

La place de l'évaluation dans la pratique de l'ergonome

François Daniellou

Nous sommes, dans la pratique de l'ergonomie, de plus en plus confrontés à cette question de l'évaluation, à de nombreux titres : évaluation de la performance, des coûts cachés, des atteintes à la santé. Mais le thème que je voudrais introduire aujourd'hui est celui de l'évaluation de l'intervention ergonomique.

1. L'évaluation, pour quoi faire ?

Une des définitions de l'évaluation de l'intervention est « porter un jugement de valeur sur une intervention ou sur n'importe laquelle de ses composantes dans le but d'aider à la prise de décision. » (Contandriopoulos et coll., 1992)¹.

Cette idée d'aide à la prise de décision me paraît importante : la question