déterminants potentiels du type d'environnement de travail. Parmi eux, citons-en deux qui rencontrent une certaine adhésion auprès des chercheurs :

- l'occasion de mise en pratique; les stagiaires ont besoin d'avoir l'occasion de mettre en œuvre ce qu'ils ont récemment appris pour pouvoir le transférer (Ford et al, 1992; Quinones et al., 1995 in Salas & Cannon-Bowers, 2001).
- le délai: le temps s'écoulant entre la fin de la formation et l'application des acquis en situation de travail a également une influence (Arthur et al., 1998 in Salas & Cannon-Bowers, 2001).

## 7.4 PRATIQUES ET TENDANCES ACTUELLES D'ÉVALUATION DES COMPORTEMENTS

### 7.4.1 Les pratiques courantes et les conseils généraux

Selon Twitchell et al. (2000), les pratiques les plus courantes d'évaluation des comportements des entreprises américaines sont l'observation (56 %), l'évaluation de la performance (49 %), les enquêtes des superviseurs (36 %), et l'anecdote (32 %).

Barzucchetti et Claude (1995) recommandent quatre façons d'évaluer les compétences en situation de travail :

- l'observation des comportements professionnels soit à l'aide d'un guide d'observation ou d'un questionnaire d'évaluation des comportements lorsque l'observation s'avère impossible; il s'agit ici de recueillir le témoignage de l'apprenant et d'autres observateurs sur ce qu'il met en application et comment, ainsi que sur les raisons des comportements;
- l'analyse des produits de l'activité du participant : dans certains cas, comme des habiletés de rédaction, il est plus simple d'analyser les produits du participant que de l'observer; il faut par ailleurs prendre en compte les variables qui peuvent influencer sur la qualité ou la quantité de sa production et qui ne dépendent pas de la formation; une grille d'analyse du produit peut être élaborée pour les fins de cette évaluation;
- l'interprétation des performances individuelles du participant : les compétences acquises devraient se traduire en performance, à certaines conditions qui doivent être prises en compte; les auteurs insistent sur les conditions liées à la personne (motivation, caractéristiques personnelles freinant la performance, persévérance) et liées à l'environnement relationnel ou logistique;
- l'analyse critique des situations problèmes : il s'agit ici d'analyser après coup les comportements émis en

réponse à un incident critique en lien avec les compétences acquises; cette évaluation tente d'isoler les comportements qui n'ont pas été adéquats, les causes de la carence et les acquis qui ont été mal utilisés.

Phillips (1997) propose entre autres les questionnaires de suivi ou d'enquête, les observations en poste de travail, les entrevues avec les participants, les «*focus groups*» de suivi. À cela s'ajoutent :

- les plans d'action que les participants doivent mettre en œuvre suite à la formation et pour lesquels ils doivent rendre compte;
- les contrats de performance qui sont en quelque sorte des ententes mutuelles où le participant et le superviseur conviennent d'une amélioration de la performance dans un domaine convenu en lien avec la formation reçue;
- les sessions de suivi de la formation, où les participants se réunissent à nouveau après un certain temps pour procéder à l'évaluation de la formation reçue et recevoir une formation complémentaire en fonction des obstacles identifiés.

Kirkpatrick (1998) insiste sur le fait qu'on ne peut pas faire l'évaluation des comportements sans avoir évalué les niveaux précédents au risque d'aboutir à un diagnostic erroné. Il décrit trois difficultés pour réaliser cette évaluation :

- les stagiaires ne peuvent changer leur comportement sans avoir l'occasion de mettre en œuvre leurs nouveaux apprentissages;
- il est impossible de prédire le moment où un changement de comportement va avoir lieu;
- les stagiaires peuvent transférer leurs apprentissages sur le lieu de travail puis en venir à la conclusion que finalement, ils préfèrent leur ancienne façon de faire.

Il donne les conseils suivant pour réaliser cette évaluation :

- «*Utiliser un groupe de contrôle dans la mesure du possible.*
- *Laisser suffisamment de temps aux stagiaires pour que des changements de comportement se produisent.*
- *Évaluer à la fois avant et après le programme, dans la mesure du possible.*
- *Interviewer ou sonder une ou plusieurs des personnes suivantes : les stagiaires, leur superviseur direct, leurs subordonnés, et d'autres observateurs de leur comportement.*
- *Obtenir 100 % de réponses ou constituer un échantillon.*

- Répéter l'évaluation à des moments appropriés.
- Considérer les coûts versus les bénéfices. » (traduction libre, Kirkpatrick, 1998, p.49)

#### 7.4.2 Déterminer le meilleur moment

À propos des délais de l'évaluation, la question du meilleur moment pour mesurer des changements de comportements va dépendre de la nature de la formation. Il y a des résultats contradictoires à ce sujet. Kirkpatrick (1998) semble plaider en faveur d'une évaluation deux à trois mois ou plus après la formation, mais reconnaît toutefois que certains participants ne vont pas changer leur comportement avant six mois. Axtell et al. (1997 in Tamkin et al., 2002) ont démontré que la quantité d'apprentissages transférés après un mois est un indicateur fiable de la quantité d'apprentissages transférés après un an.

Il a également été prouvé que la période suivant immédiatement la formation était critique vis-à-vis du transfert des apprentissages en situation de travail. En effet, durant cette période, les erreurs peuvent facilement être produites et le support du manager devient essentiel (Axtell et al., 1997 in Tamkin et al., 2002). Elle est donc plus propice à une évaluation formative des comportements.

Il est nécessaire d'insister sur le fait que seules des évaluations répétées des comportements permettront de s'assurer de leur maintien, l'apprenant pouvant après un certain temps décider de reprendre ses comportements antérieurs à la formation.

#### 7.4.3 Le choix des acteurs

Une bonne évaluation des comportements en poste de travail exige la coopération des gestionnaires et des stagiaires eux-mêmes. Des chercheurs ont montré qu'on pouvait toutefois questionner l'aptitude d'un individu à s'autoévaluer de façon impartiale (Bee & Bee, 1994 in Tamkin et al., 2002; Carless & Roberts-Thompson, 2001 in Tamkin et al., 2002).

#### 7.4.4 La nécessité de faire un diagnostic préalable

De même que pour les apprentissages, le transfert sur le lieu de travail nécessite la spécification des connaissances, des habiletés, des tâches et des comportements devant être mis en œuvre au poste de travail. Il faut ensuite déterminer au bout de combien de temps et à quelle fréquence le stagiaire doit les appliquer ainsi que l'importance de ces nouvelles connaissances. Une analyse de tâche est donc incontournable (Baldwin & Ford, 1988).

#### 7.4.5 Au-delà de l'évaluation des seuls comportements

La plupart des entreprises qui évaluent les comportements s'y limitent. Kraiger (2002) propose un ensemble d'éléments qui devraient faire partie d'une évaluation complète du transfert des apprentissages et de ses conditions. Il s'agit :

- du climat de transfert, qui est une « mesure des perceptions de l'apprenant des situations et actions qui témoignent d'un support de l'organisation au transfert de la formation » (p. 360); ceci peut être fait, par exemple, par des instruments tels que le « *Transfer climate survey* » de Tracey et al., (1995 in Kraiger, 2002);
- de l'occasion de performer, qui est une « mesure des perceptions de l'apprenant (ou de l'observateur) à savoir s'il a l'occasion ou recherche de lui-même des occasions de mettre en application les nouveaux comportements, connaissances ou habiletés » (traduction libre, p. 360); ceci peut être fait, par exemple, par la « *Opportunity to perform survey* » (Ford et al., 1992 in Kraiger, 2002);
- du changement de comportement au travail, qui « précise quelles activités ont changé ou sont émises plus souvent suite à la formation » (traduction libre, p. 360); ceci peut être fait à l'aide de grilles d'observation du comportement;
- du maintien des habiletés, qui « détermine si l'émission et la fréquence des nouveaux comportements sont maintenus avec le temps » (traduction libre, p. 360); ceci peut être fait par une observation en deux temps de la performance au travail;
- de la généralisation des habiletés, qui « mesure si l'apprenant peut mettre en application ses nouveaux apprentissages dans des situations différentes de celles de la formation » (traduction libre, p. 360); ceci peut être entre autres évalué par des « *Probed protocol analysis* » (Kraiger et al., 1993).

Dans ce dernier cas, étant donné que les comportements impliquent aussi des connaissances et que la généralisation des apprentissages à d'autres tâches peut être assez complexe, la « *probed protocol analysis* » consiste à appliquer des protocoles où un expert en la matière demande à l'apprenant d'expliquer à haute voix et de manière détaillée ce qu'il est en train de faire. Ceci rejoint le questionnaire d'évaluation des comportements de Barzuchetti et Claude (1995).

Holton et al. (2000) propose aussi un outil pour conduire une analyse des facteurs influençant le transfert : « *l'inventaire du système de transfert des apprentissages* » (« *Learning Training System Inventory* » ou LTSI). Le LTSI est un outil d'évaluation comprenant soixante-deux questions regroupées dans une structure comprenant seize



facteurs tels que la motivation à transférer, le support des collègues de travail et celui du manager, la perception de la validité des contenus, l'existence d'occasions permettant d'appliquer les nouveaux acquis, etc... Il peut être utilisé de plusieurs façons, notamment : pour évaluer les problèmes de transfert potentiels en amont d'une formation; en tant qu'élément d'évaluation dans le cadre du suivi de programmes de formation existants; ou pour cibler les actions à entreprendre pour améliorer un transfert.

## 7.5 LIMITES DES PRATIQUES D'ÉVALUATION DES COMPORTEMENTS

Les principales limites des pratiques courantes d'évaluation des comportements recensées par la littérature sont les suivantes :

- les observateurs, qui sont souvent les supérieurs immédiats ne sont pas nécessairement entraînés; de plus, le manque de soutien du gestionnaire immédiat ou superviseur est souvent dénoncé comme un des problèmes les plus récurrents de l'évaluation de la formation. Rae (1999 in Tamkin et al., 2002) défend le point de vue selon lequel les départements de formation devraient éduquer les superviseurs sur ce sujet et sur le rôle qu'ils ont à jouer. Une des façons de gagner l'adhésion des superviseurs est, pour le responsable de formation, de travailler avec eux en devenant un consultant interne et de jouer ainsi un rôle plus actif vis-à-vis des questions politiques et des changements organisationnels à long terme (Newby, 1992 in Tamkin et al., 2002);
- les pratiques usuelles, qui se limitent aux observations des comportements, ne permettent pas l'évaluation de transferts d'apprentissages complexes («*soft skills*»); le recours à de nouvelles méthodes et techniques, telles que celles proposées par Barzucchetti et Claude (1995) et Kraiger et al. (1993) devrait être encouragé;
- la majorité des recueils de données sont réalisés dans le court terme, peu de temps après la fin des formations et n'ont pas été répétés après un laps de temps plus long; la rétention sur le long terme n'est donc pas connue; l'idéal serait de réaliser un recueil de données à la fois dans le court terme et dans le long terme (Gaudine & Saks, 2004);
- bien qu'il existe des outils à cet effet, il semble y avoir peu d'évaluation du climat de transfert qui permettrait de comprendre l'influence de certains facteurs, entre autres reliés à l'environnement, et d'approfondir la réflexion sur les meilleurs façons de concevoir une formation qui prenne en considération et agisse sur ces facteurs;

- lorsque l'évaluation des comportements est faite lors de l'évaluation annuelle de la performance ou du rendement, elle est alors très sommaire, et permet peu de bien évaluer l'apport de la formation à la performance;
- une des principales sources d'information à propos des changements comportementaux est le stagiaire lui-même; comme on l'a vu plus haut, on peut remettre en cause la capacité des personnes à évaluer leurs propres changements comportementaux;
- enfin, comme le souligne Denney (2001), les pratiques actuelles d'évaluation des comportements limitent l'effet de la formation à la seule mise en application de certains comportements prédéterminés et sont de ce fait très normatives et limitatives; dans une perspective de développement du capital humain et d'organisation apprenante, il y aurait certainement avantage à laisser place, par exemple, aux apprentissages et comportements non prévus qui contribuent au développement personnel et professionnel de l'apprenant.

Néanmoins, les professionnels de la formation doivent prendre conscience du fait que le transfert ne dépend pas que de la formation et qu'une évaluation bien faite du transfert et des conditions critiques qui l'accompagnent devrait amener une réflexion impliquant un cercle élargi d'acteurs de l'organisation ne se limitant pas aux seuls formateurs. En ce sens, l'évaluation des comportements doit se poursuivre et se développer de manière à ce que l'organisation soit mieux mobilisée pour faire en sorte que la formation génère les résultats attendus.

## 8. L'ÉVALUATION DES RÉSULTATS DE LA FORMATION SUR L'ORGANISATION

### 8.1 EN QUOI CONSISTE L'ÉVALUATION DES RÉSULTATS DE LA FORMATION

Les résultats de la formation sur l'organisation désignent les résultats finaux atteints grâce à la participation des stagiaires au programme de formation. On désigne par là le dernier stade de la chaîne des relations de cause à effet de la formation : les stagiaires ont transféré leurs nouveaux acquis au travail, ces nouveaux comportements ont augmenté leur niveau de performance, et l'efficacité de l'unité de travail ou de l'organisation en a été améliorée (Kraiger, 2002). On parle aussi de la contribution de la formation à la performance de l'entreprise, des « *résultats organisationnels* » ou des « *résultats d'affaires* » de la formation.

Les résultats de la formation correspondent au niveau quatre de Kirkpatrick. Ils peuvent s'observer à travers tous les indicateurs directs des performances collectives de l'entreprise tels que par exemple un accroissement de la production, une amélioration de la qualité, une réduction des coûts ou une augmentation des profits mais aussi à travers des indicateurs indirects tels qu'une amélioration de l'ambiance de travail, une augmentation de l'adaptabilité des employés, ou une évolution de la culture d'entreprise (Meignant, 2003). Tous ces effets peuvent aussi être rassemblés sous la désignation « *impact* » de la formation. Selon les modèles d'évaluation de la formation, ils peuvent aussi comprendre le retour sur l'investissement ou le rendement financier, mais ce type de résultat sera ici traité séparément à la section suivante.

Notons enfin que l'expression « *résultats de la formation* » peut être comprise de deux façons. Elle peut désigner les conséquences d'un programme de formation particulier ou bien désigner la contribution du service de formation à l'organisation. Ainsi, bien que l'effort d'évaluation soit généralement dirigé vers un programme de formation particulier, il faut avoir en tête qu'un objectif sous-jacent pourrait être l'évaluation de la contribution du service formation à l'entreprise et l'efficacité de la gestion de ses responsables (Smith, 2001).

### 8.2 POURQUOI ÉVALUER LES RÉSULTATS DE LA FORMATION

La littérature suggère qu'il y a un intérêt grandissant dans les organisations vis-à-vis de la justification de l'investissement dans la formation. Cet intérêt se traduit par une attente de résultats d'évaluation sous la forme d'indicateurs chiffrés et si possible, financiers. Productivité accrue, profits ou sécurité en hausse, réduction des erreurs

ou des plaintes des clients, augmentation des parts de marché, ou chiffres permettant le calcul du retour sur investissement sont parmi les indicateurs attendus de cette justification de l'investissement dans la formation, car ils sont aussi les indicateurs que les cadres opérationnels utilisent pour mesurer leurs propres résultats (Tamkin et al., 2002; Smith, 2001; Redshaw, 2000). Mais les résultats de la formation ne se traduisent pas tous en données chiffrées et les responsables de formation, pour produire les données qui leur permettraient d'assurer leur légitimité, se trouvent face à une difficulté technique et méthodologique importante.

Au-delà de la seule justification, l'évaluation des résultats de la formation revêt un intérêt certain. En effet, on ne forme pas pour former, mais pour contribuer au développement de l'entreprise et de ses employés. Une démarche d'évaluation des résultats est certainement un moment privilégié pour réfléchir sur ces finalités et leur formulation précise dans le contexte de l'organisation, voir comment et dans quelle mesure la formation y contribue et quelles sont les principales contraintes qui interfèrent et qui pourraient être levées. Un tel exercice d'évaluation de programme n'a certainement pas à être mené fréquemment, mais peut être d'un apport important à une réflexion stratégique sur la fonction formation de l'organisation.

### 8.3 PRATIQUES ET TENDANCES ACTUELLES D'ÉVALUATION DES RÉSULTATS DE LA FORMATION

#### 8.3.1 Les principales difficultés de l'évaluation des résultats de la formation.

Les résultats sur l'organisation sont reconnus comme étant les plus difficiles à évaluer pour plusieurs raisons :

- la formation est un facteur parmi de nombreux autres; plus on avance dans la chaîne des effets et plus les bénéfiques de l'intervention analysée deviennent difficiles à mesurer parce que l'effet de la seule formation devient très difficile à isoler (Tamkin et al., 2002; Redshaw, 2000; Meignant, 2003);
- d'autre part, les résultats de la formation au niveau organisationnel peuvent se produire après un délai important rendant également très difficile la reconnaissance des résultats de la formation (Redshaw, 2000); cette difficulté est renforcée par la pression des cadres opérationnels ayant tendance à suivre la performance de l'entreprise à court terme, sur une base mensuelle ou semestrielle (Watson, 1998), alors qu'il

est possible qu'à court terme, le transfert se manifeste peu ou pas, ou même par une baisse de performance;

→ il y a un travail d'explication important à mener auprès des gestionnaires de l'entreprise afin de leur expliquer pourquoi on ne peut pas évaluer les résultats de la formation en faisant une simple comparaison des entrées et des sorties (Redshaw, 2000; Pulley, 1994).

### 8.3.2 Comment dépasser les difficultés de l'évaluation des résultats de la formation

#### *La recherche de signes plutôt que de preuves.*

La fiabilité de l'évaluation repose-t-elle sur le fait que l'on puisse démontrer de façon évidente que la formation est la seule cause du résultat observé au niveau de l'organisation? Cette question trouve des réponses partagées dans la littérature. Elle met en évidence la difficulté de vouloir isoler les résultats de la formation sur une chaîne d'impacts et sur des objectifs à moyen ou long terme. Cette difficulté expliquerait la réticence des évaluateurs à mettre en œuvre une évaluation des résultats (Meignant, 2003; Haccoun et al., 1997).

Phillips (1997) propose 10 stratégies afin d'isoler l'effet de la formation. Une fois l'identification des facteurs ayant contribué à l'amélioration de la performance effectuée, ces stratégies sont : l'utilisation de groupes contrôle; l'analyse de tendance («*trend-line analysis*») où l'on projette de manière linéaire la performance sur la base de résultats passés, l'écart par rapport à cette projection pouvant être attribué à la formation; les méthodes de prédiction des changements de performance, qui permettent une estimation lorsque plus d'une variable (la formation) a pu avoir une influence sur la performance; l'estimation de l'impact de la formation par le participant, le superviseur ou le gestionnaire; l'avis des clients; l'estimation de l'impact de la formation par un expert ou les subordonnés du participant; et le calcul de l'impact des autres facteurs, lorsque cela s'avère possible. Si ces stratégies peuvent améliorer l'évaluation des résultats, il n'en demeure pas moins qu'elles présentent toutes des faiblesses et augmentent la lourdeur de l'opération.

En ce sens, lorsqu'il s'agit d'évaluation d'impact, une vision réaliste qui cherche patiemment à accumuler des preuves partielles peut parfois être préférable à une vision perfectionniste qui risque fort de devenir très lourde techniquement tout en masquant plus ou moins habilement une bonne part d'insuffisances.

#### *La compréhension de l'environnement et du fonctionnement de l'entreprise.*

Pour optimiser les résultats d'un programme de formation et pouvoir ensuite les évaluer de façon fiable et efficace, il est nécessaire d'avoir une compréhension réelle de l'environnement et du fonctionnement de l'organisation.

Pour ce faire, il est recommandé de garder en tête qu'il ne faut pas chercher de relations causales directes entre un programme de formation et une amélioration de résultats opérationnels car ce lien existe rarement (Spitzer, 1999; Kirkpatrick, 1998). La formation étant rarement la seule responsable d'un effet sur les affaires, il faut davantage procéder à l'identification des différentes étapes de la chaîne de causes à effets qui relie la formation à ses résultats sur l'organisation (Spitzer, 1999; Williams, 1996).

De nombreux chercheurs recommandent à cet effet que les preneurs de décision et les cadres opérationnels<sup>3</sup> de l'entreprise soient impliqués dans l'évaluation de la formation (Redshaw, 2000; Spitzer, 1999; Pulley, 1994; Torres & Preskill, 2001; Kirkpatrick, 1998). Il est notamment admis que le processus d'évaluation gagne à débiter à partir du moment où un besoin de formation a été identifié, ce qui nécessite déjà une rencontre formateurs / cadres opérationnels. Il est donc nécessaire que les preneurs de décision destinataires des résultats de l'évaluation soient clairement identifiés (Pulley, 1994).

Il est également recommandé que la coopération avec les cadres opérationnels soit suivie dans le temps. Pour ce faire, un mode de fonctionnement permettant la mise en place de bilans d'évaluation communs (et non de bilans de formation) devrait être organisé (Redshaw, 2000). Le signe que l'évaluation est bel et bien centrée sur les critères d'affaires qui importent à l'entreprise pourrait être que le directeur de la formation et le directeur commercial se rencontrent pour parler de l'amélioration des performances (Watson, 1998).

Il faut de plus collecter les données permettant de montrer que la formation a un impact positif sur la performance d'un individu, sur son équipe de travail ou encore sur l'organisation. La performance est généralement mesurée au niveau individuel par le biais d'estimations de performance traditionnelles (par exemple, on demande aux stagiaires et à leurs superviseurs de donner une estimation du pourcentage d'accroissement de la performance imputable à la formation), par des systèmes d'évaluation à 360 degrés, de sondage ou par le biais de programmes de certification plus formels (Kraiger, 2002).

Certains auteurs soulignent qu'il n'est pas souhaitable de créer de nouveaux instruments de mesure ou indicateurs spécifiquement dédiés à la formation et qu'il est plus judicieux et cohérent d'avoir recours aux indicateurs de performance organisationnelle déjà existants au niveau opérationnel (Spitzer, 1999; Redshaw, 2000; Watson, 1998; Kraiger, 2002). D'ailleurs, le recours aux indicateurs les plus surveillés dans l'entreprise peut être garant du lien existant entre l'évaluation de la formation et les objectifs de l'organisation. Comme le souligne Redshaw (2000) :

<sup>3</sup> Swanson, (1994 in Holton, 1996) propose des démarches d'analyses organisationnelles qui permettent l'identification de priorités de développement.

« ... l'évaluation (...) des résultats de la formation a plus de chances d'être un succès quand :

- la réflexion sur l'évaluation commence au moment où un besoin de formation est identifié;
- les cadres opérationnels sont impliqués dans le calendrier et le processus de mise en œuvre,
- une attention particulière est accordée à des critères autres que ceux liés à la formation;
- l'évaluation se concentre aux niveaux opérationnels. » (traduction libre, p. 248 )

### 8.3.3 Le choix des indicateurs et des méthodes

Le choix des indicateurs est primordial. Il est ainsi recommandé de cerner les objectifs que l'organisation poursuit et de baser l'évaluation sur les indicateurs que l'organisation valorise le plus (Spitzer, 1999; Holton, 1996; Torres & Preskill, 2001). On parle également d'« alignement » des objectifs de la formation à ceux de l'entreprise (Watson, 1998). Par la suite, la notion de performance devra être précisée : « Une évaluation de la performance de l'entreprise ne peut être efficace que dans la mesure où la performance en question a été clairement définie » (traduction libre, Sulsky & Keown, 1998<sup>4</sup>, p. 57). Nous ajouterons que ce choix est d'autant plus complexe que les formations sont non-techniques.

Pour connaître ces indicateurs, il faut se soucier des « valeurs d'affaire » de l'entreprise plutôt que de rester focalisé sur les « valeurs de la formation ». À cet effet, Spitzer (1999) précise qu'il faut avoir pour objectif de vouloir aider les personnes de l'entreprise à dépasser les problèmes qu'elles rencontrent et à atteindre les objectifs qui ont un sens réel pour elles et pour l'entreprise. En ce sens, les cadres « clients » de la formation doivent impérativement être rencontrés et impliqués dans le processus évaluatif.

Selon Redshaw (2000) et Pulley (1994), les évaluateurs devraient par la suite se poser les questions suivantes :

- qu'est-ce que ces preneurs de décision souhaitent améliorer?
- à partir de quel moment reconnaîtront-ils que l'opération est un succès?
- quels seront leurs indicateurs et critères pour mesurer ce succès?
- quels biais risquent d'influencer leur jugement à propos du programme?

Ces questions devraient permettre d'identifier les informations dont les preneurs de décision ont besoin. Elles pourront ensuite servir de base à la définition des

objectifs de formation et à son évaluation, tout en ayant identifié tous les facteurs qui ne relèvent pas de la formation, mais qui peuvent avoir un effet sur le résultat visé.

Une fois la liste des indicateurs arrêtée, Smith (2001) propose les questions suivantes afin de faire une dernière validation de leur pertinence :

- est-ce qu'au moment où les données sur l'indicateur retenu seront disponibles, l'information sera toujours utile?
- est-ce que suffisamment de données pourront être récoltées pour que la mesure soit fiable?
- est-ce que les biais pouvant affecter l'indicateur retenu ont bien été identifiés et isolés?

#### *Rechercher des indicateurs chiffrés à tout prix?*

Il existe dans les organisations une forte pression à exprimer les résultats d'évaluation de la formation sous la forme de données chiffrées. Mais selon Pulley (1994), cette tendance va à l'encontre de la logique de la démarche de l'évaluateur qui devrait choisir tour à tour les meilleurs indicateurs, qu'ils soient chiffrés ou non, en fonction des besoins de l'environnement organisationnel et de ce qui a le plus de sens pour les destinataires de l'évaluation. Elle fait remarquer que chiffre n'est pas synonyme d'objectivité. Elle propose une démarche d'évaluation « responsive » qu'elle décrit comme suit :

« En utilisant l'évaluation « responsive », vous ferez face à des situations où ce sont les informations quantitatives et orientées vers les résultats, qui sont les plus pertinentes, et vous ferez face à des moments où ce sont les informations qualitatives, orientées sur les processus qui seront les plus efficaces. La clé est d'être attentif à l'environnement et d'utiliser les techniques appropriées au bon moment. » (traduction libre, p. 20).

Cette position est largement partagée en évaluation de programme. En effet, plusieurs auteurs (entre autres, Phillips, 1997; Patton, 1997; Fitzpatrick, Sanders, & Worthen, 2002) préconisent le recours à des méthodes qualitatives et quantitatives, l'accent étant d'abord et avant tout mis sur une réponse adéquate aux questions auxquelles se doit de répondre l'évaluation.

Selon Kraiger (2002), les indicateurs de nature quantitative habituellement utilisés pour mesurer l'impact de la formation sur l'organisation, le bureau ou le groupe comprennent entre autres les chiffres de production, de ventes, les dépassements de délais, les durées d'arrêt de production, les temps de cycle, les échéances rencontrées, le taux de roulement de personnel, le nombre de plaintes de clients, le taux de réduction des erreurs, le chiffre d'affaires et le montant d'économies réalisées.

<sup>4</sup> Sulsky & Keown (1998) ont également écrit sur les systèmes de surveillance de la performance électroniques installés en entreprise et sur leur difficulté à capturer toutes les facettes de la performance du fait de son caractère complexe et multidimensionnel.



L'évaluation qualitative se réfère «aux nuances, aux détails et aux caractéristiques uniques qui font une différence par rapport à une échelle numérique» (traduction libre, Pulley, 1994). Elle est généralement descriptive, et peut se donner à lire sous la forme d'anecdotes. Elle s'avère apporter un bon complément aux données quantitatives ou peut dans certains cas apporter des explications permettant de faire une juste interprétation des chiffres. Il faut noter également qu'une recherche suggère que les croyances et les actions des personnes sont généralement plus influençables par les anecdotes et les métaphores que par les statistiques (Pulley, 1994).

Les indicateurs qualitatifs des résultats de la formation incluent, selon Kraiger (2002), la satisfaction des employés, l'implication dans l'entreprise, la satisfaction et la fidélisation des clients, et la qualité des idées nouvelles. À cela, Meignant (2003) ajoutera l'amélioration de l'adaptabilité des employés ou une évolution de la culture d'entreprise.

#### *Les méthodes d'évaluation de l'impact*

Les méthodes d'évaluation sont multiples. Phillips (1997) propose certaines méthodes déjà utilisées pour l'évaluation des comportements, à savoir les questionnaires de suivi, les plans d'action, les contrats de performance et les sessions de suivi de la formation. À celles-ci, il ajoute le suivi de la performance, selon des indicateurs préférentiellement déjà en place dans l'organisation, de manière à maintenir les coûts de l'opération à un niveau acceptable.

On pourrait aussi mentionner les évaluations «à 360°», qui sont plus souvent une évaluation de la performance que de la stricte mise en application des comportements appris en formation. Ces instruments, également appelés «bilans multiperspectives» ou «diagnostics à évaluateurs multiples», se présentent sous la forme d'un questionnaire auquel la personne à évaluer doit répondre ainsi que son responsable, et certains de ses collègues et collaborateurs. «Le principe est de comparer les réponses, guidées par des perceptions et des systèmes de représentation différents»<sup>5</sup>.

Mais quelles que soient les méthodes proposées, il importe d'abord et avant tout de bien définir ce sur quoi doit porter l'évaluation et asservir le choix des méthodes à ce que l'on souhaite évaluer. Ensuite, il faudra aussi mettre à profit les données déjà disponibles dans l'organisation à partir des indicateurs de gestion déjà en place. Ceux-ci devraient normalement refléter de manière relativement adéquate une bonne part des préoccupations de résultats de l'organisation. Ce n'est que par la suite que se pose le problème de la mise en place de mesures complémentaires.

### 8.3 LIMITES DES PRATIQUES D'ÉVALUATION DES RÉSULTATS DE LA FORMATION

Le principal problème auquel se heurte l'évaluation des résultats est certainement le fait que sa complexité et sa lourdeur font en sorte qu'elle est peu pratiquée. Les statistiques sur les pratiques d'évaluation de la formation déjà évoquées au début de ce document sont éloquentes à cet effet : selon les années, entre 7 et 11 % des entreprises américaines procéderaient à l'évaluation des résultats de la formation. Or, ces évaluations sont les plus susceptibles de supporter une réflexion en profondeur sur l'ingénierie de la formation : dans quelle mesure produit-on les impacts prévus et comment pourrait-on optimiser l'apport de la formation à l'organisation? En effet, les évaluations de la satisfaction, des apprentissages et des comportements ne peuvent répondre à ces questions de manière suffisante.

L'évaluation d'impact oblige à bien préciser ce que l'on souhaite comme apport de la formation au développement de l'organisation. Or, ceci peut être l'objet de débats fort intéressants sur ce que l'on entend par efficacité organisationnelle et quelles sont les approches de formation qui sont les plus susceptibles d'y contribuer :

*«Plutôt que d'essayer de former des personnes de façon à ce qu'elles sachent reproduire des comportements préétablis plus efficaces que ceux qu'elles adoptaient par le passé, l'efficacité organisationnelle peut être augmentée si les opportunités d'apprentissage favorisent l'acquisition et l'application de connaissances de façon à ce que l'autonomie et la rationalité des personnes soient respectées. En effet, il a été observé que l'efficacité organisationnelle est le résultat de décisions complexes et inter-corrélées que prennent les personnes plutôt que des simples performances comportementales. (...) L'atteinte d'objectifs préétablis, généralement formulés en termes comportementaux (...) agit comme une barrière vis-à-vis de l'efficacité de l'organisation.»* (traduction libre, Barrie & Pace, 1997, p.339).

Ceci rejoint des préoccupations aussi exprimées par Holton (1996), à savoir qu'on ne peut pas faire abstraction des besoins et des intérêts individuels des travailleurs dans la recherche de l'amélioration des pratiques de travail parce que les personnes ne sont pas de simples exécutants. Lee et Pershing (2000 in Tamkin et al., 2002) soulignent pour leur part que la capacité d'adaptation des personnes au changement de l'organisation dans son ensemble peut être un objectif plus pertinent dans le contexte économique actuel.

L'évaluation des résultats de la formation apparaît donc comme un moyen privilégié de provoquer certains débats de cette nature, par l'obligation qu'elle comporte de définir précisément les attentes des différents partenaires à l'égard de la formation au sein de l'organisation et leurs contributions respectives permettant d'assurer son impact.

<sup>5</sup> Farchy, I., Centre Inffo, [http://www.centre-inffo.fr/article.php3?id\\_article=748#top](http://www.centre-inffo.fr/article.php3?id_article=748#top).

## 9. L'ÉVALUATION DU RENDEMENT FINANCIER DE LA FORMATION

### 9.1 EN QUOI CONSISTE L'ÉVALUATION DU RENDEMENT FINANCIER

Sous le terme de rendement financier, se regroupent ici l'analyse bénéfices-coûts (B/C) et le retour sur l'investissement (ROI) qui constituent les deux façons les plus répandues de l'évaluer. Ces deux mesures du rendement financier répondent aux équations suivantes :

$$B/C = \text{bénéfice total} / \text{coût total}$$

$$\text{ROI} = (\text{bénéfice total} - \text{coût total}) / \text{coût total} \times 100\%$$

Dans un cas comme dans l'autre, le calcul du rendement financier est indissociable de l'attribution d'une valeur monétaire aux résultats de la formation, ce qui oblige les évaluateurs à extrapoler la valeur financière des résultats de la formation qu'ils observent (augmentation de la productivité, baisse du taux de roulement des employés, etc...).

Les sciences économiques ont procédé à des études tentant de situer l'apport de la formation tant au niveau macro que micro. Les études de niveau macro se sont plus particulièrement intéressées à l'apport de la formation à la performance économique nationale. Elles ont généralement démontré que la formation, via l'amélioration de la qualité du travail, constituait un des apports les plus importants à la croissance économique (Sturm, 1993 in Wang, Dou, & Li, 2002 ; Bartell, 2000). Les études micro, pour leur part, se sont limitées à l'apport de la formation à la rentabilité d'entreprises. Celles-ci ont aussi démontré un apport positif de la formation sur la productivité (Wang et al., 2002).

Le calcul du retour sur investissement dans le domaine des ressources humaines est complexe car il porte en partie sur des comportements humains. Cela ne signifie toutefois pas que les postulats sur lesquels repose le ROI soient automatiquement subjectifs. Certains auteurs, tels que Wang et al., (2002) ont développé des approches qu'ils estiment fiables et objectives, alors que la plupart acceptent une certaine subjectivité (Phillips, 1997, 1996; Parry, 1996; Kirkpatrick, 1998; Goldwasser, 2001; Hassett, 1992; Pine & Tingley, 1993) et défendent l'idée selon laquelle le calcul du ROI d'une formation est moins précis que le ROI d'une dépense de capital.

### 9.2 POURQUOI ÉVALUER LE RENDEMENT FINANCIER DE LA FORMATION

Le terme «*retour sur investissement*» a commencé à être utilisé dans le domaine de l'évaluation de la forma-

tion au début des années quatre-vingt. Cet intérêt peut s'expliquer par le besoin grandissant des départements de ressources humaines et de formation de justifier leur existence du fait de la morosité économique générale d'alors, ou encore par la popularité des approches dites de «*qualité totale*» (Pine & Tingley, 1993).

Le point de vue selon lequel la détermination du retour sur investissement (ROI) devrait être une priorité pour tous les responsables des ressources humaines est de plus en plus présent dans la littérature (Phillips & Stone, 2002; Phillips, 1997; Hassett, 1992; Parry, 1996). L'intérêt croissant des organisations pour des preuves chiffrées de la contribution de la formation aux résultats d'affaires indiquerait que les entreprises seraient ainsi passées d'une attitude où la formation se justifie d'elle-même à une autre où la formation est dédiée à l'obtention de résultats.

Phillips (1997) est certainement un des promoteurs les plus visibles et constants de ce niveau d'évaluation. Selon cet auteur, l'évaluation du retour sur l'investissement est de plus en plus nécessaire dans un contexte où les entreprises consacrent des sommes importantes à la formation<sup>6</sup>, où le niveau de compétition et de réduction des dépenses est élevé et où les départements de ressources humaines sont de plus en plus amenés à démontrer leur contribution aux résultats financiers.

Ceci ne signifie pas pour autant que l'évaluation du rendement financier est un incontournable. D'abord, le type d'organisation et sa finalité sont décisifs. «*Le type d'approche utilisé par les organisations pour évaluer et mesurer le rendement de la formation peut varier selon la nature et la fonction de l'organisation.(...) Pour les agences gouvernementales et les institutions sans but lucratif par exemple, la mesure du ROI se focalise souvent plus sur l'efficacité de l'accomplissement d'une mission spécifique que sur sa valeur monétaire.*» (traduction libre, Wang et al., 2002, p. 208).

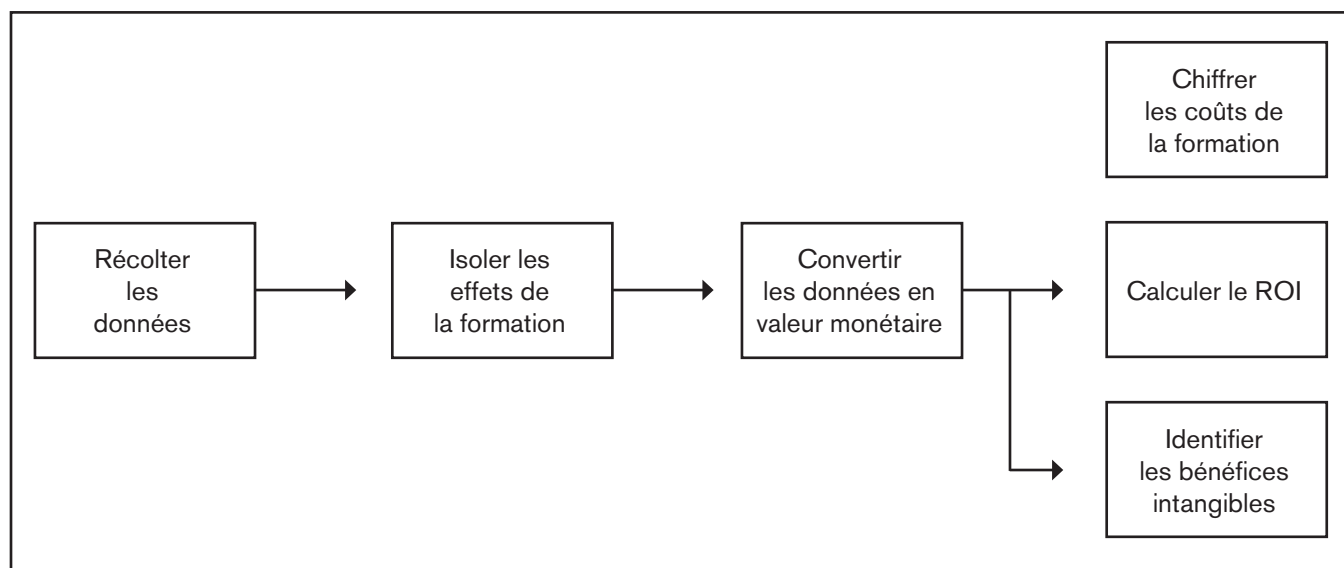
Ensuite, le type d'objectif visé par la formation est important. Plusieurs auteurs estiment qu'il est difficile, voire impossible de mesurer des résultats financiers pour des programmes de formation portant sur des sujets tels que par exemple la satisfaction des clients, le leadership, la communication, la motivation, la résolution de problème, ou la gestion du temps («*soft-skills training*») (Abernathy, 1999; Kirkpatrick, 1998; Pine & Tingley, 1993; Wang et al., 2002).

Hassett (1992) précise que dans la mesure où l'apport de la formation dépasse la seule contribution aux objec-

<sup>6</sup> Selon le rapport 2006 de l'ASTD, ce montant se chiffrerait à 109,25 milliards de dollars US.

## DIAGRAMME 4

### Les étapes du calcul du ROI de la formation (Phillips, 1997)



tifs de l'entreprise déjà pris en compte par les profits et pertes à court terme, il vaut mieux s'attarder aussi aux autres résultats non économiques tels que la conformité aux procédures ou un respect accru des règles de sécurité.

En somme, les formations qui se prêtent le mieux à l'évaluation du rendement financier sont celles qui « ont un cycle de vie très long, sont liées de très près à des objectifs opérationnels, sont centrées sur des objectifs stratégiques de l'entreprise, sont coûteuses et ont une forte visibilité » (traduction libre, Phillips cité par Worthen, 2001). Et encore là, leur rendement ne sera évalué qu'à des moments critiques justifiant une démarche de cette importance.

### 9.3 PRATIQUES ET TENDANCES ACTUELLES D'ÉVALUATION DU RENDEMENT FINANCIER

Comme en témoignent les résultats des enquêtes de l'ASTD<sup>7</sup> sur les pratiques d'évaluation des entreprises, l'évaluation du rendement financier de la formation demeure encore aujourd'hui très marginale. Il semble que la majorité des pratiques d'évaluation du rendement de la formation visent à démontrer quels résultats ont été obtenus suite à la formation, en comparaison avec les résultats enregistrés avant formation (Parry, 1996).

Les techniques développées pour calculer le ROI de la formation sont très nombreuses (Mosier, 1990 in Holton, 1996). Cette section du rapport propose d'abord l'approche de Phillips (1997), puis deux approches simplifiées. Suit l'approche systémique de Wang et al., (2002). Elle se termine par des critères afin d'aider au choix de la bonne approche.

#### 9.3.1 Le calcul du ROI selon Phillips (1997)

On peut visualiser les étapes principales du processus d'évaluation du ROI proposé par Phillips de la façon illustrée au diagramme suivant. Selon l'auteur, ce modèle serait actuellement mis en pratique par des milliers d'organisations.

La première étape consiste à récolter les données sur les résultats (niveau 4) après la formation. Ces données peuvent être de nature quantitative (qualité des produits, temps, coûts, etc.) ou qualitative (habitudes de travail, climat de travail, etc.). Elles peuvent découler de comparaisons pre- et post formation ou avec un groupe contrôle.

La seconde étape consiste à isoler l'effet de la formation. Les 10 stratégies que propose l'auteur pour y arriver ont déjà été décrites à la section 6.2.2. Rappelons que certaines d'entre elles peuvent faire appel à la subjectivité de certains acteurs.

Les données sont ensuite converties en valeur monétaire. À cet effet, l'auteur propose 10 façons :

- ↳ les résultats sont convertis en contribution aux profits ou en économies de coûts, ces données étant déjà disponibles dans l'organisation;
- ↳ les améliorations de la qualité sont traduites en économies de coûts;
- ↳ lorsque la formation permet des économies de temps des employés, celles-ci sont traduites en valeur monétaire selon les coûts de main d'œuvre;
- ↳ il est aussi possible de faire appel à un historique des coûts pour évaluer l'apport d'une amélioration;
- ↳ dans certains cas, les jugements d'experts internes ou externes peuvent être sollicités pour évaluer la valeur d'une amélioration;

- on peut avoir recours à des bases de données externes sur la valeur ou le coût de certains objets et partir de là pour estimer l'apport financier de la formation;
- enfin, on peut s'appuyer sur une estimation de la valeur de l'amélioration selon :
  - les participants à la formation;
  - leur superviseur;
  - la haute gestion;
  - ou le personnel des ressources humaines.

La quatrième étape consiste à chiffrer les coûts de la formation. Cette opération doit prendre en compte :

- le coût de conception et élaboration du programme, préférablement étalé sur sa durée de vie;
- le coût de tout le matériel distribué aux participants;
- le coût du formateur, incluant son temps de préparation et de formation;
- le coût des locaux de formation;
- les coûts de logements, repas et déplacements des participants, s'il y a lieu;
- les salaires et avantages sociaux des participants;
- les coûts d'encadrement et frais généraux attribués au programme;
- les coûts éventuels d'analyse des besoins et d'évaluation.

La cinquième étape consiste à calculer soit le rapport bénéfico-coûts, soit le rendement sur l'investissement selon les formules déjà définies. La sixième étape consiste enfin à identifier les bénéfices intangibles tels que la satisfaction au travail, l'amélioration du travail d'équipe ou les divers effets sur le climat de travail. Ces derniers bénéfices n'ont pas été traduits en valeur monétaire, mais doivent néanmoins être pris en considération.

L'auteur attire aussi l'attention sur certaines conditions critiques de mise en œuvre de l'évaluation du rendement financier qui peuvent en déterminer la réussite. Celles-ci concernent :

- l'existence d'un énoncé de politique relatif à la formation et au développement centrés sur les résultats;
- la mise à disposition de procédures et modes d'emploi pour différents aspects du processus d'évaluation du rendement financier;
- le développement des compétences du personnel impliqué dans l'évaluation;
- des stratégies afin d'obtenir l'adhésion des gestionnaires;
- des mécanismes d'appui technique à l'élaboration des instruments, à l'analyse des données et à la définition d'une stratégie d'évaluation;

- un appui à l'amélioration de la culture de l'organisation en vue d'accorder plus d'importance aux résultats.

### 9.3.2 Deux méthodes simples

Voici deux méthodes plus simples d'évaluation du rendement financier, soit l'analyse d'investissement en formation de Hasset (1992) et la méthode du retour sur les attentes de Hodges (Goldwasser, 2001).

L'analyse de l'investissement en formation de Hasset (1992) est présentée comme un compromis entre les méthodes complexes et les méthodes simples, avec des résultats relativement fiables. Les étapes sont les suivantes :

- déterminer les informations dont l'entreprise a besoin et utiliser la même terminologie que celle utilisée par les autres départements;
- utiliser la méthode la plus simple et la moins chère possible pour trouver l'information dont on a besoin; celle-ci devrait porter sur les revenus attribuables à la formation découlant de la hausse des ventes, une productivité accrue, une baisse des erreurs, la rétention des clients et la rétention des employés; faire appel à un gestionnaire ou à un expert de l'entreprise pour quantifier les effets de la formation;
- procéder à l'analyse le plus rapidement possible;
- diffuser les résultats, que la formation soit rentable ou pas, afin de faire passer le message d'un service de formation soucieux de prendre en considération les résultats globaux de l'entreprise.

La méthode du « *retour sur les attentes* » (ROE : Return On Expectations) de Hodges (Goldwasser, 2001) consiste à recueillir les attentes du ou des commanditaires de la formation lors d'un entretien d'une quinzaine de minutes et d'établir à partir de là, les objectifs d'apprentissage de la formation. Une fois la formation terminée, ces personnes sont à nouveau interrogées et on leur demande de quantifier les résultats obtenus. « *Si le vice-président dit que ses employés perdent moins de temps, je vais lui demander d'attacher une valeur monétaire à ce temps gagné.* » Ces données sont ensuite utilisées comme des preuves crédibles dans un calcul de ROI.

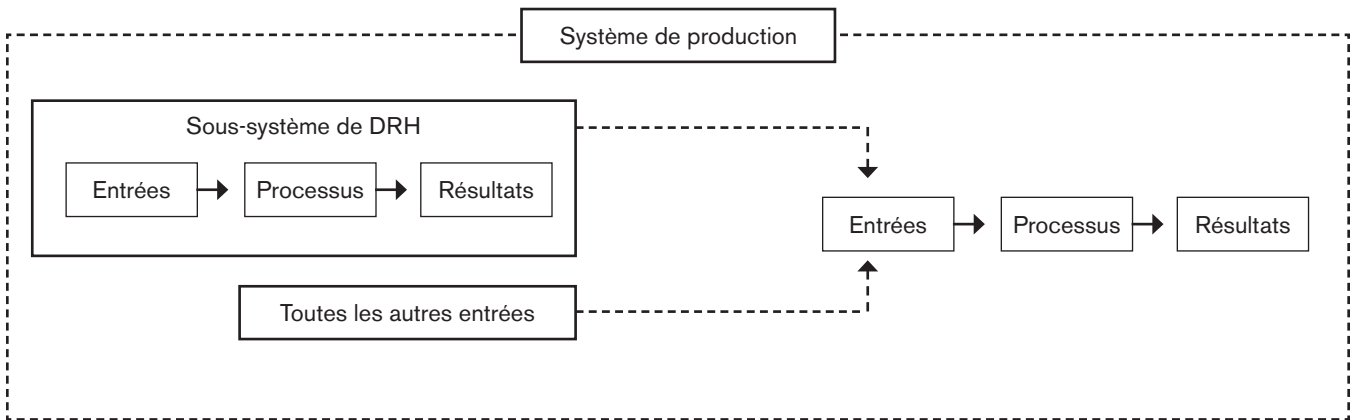
La méthode a l'avantage d'être rapide et peu coûteuse. Elle est assez similaire à l'approche développée par Andersen Consulting Education basée sur les attentes des managers. L'idée maîtresse est que ces derniers savent quels sont les compétences et les comportements dont leurs personnels ont besoin pour faire leur travail, et qu'en les formant dans ces domaines, des résultats financiers vont être générés (Abernathy, 1999).

Par ailleurs, force est de constater que ces méthodes sont loin d'être rigoureuses et nous amènent progressivement à glisser vers une simple évaluation informelle du rendement.



## DIAGRAMME 5

### Vue systémique de l'apport du DRH et de son impact potentiel (Wang et al., 2002)



#### 9.3.3 L'approche systémique de Wang et al. (2002)

Wang et al. (2002) proposent une approche systémique de l'apport de la formation à la production qui devrait, par surcroît, éliminer la subjectivité de l'évaluation du rendement financier de la formation.

Les «entrées» du sous-système DRH comprennent les employés et les investissements dans les RH en terme de dollars, de temps, d'équipement, etc... Le terme «processus» se réfère aux interventions des RH telles que les formations. Les «résultats» renvoient généralement à une augmentation de la performance de la part des employés impliqués :

*«Ce sont les relations entre le sous-système DRH et le système de production, et les interactions entre les deux qui sont à l'origine des défis de la mesure du ROI. En d'autres termes, une bonne entrée du sous-système DRH dans le système de production est une condition nécessaire, mais pas suffisante pour produire le résultat final. Les résultats du sous-système de DRH ne sont qu'une des composantes des entrées du système de production. Le résultat final de l'entreprise dépend non seulement de ces entrées, mais aussi des autres entrées du système et de l'efficacité des processus de l'entreprise.»*  
(traduction libre, Wang et al., 2002, p. 213).

Il s'agit donc, pour arriver à la mesure du ROI d'un programme DRH qui soit complète, d'identifier et de mesurer l'impact du sous-système DRH sur les résultats du système de production. Ceci se fait à l'aide d'équations mathématiques.

Cette approche, selon les auteurs, aboutit à un résultat objectif, car ses équations sont alimentées par des variables objectives. Mais il est recommandé de combiner ces données objectives avec une enquête organisationnelle (de type évaluation à 360°) ou un sondage.<sup>8</sup> Cette recherche de la variété des sources des données et de leur

forme (entretiens, «focus groups», sondages, ...) est une des recommandations que l'on retrouve fréquemment dans la littérature (Phillips, 1996; Delahoussaye, 2002).

Selon les auteurs, les avantages de cette approche systémique sont multiples :

- ↳ la subjectivité est grandement réduite; les variables utilisées sont issues des bases de données de l'organisation ou obtenues auprès des responsables à même de faire des estimations sur la base d'expériences de travail;
- ↳ les bénéfices non monétaires des interventions de DRH, autrement très difficiles à quantifier par l'approche conventionnelle de calcul du ROI, peuvent être évalués;
- ↳ il est possible de séparer objectivement les impacts liés au DRH des impacts qui n'en sont pas liés;
- ↳ cette approche peut être utilisée pour faire des prévisions pour de futurs investissements sur des programmes de DRH et ce, avec une précision statistique;
- ↳ enfin, ce système permet l'évaluation des résultats de plusieurs programmes de DRH conjointement.

Toujours selon les auteurs, ses principales limites sont par ailleurs les suivantes :

- ↳ «Ce système ne se prête peu ou pas du tout à des évaluations portant sur un unique programme de DRH proposé sur du court terme. Ceci est lié au fait qu'un programme proposé sur une période courte, lorsqu'il est examiné dans un système plus large, peut avoir un impact négligeable sur les résultats finaux de l'ensemble.» (traduction libre, Wang et al., 2002, p. 219);
- ↳ en conséquence, ce système peut être vu comme coûteux et surtout applicable dans des organisations dépassant une certaine taille;

<sup>8</sup> Deux exemples de mise en œuvre de cette approche sont présentés dans l'article.

- sa mise en oeuvre nécessite que les personnels en charge aient certaines compétences dans le domaine de la pensée systémique et des analyses quantitatives, ce qui fait souvent défaut aux praticiens des ressources humaines.

#### 9.3.4 Quelques critères de sélection de la bonne méthode.

L'identification de la méthode la plus adéquate pour faire une évaluation du rendement financier mérite réflexion. Le choix de cette méthode devrait se faire entre autres en fonction des critères suivants :

- le niveau d'attente des gestionnaires clients de l'évaluation du rendement financier; devant la gamme étendue des méthodes proposées et de leurs niveaux de rigueur, il convient dès le départ de préciser avec ceux qui souhaitent une telle évaluation leur niveau d'attente à cet égard; rien ne sert de mettre en place un processus qui risque d'être disqualifié en fin de parcours;
- les données disponibles : de nombreux chercheurs écartent l'argument selon lequel le recueil de données chiffrées n'est pas toujours possible (Phillips, 1996; Spitzer, 1999; Newby, 1992 in Tamkin et al., 2002) et défendent le point de vue selon lequel chaque entreprise dispose de centaines d'indicateurs chiffrés de son efficacité organisationnelle : l'efficacité de la production, les niveaux des inventaires, le nombre d'accidents, le nombre d'appels non aboutis, etc... Notons également que « *les experts de l'évaluation ont commencé à questionner la pertinence du choix d'un critère unique, par exemple l'augmentation des profits, pour justifier les dépenses de formation.* » (traduction libre, Goldwasser, 2001);
- l'ampleur du programme de formation et la pertinence de l'évaluation, la démarche pouvant aisément devenir très coûteuse et même démesurée eu égard au budget total du programme; avant d'entreprendre une évaluation, il est donc important d'identifier la pertinence et la valeur stratégique des informations qu'on en attend et de la comparer avec une estimation du coût de l'évaluation.

Il faut arriver à trouver le juste équilibre entre une évaluation du rendement de la formation rigoureuse mais coûteuse, exigeante et longue (les résultats de l'évaluation seront-ils encore utiles au moment où ils seront prêts?) et une forme plus simple mais moins rigoureuse. En effet, Bee et Bee (1994 in Tamkin et al., 2002) suggèrent une approche pragmatique plutôt que perfectionniste. Ils disent que toutes les analyses bénéfices-coûts sont imparfaites, mais que si leurs limites sont rendues explicites, il est alors possible d'en tirer des informations pertinentes pour l'entreprise.

#### 9.4 LIMITES DES PRATIQUES D'ÉVALUATION DU RENDEMENT FINANCIER

Le calcul du ROI d'une formation est sujet à de nombreuses contraintes, et la littérature sur le sujet reflète bien les dilemmes auxquels sont soumis les évaluateurs.

D'entrée de jeu, ce niveau d'évaluation supplémentaire n'est ni possible ni pertinent pour certaines formations qui ne visent pas un apport financier. Et les valeurs monétaires propres à l'évaluation du rendement financier n'expriment pas toujours de manière adéquate la véritable valeur ajoutée de certaines formations.

Mais ces réserves étant exclues, une des questions les plus récurrentes sur le sujet est la suivante : en prenant en compte les contraintes de l'entreprise (de temps et de coûts notamment), est-il possible de calculer le ROI d'une formation de façon suffisamment fiable et rigoureuse pour que les conclusions de cette évaluation puissent lui être utiles et que l'opération s'avère rentable?

Certaines méthodes plus objectives et rigoureuses sont proposées. Elles sont en conséquence plus lourdes et plus coûteuses et exigent des compétences peu souvent disponibles à l'interne. D'autres proposent des méthodes simplifiées. Mais cette simplification se fait de toute évidence au prix de la rigueur, en faisant dans bien des cas appel à la subjectivité, que ce soit pour isoler l'effet de la formation ou attribuer une valeur financière aux résultats. Or, ces évaluations moins rigoureuses peuvent-elles convaincre ceux qui veulent des données dures pour démontrer et quantifier l'apport de la formation?

En somme, ce n'est pas sans raison qu'il se fait si peu d'évaluation du rendement financier de la formation dans les entreprises, malgré la pertinence des informations qui pourraient être générées.

La voie la plus simple consiste certainement à maintenir de manière constante un pont entre les systèmes de formation et de production. Ainsi, il faudrait que les décideurs des systèmes de formation et de production conviennent de manière très explicite, dès la conception de la formation, de l'arrimage des objectifs et stratégies de formation aux améliorations de performance (et non seulement de comportement) souhaitées par le système de production. Ceci étant fait, le système de production devrait, dès le départ, être en mesure de rattacher les effets escomptés de la formation à certains indicateurs qu'il utilise déjà dans sa gestion courante. Lorsqu'une évaluation du rendement financier sera requise, il sera donc capable, sans coûts additionnels, de fournir des données sur les indicateurs quantitatifs qui témoigneront de l'évolution de la productivité et de la qualité et qu'il sera en mesure de traduire en valeur économique. Cette trame de fond devrait valoir pour toutes les formations visant un impact financier.

## 10. CONCLUSION

Il existe une littérature scientifique relativement importante sur l'évaluation de la formation dans l'entreprise. Ces quelques pages en résument ce qui nous est apparu l'essentiel pour les formateurs et cadres qui œuvrent dans ce domaine. Elle ne prétend pas être exhaustive, ce qui aurait dans certains cas obligé à faire état de débats ou de nuances qui sont peu susceptibles de guider l'action ou les décisions à court ou moyen terme de ces acteurs. Par ailleurs, nous n'avions pas non plus l'ambition de présenter un manuel technique expliquant de manière précise comment évaluer la formation dans l'entreprise, bien que le besoin existe en langue française.

Il est important de reconnaître que cette littérature est fortement imprégnée du modèle de Kirkpatrick (1959). En effet, celle-ci est principalement centrée sur «*ce que l'on doit évaluer*», ces objets d'évaluation correspondant dans la plupart des cas aux fameux quatre niveaux. Or, l'évaluation de la formation se situe à l'intérieur d'un plus vaste domaine qui est celui de l'évaluation de programme, qu'il s'agisse de programme de formation, d'éducation, de santé, de développement social ou autres. Et le «*comment évaluer*» et «*à quelles fins évaluer*» occupe une place importante dans cette littérature. Ceci témoigne entre autres de la préoccupation d'assurer l'utilité de l'évaluation qui est apparue ces dernières années comme un enjeu important (Sridharan, 2003; Leviton, 2003). Par exemple, le modèle centré sur l'utilisation de Patton (1997) ne définit pas du tout a priori ce qui doit être évalué, mais propose plutôt une démarche qui vise à évaluer ce qui correspond de manière précise à ce dont les utilisateurs de l'évaluation ont besoin. Cela pourra éventuellement coïncider avec les niveaux de Kirkpatrick, mais on évitera certainement la situation que Kraiger (2002) évoque avec un peu d'ironie :

*«Quand des gestionnaires de formation disent «Nous avons toujours fait le niveau 2, et maintenant nous devons faire le niveau 3», je leur demande pourquoi ils doivent mesurer les comportements, ou comment ils tireront avantage de cette information. Dans presque tous les cas, ils ne peuvent me donner une raison autre que «C'est parce que c'est ce qui vient après». Si je leur demande comment l'information leur sera utile, ou ce qu'ils espèrent gagner de l'évaluation, ils sont habituellement plus capables de donner des raisons. Fréquemment, ces raisons peuvent ou pas nécessiter une évaluation du niveau 3, et le plan d'évaluation que nous développons sera fondé sur ce qu'ils espèrent faire plutôt que sur le niveau qui vient après.»* (traduction libre, p.335).

La version de l'évolution de l'évaluation que proposent Guba et Lincoln (1989) est aussi fort éloquente quant à l'importance des finalités de l'évaluation et des processus qui en découlent. Selon ces auteurs, la première génération de l'évaluation (1880-1930) s'est d'abord centrée sur la mesure, s'inscrivant ainsi dans un contexte où les sciences sociales visaient à être pleinement reconnues comme «*sciences*». Pendant la seconde génération (1930-1960), on ajoute à la mesure la description qui permet d'associer les résultats au processus qui les génère et de mieux les mettre en contexte. Lors de la troisième génération (1960-1990), les évaluateurs sont amenés à se départir de leur position «*objective*» pour reconnaître et intégrer la nécessité de porter des jugements sur la valeur des programmes, jugements qui, selon les modèles qui seront par la suite élaborés, pourront être portés par eux ou d'autres acteurs des programmes. Enfin, la quatrième génération s'écarte définitivement du positivisme pour définir le processus évaluatif comme consistant à répondre aux besoins et construire avec les acteurs des programmes. Les approches de cette génération reconnaissent que les réalités vécues par les différents acteurs peuvent être différentes. Ce qui compte n'est plus l'objectivité en soi, mais la possibilité de construire avec les parties prenantes une vision consensuelle de la réalité de manière à faire avancer les choses.

Dans cette perspective, Preskill et Torrès proposent la définition suivante de l'évaluation :

*«Nous voyons l'enquête évaluative comme un processus continu pour approfondir et comprendre les enjeux critiques de l'organisation. C'est une approche de l'apprentissage qui est pleinement intégrée aux pratiques de travail de l'organisation et comme telle, elle suscite : (1) l'intérêt des membres de l'organisation et leur habileté à explorer les enjeux critiques en utilisant la logique évaluative; (2) l'implication des membres de l'organisation dans les processus évaluatifs, et (3) la croissance personnelle et professionnelle des individus dans l'organisation.»* (1999 in Russ-Eft & Preskill, 2001, p.5) (traduction libre).

On comprendra qu'une telle définition de l'évaluation est bien loin de celle de Kirkpatrick (1959). Il ne s'agit plus d'évaluer selon des critères prédéterminés (satisfaction, apprentissages, comportements et résultats) qui correspondent surtout aux préoccupations de l'organisation, mais bien d'accompagner l'organisation et ses acteurs dans leur réflexion sur leurs pratiques.

D'une certaine façon, la définition de l'évaluation de Preskill et Torrès reflète l'évolution actuelle de l'ingénierie de la formation dans l'entreprise. D'une formation en salle reproduisant le modèle scolaire, on passe de plus en plus à une formation où la participation active de l'apprenant est requise et où l'on accorde une importance grandissante à son appropriation de ce qui lui est proposé. L'«*empowerment*» des employés au sein de l'organisation et l'évolution de l'organisation vers une organisation apprenante et la mise en place de communautés de pratique sont, depuis quelques années, des thèmes d'actualité. Si nous sommes portés à croire que les pratiques informelles d'évaluation de la formation font, du moins en partie, état de cette évolution des approches de formation, les pratiques formelles d'évaluation sont pour une bonne part restées près d'un modèle de mesure qui se veut la plus objective possible, et dont les préoccupations se limitent à celles de l'organisation et au rendement de la formation. Sans nier la pertinence et les apports de telles pratiques, il n'en demeure pas moins que celles-ci font peu de place, par exemple, à l'autoévaluation et à la prise en compte du point de vue de l'apprenant qui poursuit une démarche de développement personnel et professionnel. Pourtant, ces dernières sont essentielles pour supporter le développement de l'autonomie, de pratiques réflexives et d'autocontrôle de la part des employés. De plus, elles peuvent contribuer, de toute évidence, à identifier les écarts entre les attentes formelles de performance et les contraintes du milieu de travail de même que les voies de solution appropriées.

La conjoncture démographique et économique devrait inciter les entreprises à accorder plus d'importance à l'apprenant et son développement personnel. Les écrits de Davenport (1999) à ce sujet sont fort intéressants. Selon cet auteur, l'employé n'est pas «*le bien le plus précieux*» de l'entreprise qui investit en lui par la formation. C'est plutôt lui qui investit son capital humain dans l'entreprise. Dans la mesure où ce travailleur investisseur aura la conviction que ce placement lui est favorable, son engagement sera assuré. Une telle approche est loin d'être en contradiction avec les intérêts de l'entreprise qui peut ainsi recruter et fidéliser les meilleurs éléments, ou encore redynamiser ceux que l'on aura progressivement démotivés. Il est évident que la littérature actuelle sur l'évaluation de la formation fait peu de place à des processus évaluatifs également centrés sur le développement de l'entreprise et du capital humain des employés.

Enfin, le déplacement vers l'apprenant et le contexte de ses pratiques, entre autres en réaction à la problématique du transfert des apprentissages, incite de plus en plus à se préoccuper de l'évaluation de la performance de l'employé en poste de travail. Rummel et Brache (1995 in Broad, 2005) identifient 6 déterminants de la performance : (i) des spécifications de performance claires; (ii) des ressources et un support nécessaire; (iii) un système de récompense approprié; (iv) des feedback

pertinents et à temps; (v) des capacités individuelles adéquates pour la tâche; et (vi) les connaissances et habiletés requises. Selon ces auteurs, les deux derniers déterminants, sur lesquels la formation peut agir, ne comptent que pour 15 à 20 % des problèmes de non-performance.

De tels constats ne devraient pas, pour autant remettre en question la pertinence de l'évaluation de la formation, mais plutôt inciter à élargir ces pratiques vers une évaluation de la performance qui tiendrait compte des différents facteurs qui l'influencent, incluant la formation. En effet, quoi qu'il en soit, il demeure difficilement acceptable que l'on poursuive des activités de formation avec l'ampleur qu'elles ont actuellement sans évaluer les produits et résultats de ces actions et qu'on puisse ainsi en améliorer la qualité. Par ailleurs, ces constats remettent en cause une évaluation de la formation trop exclusivement centrée sur les seuls produits et résultats exclusivement attribuables à la formation. Il serait certainement souhaitable que l'on décloisonne l'évaluation de la formation et qu'elle prenne de plus en plus en compte l'évaluation du contexte des pratiques et en arrive ainsi à accroître ses interactions avec les autres fonctions de l'organisation.

Il va de soi qu'une approche plus «*sur mesure*» de l'évaluation et moins limitée à la seule formation est de loin susceptible de mieux profiter à l'organisation dans ses processus de prise de décision. Et ce, non seulement pour les décisions qui concernent la formation, mais aussi celles qui conditionnent la performance effective des employés en poste de travail et la gestion plus générale des ressources humaines et de la productivité de l'entreprise.



# BIBLIOGRAPHIE

- Abernathy, D. J. (1999). Thinking outside the evaluation box. *Training & Development*, 53(2), 18-23.
- Alliger, G., Tannenbaum, S. I., Bennett, W., Traver, H., & Shortland, A. (1997). A meta-analysis on the relations among training criteria. *Personnel Psychology*, 50, 341-358.
- Alliger, G. M., & Horowitz, H. M. (1989). IBM takes the guessing out of testing. *Training & Development*, 43(1), 69.
- Alliger, G. M., & Janak, E. A. (1989). Kirkpatrick's levels of training criteria: thirty years later. *Personnel Psychology*, 42(2), 331.
- Ammons, D. N., & Niedzielski-Eichner, P. A. (1985). Evaluating Supervisory Training in local Government: Moving Beyond Concept to a Practical Framework. *Public Personnel Management*, 14(3), 311-330.
- Archambault, G. (1997). La formation de suivi et le transfert des apprentissages. *Gestion*, 22(3), 120.
- Arvey, R., & Cole, D. A. (1989). Evaluating change due to training. In Jossey-Bass (Ed.), *Training and development in organizations* (pp. 89-117). San Francisco, CA: Goldstein, I.
- Aubret, J., & Gilbert, P. (2003). *L'évaluation des compétences*. Sprimont, Belgique : Mardaga éditeur.
- Aubret, J. T., & Gilbert, P. (2003). *L'évaluation des compétences* (Mardaga ed.). Sprimont, Belgique.
- Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: a review and directions for future research. *Personnel Psychology*, 41(1), 63.
- Barrie, J., & Pace, W. R. (1997). Competence, efficiency, and organizational learning. *Human Resource Development Quarterly*, 8(4), 335-342.
- Barzucchetti, C., & Claude, J. (1995). *Évaluation de la formation et performance de l'entreprise* (Éditions Liaisons ed.). Rueil-Malmaison, France.
- Bassi, L. J., & Van Buren, M. E. (1999). The 1999 ASTD state of the industry report. *Training and development*, 53(1).
- Bedingham, K. (1997). Proving the effectiveness of training. *Industrial and Commercial Training*, 29(3), 88-91.
- Beehr, T. A., Ivanitskaya, L., Hansen, C. P., Erofeev, D., & Gudanowski, D. M. (2001). Evaluation of 360 degree feedback ratings: relationships with each other and with performance and selection predictors. *Journal of Organizational Behavior*, 22, 775-788.
- Bernthal, P. R. (1995). Evaluation that goes the distance. *Training & Development*, 49(9), 41.
- Betcherman, G. K., Davidman, K., & McMullen, K. (1998). *La formation et la nouvelle économie. Un rapport de synthèse*. Ottawa : Réseaux canadiens de recherche en politiques publiques
- Betcherman, G. N., Leckie, K., & McMullen, K. (1997). *Developing skills in the canadian workplace. The results of the Ekos Workplace Training Survey*. Ottawa: Canadian Policy Research Network.
- Billet, S., Cooper, M. (1997) returns to enterprises from investment in VET: Review of research. National Center for Vocational Education Research.
- Bloom, B. (1956). *Taxonomy of educational objectives: the cognitive domain*. New York: McKay.
- Blouin, C., Lapierre, G., Lapierre, R., & Pronovost, D. (2000). *Évaluation formative de la Loi favorisant le développement de la main d'œuvre. Résultats des entrevues, volet qualitatif*. Québec : Ministère de la Solidarité sociale, Direction de la Recherche, de l'évaluation et de la statistique.
- Boverie, P., Mulcahy, D. S., & Zondlo, J. A. (1994). *Evaluating the Effectiveness of Training Programs*. Paper presented at the The 1994 Annual: Developing Human Resources.
- Boyle, M. A., & Crosby, R. (2004). Academic Program Evaluation: Lessons from Business and Industry. 34(3).
- Bramley, P. (1996). Training in its organizational context. In R. Bennett (Ed.), *Evaluating training effectiveness. Second Edition* (pp. 64-70). Cambridge: McGraw-Hill Publishing Company.
- Brinkerhoff, R. O. (1989). Using evaluation to transform training. In A. E. Association (Ed.), *Evaluating training programs in business and industry* (American Evaluation Association ed., Vol. 44, pp. 5-20). San Francisco, CA: Jossey-Bass Publishers.
- Brinkerhoff, R. O. (1989). *Evaluating training programs in business and industry* (Vol. 44). San Francisco: Jossey-Bass.
- Brinkerhoff, R. O., & Apking, A. M. (2001). *High impact learning. Strategies for leveraging business results from training*. New York: Basic Books.
- Broad, M. L. (2005). *Beyond transfer of training : engaging systems to improve performance* (Pfeiffer ed.). Canada: John Wiley & Sons Inc
- Brown, K. G. (in press). Training evaluation. In *The Encyclopedia of Industrial and Organizational Psychology* (S.Rogelberg ed.).
- Brown, K. G., & Gerhardt, M. G. (2002). Formative evaluation: an integrative practice model and case study. *Personnel Psychology*, 55(4), 951.
- Bushnell, D. S. (1990). Input, process, output: a model for evaluating training. *Training and Development*, 44(3), 41-43.
- Carnevale, A., & Schulz, E. (1990). Return on investment: accounting for training. *Training and Development Journal*, 44(7), 23-29.
- Davenport, T. (1999). *Human capital: what it is and why people invest it*. San Francisco: Jossey-Bass.
- Delahoussaye, M. (2002). Show me the results. *Training*, 39(3), 28-29.
- Delahoussaye, M., Ellis, K., & Bolch, M. (2002). Measuring corporate smarts. *Training*, 39(8), 20.
- Dennerly, M. (2001). *Évaluer la formation. Des outils pour optimiser l'investissement formation* (ESF Éditeur ed.).
- Dionne, P. (1996). The evaluation of training activities: A complex issue involving different stakes. *Human Resource Development Quarterly*, 7(3), 279-286.
- Dunberry, A. (2006). *Analyse des pratiques d'évaluation de la formation et de son rendement dans des entreprises performantes. Rapport détaillé*: CIRDEP UQAM.
- Erickson, P. R. (1990). Evaluating training results. *Training and Development Journal*, 44(1), 57.
- Figgis, J., Alderson, A., Blackwell, A., Butorac, A., Mitchell, K., & Zubrick, A. (2001). *What convinces enterprises to value training and learning and what does not? A study in using case studies to develop cultures of training and learning*: Leabrook.
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2002). Alternative evaluation approaches : a summary and comparative analysis. In



- Evaluation. *Alternative approaches and practical guidelines*. (Alyn & Bacon ed., pp. 152-167).
- Garand, D. J., & Fabi, B. (1996). La formation et le développement des ressources humaines dans les PME : un perpétuel défi. In L. J. Filion & D. Lavoie (Eds.), *Systèmes de soutien aux sociétés entrepreneuriales. Actes du 13<sup>ème</sup> colloque du Conseil canadien de la PME et de l'entrepreneuriat (CCPME)* (Vol. 2, pp. 29-45). Montréal : GREPME, DAE-UQTR.
- Garnier, B., Gasse, Y., & Parent, A. (1991). Évaluation des retombées d'un programme de formation en gestion de dirigeants de PME. *Relations Industrielles*, 46(2), 357-375.
- Gaudine, A. P., & Saks, A. M. (2004). A longitudinal quasi-experiment on the effects of posttraining transfer interventions. *Human Resource Development Quarterly*, 15(1), 57-76.
- Geertshuis, S., Holmes, M., Geertshuis, H., Clancy, D., & Bristol, A. (2002). Evaluation of workplace learning. *Journal of Workplace Learning*, 14(1), 11-18.
- Ghesquière, I., & Vermelle, M.-C. (2002, 6-8 juin). *Question et évaluation des compétences : le cas d'un constructeur automobile*. Paper presented at the La co-ordination des compétences et des connaissances dans l'industrie automobile, Paris, France.
- Goldstein, I. L., & Ford, J. K. (2002). *Training in organizations. Fourth Edition*. (Wadsworth ed.).
- Goldwasser, D. (2001). *Beyond ROI Training*, 38(1).
- Grove, D., & Ostroff, C. (1990). Program evaluation. In K. Wexley & J. Hinrichs (Eds.), *Developing human resources*. Washington, DC: BNA Books.
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park: Sage.
- Haccoun, R. R., Jeanrie, C., & Saks, A. M. (1997). Concepts et pratiques contemporaines en évaluation de la formation : vers un modèle de diagnostic des impacts *Gestion*, 22(31), 108.
- Hadden, D., D. (2002). *Beyond ROI: measuring and optimizing the true impact of training and IT interventions on the enterprise*. Durham, NC.
- Halverson, S. K., Tonidandel, S., Barlow, C., & Dipboye, R. L. (2002, Avril). *Self-other agreement on a 360-Degree leadership evaluation*. Paper presented at the 17th Annual Conference of the Society for Industrial and Organizational Psychology, Toronto, Canada.
- Hamblin, A. C. (1974). *Evaluation and control of training*. London: McGraw-Hill.
- Hassett, J. (1992). Simplifying ROI. *Training*, 29(9), 53.
- Holton, E. F., Bates, R. A., & Ruona, W. E. A. (2000). Development of a generalized learning transfer system inventory. *Human Resource Development Quarterly*, 11(4), 333-359.
- Holton, E. F. (1996). The flawed four-level evaluation model *Human Resource Development Quarterly*, 7(1), 5-21.
- Joint Committee on standards for educational evaluation. (1994). *The program evaluation standards : how to assess evaluations of educational programs (2nd ed.)*. Thousand Oaks, CA: Sage.
- Kaufman, R., & Keller, J. M. (1994). Levels of evaluation: Beyond Kirkpatrick. *Human Resource Development Quarterly*, 5(4), 371-380.
- Kaufman, R., & Keller, J. M., Watkins, R. (1995) What works and what doesn't: evaluation beyond Kirkpatrick. *Performance and instruction*, 35 (2), 8-12.
- Kearns, P., & Miller, T. (1997). *Measuring the impact of training and development on the bottom line*. London, UK: Pitman Publishing.
- Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. *Journal of the American Society for Training and Development*, 13(11), 3-9.
- Kirkpatrick, D. L. (1996). Invited reaction: reaction to Holton article. *Human Resource Development Quarterly*, 7(1), 23-25.
- Kirkpatrick, D. L. (1998). *Evaluating training programs. Second Edition*. San Francisco, CA: Berrett-Koehler Publishers Inc.
- Kraiger, K. (2002). Decision-based evaluation. In *Creating, implementing and managing effective training and development. State-of-the-art lessons for practice*. (Society for Industrial and Organizational Psychology ed., pp. 331-375): Jossey-Bass.
- Kraiger, K. (2002). *Creating, Implementing and Managing Effective Training and Development. State-of-the-Art Lessons for Practice*. San Francisco: Jossey-Bass.
- Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive skill-based and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology*, 78(2), 311-328.
- Krathwohl, D.R., Bloom, B.S., & Maisa, B.B. (1964). *Taxonomy of educational objectives: the classification of educational goals*. New York: Longman.
- Lam, A. (2002). *Alternative societal models of learning and innovation in the knowledge economy*. Paper presented at the Conference Name. Retrieved Access Date. from URL.
- Lapierre, G., Méthot, N., & Tétreault, H. (2005). *L'évaluation de la loi favorisant le développement de la formation de la main d'œuvre. Rapport de la deuxième enquête auprès des employeurs assujettis à la loi*. Québec : Direction de l'évaluation. Direction générale adjointe de recherche, de l'évaluation et de la statistique. Ministère de l'Emploi, de la Solidarité sociale et de la Famille.
- Latham, G. P. (1988). Human resource training and development. *Annual Review of Psychology*, 39, 545-582.
- LeBoterf, G. (2002). *Ingénierie et évaluation des compétences* (Éditions d'Organisation ed.). Paris, France.
- Leckie, N., Léonard, A., Turcotte, J., & Wallace, D. (2001). *Pratiques de ressources humaines : perspectives des employeurs et des employés*. Ottawa: Statistiques Canada et Développement des ressources humaines Canada.
- Lee, S. H., & Pershing, J. A. (1999). Effective reaction evaluation in evaluating training programs. *Performance Improvement*, 38(8), 32-39.
- Leverentz, A. M. (1999, August). *One hundred percent satisfaction: how a program fails while pleasing everyone*. Paper presented at the Paper presented at the American Sociological Association Annual Meetings, Chicago, IL.
- Leviton, L. C. (2003). Evaluation use: advances, challenges and applications. *American Journal of Evaluation*, 24(4), 525-535.
- Lévy-Leboyer, C. (2000). *Le 360° : outil de développement personnel* (Éditions d'Organisation ed.). Paris, France.
- Lewis, P., & Thornhill, A. (1994). The evaluation of training; an organizational culture approach. *Journal of European Industrial Training*, 18(8), 25-32.
- Longenecker, C. O., Simonetti, J. L., & LaHote, D. (1998). Increasing the ROI on management education efforts. *Career Development International*, 3(4), 154-160.
- Longenecker, C. O., Simonetti, J. L., & LaHote, D. (1998). Increasing ROI on management education efforts. *Career Development*, 3(4), 154.
- Malglaive, G. (1998). La définition des objectifs pédagogiques. In *Enseigner à des adultes* (pp. 124-130). Paris, France : Presses Universitaires de France.

- Mark, M. M., Henry, G. T., & Julnes, G. (2000). *Evaluation. An integrated framework for understanding, guiding and improving policies and programs*. San Francisco: Jossey-Bass.
- Martocchio, J. J., & Hertenstein, E. J. (2003). Learning orientation and goal orientation context: relationships with cognitive and affective learning outcomes. *Human Resource Development Quarterly*, 14(4), 413-434.
- Mathieu, J. E., Martineau, J. W., & Tannenbaum, S. I. (1993). Individual and situational influences on the development of self efficacy: implications for training effectiveness. *Personnel Psychology*, 46(1), 125.
- McGee, P. The productivity model: a conceptual framework. Retrieved 04-07-06, from <http://www.internettraining.com/Productivity/Productivity.htm>
- McIntyre, D. (1994). *Training and development 1993: Policies, Practices and Expenditures*: The Conference Board of Canada.
- Meignant, A. (2003). Évaluer et utiliser les résultats. In *Manager la formation* (Les Éditions Liaisons ed., pp. 341-374). Paris, France: Éditions Liaisons.
- Mezirow, J. (1997). Transformative Learning: Theory to Practice. *New Directions for Adult and Continuing Education*, (74), 5-12.
- Molenda, M., Pershing, J., & Reigeluth, C. (1996). Designing instructional systems. In R. Craig (Ed.), *The ASTD Training and development handbook: a guide to human resource development (4th edition)*. New York: McGraw Hill.
- Morgan, R. B., & Casper, W. J. (2000). Examining the factor structure of participant reactions to training: a multidimensional approach *Human Resource Development Quarterly*, 11(3), 301-317.
- Moy, J., & McDonald, R. (2000). *Analysing enterprise returns on training*: Leabrook.
- Mulder, M. (2001). Customer satisfaction with training programs. *Journal of European Industrial Training*, 25(6), 321-331.
- Nadeau, M. A. (1988). *L'évaluation de programme. Théorie et pratique*. Québec: Les presses de l'Université Laval.
- Paradas, A. (1993a). *Contribution à l'évaluation de la formation professionnelle en PME*. Unpublished Thèse, Université de Montpellier I-II, Montpellier.
- Paradas, A. (1993b). *Les conditions de succès d'une politique de formation en PME*. Paper presented at the Congrès international francophone de la PME : compétitivité des PME et marchés sans frontières.
- Parry, S. B. (1996). Measuring training's ROI. *Training & Development*, 50(5), 72-77.
- Patton, M. Q. (1997). *Utilization-focused evaluation: the new century text* (SAGE Publications ed.): Edition 3.
- Pershing, J. A., & Pershing, J. L. (2001). Ineffective reaction evaluation *Human Resource Development Quarterly*, 12(1), 73-90.
- Phillips, J. J. (1996). ROI : The search for best practices. *Training & Development*, 50(2), 42-47.
- Phillips, J. J. (1996). Was it the training? *Training & Development*, 50(3), 28-32.
- Phillips, J. J. (1997). *Return on investment in training and performance improvement programs. A step-by-step manual for calculating the financial return* (Gulf Publishing ed.). Houston, TX: Gulf Publishing.
- Phillips, J. J. (1997). *Handbook of training evaluation and measurement methods. Third edition*.
- Pine, J., & Tingley, J. C. (1993). ROI of soft-skills training. *Training*, 30(2), 55.
- Plant, R. A., & Ryan, R. J. (1994). Who is evaluating training? *Journal of European Industrial Training*, 18(5), 27-30.
- Pulley, M. L. (1994). Navigating the evaluation rapids *Training & Development*, 48(9), 19.
- Redshaw, B. (2000). Evaluating organisational effectiveness. *Industrial and Commercial Training*, 32(7), 245-248.
- Reitz, P. (1997). L'évaluation de la formation en entreprise selon le modèle de Donald L. Kirkpatrick: un regard critique. *The Canadian Journal of Program Evaluation*, 12(2), 1-19.
- Richman-Hirsch, W. L. (2001). Posttraining Interventions to Enhance Transfer: The Moderating Effects of Work Environments. *Human Resource Development Quarterly*, 12(2), 105-120.
- Russ-Eft, D., & Preskill, H. (2001). *Evaluation in organizations: a systematic approach to enhancing learning performance and change*. Cambridge, MA: Basic Books.
- Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: a decade of progress. *Annual Review of Psychology*, 52, 471-499.
- Sarnin, P. (1998). La formation professionnelle continue dans les PME européennes: enjeux et problématiques à travers quatre études de cas. *Revue internationale PME*, 11(2/3), 127-160.
- Schon, D. A. (1994). *Le praticien réflexif. À la recherche du savoir caché dans l'agir professionnel*: Les éditions logiques.
- Scriven, M. (1993). *Hard-won lessons in program evaluation* (American Evaluation Association ed. Vol. 58). San Francisco, CA: Jossey-Bass Publishers
- Scriven, M. (2005). Key Evaluation Checklist (KEC). from <http://www.mich.edu/evalctr/checklists/>
- Smith, A. (2001). *Return on investment in training: research readings* (Australian National Training Authority ed.). Leabrook, SA, Australia: National Center for Vocational Education Research Ltd.
- Smith, M. E. (1996). Measuring results. In ASTD (Ed.), *Program design and development* (pp. 320-336). New York: The ASTD training and development handbook. 4th ed.
- Spitzer, D. R. (1999). Embracing evaluation. *Training*, 36(6).
- Spitzer, D. R. (1999). Embracing evaluation. *Training*, 36(6), 42.
- Sridharan, S. (2003). Introduction to Special Section on "What is a useful evaluation?" *American Journal of Evaluation*, 24(4), 483-487.
- Stufflebeam, D.L. (2001). Evaluation models. *New directions for evaluation*, 89, 7-99.
- Stufflebeam, D. L., Foley, W. J., Gephart, W. J., Guba, E. J., Hammond, R. L., Merriman, H. O., et al. (1971). *Educational evaluation and decision making*. Itasca: Peacock.
- Sugrue, B., & Kim, K. H. (2004). *2004 state of the industry report*: American Society for Training and Development.
- Sugrue, B. & Rivera, (2005). *2005 state of the industry report*: American Society for Training and Development.
- Sulsky, L. M., & Keown, J. L. (1998). Performance appraisal in the changing world of work: implications for the meaning and measurement of work performance. *Canadian Psychology*, 39(1-2), 52-59.
- Tamkin, P., Yarnall, J., & Kerrin, M. (2002). *Kirkpatrick and Beyond: A review of models of training evaluation*. Brighton, UK: The Institute for Employment Studies.
- Tan, J. A., Hall, R. J., & Boyce, C. (2003). The role of employee reactions in predicting training effectiveness *Human Resource Development Quarterly*, 14(4), 397-411.

- Tannenbaum, S. I., Mathieu, J. E., Salas, E., & Cannon-Bowers, J. A. (1991). Meeting trainees' expectations: the influence of training fulfillment on the development of commitment self-efficacy and motivation. *Journal of Applied Psychology, 76*(6), 759-769.
- Tannenbaum, S. I., & Yukl, G. (1992). Training and development in work organizations. *Annual Review of Psychology, 43*.
- Torres, R. T., & Preskill, H. (2001). Evaluating and organizational learning: past, present and future. *American Journal of Evaluation, 22*(3), 387-395.
- Tracey, J. B., Hinkin, T. R., Tannenbaum, S., & Mathieu, J. E. (2001). The influence of individual characteristics and the work environment on varying levels of training outcomes *Human Resource Development Quarterly, 12*(1), 5-23.
- Twitchell, S., Holton, E. F., & Trott, J. W. (2000). Technical Training Evaluation Practices in the United States. *Performance Improvement Quarterly, 13*(3), 84-109.
- Wang, G. G., Dou, Z., & Li, N. (2002). A system approach to measuring return on investment for HRD interventions *Human Resource Development Quarterly, 13*(2), 203-224.
- Warr, P., Allan, C., & Birdi, K. (1999). Predicting three levels of training outcomes. *Journal of Occupational and Organizational Psychology, 72*, 351-375.
- Watson, S. C. (1998). Five easy pieces to performance measurement. *Training & Development, 52*(5), 44-48.
- Williams, L. A. (1996). Measurement made simple. *Training & Development, 50*(7), 43-45.
- Worthen, B. R. (2001). Wither evaluation? That all depends. *American Journal of Evaluation, 22*(3), 409-418.
- Worthen, B. (2001). Measuring the ROI of training. *CIO, 14*(9), 128.
- Yamnil, S., & McLean, G. N. (2001). Theories supporting transfer of training *Human Resource Development Quarterly, 12*(2), 195-207.
- Zarifian, P. (1999). La compétence : définition, enjeux et difficultés. In *Objectif compétence*. Paris, France : Éditions Liaisons.

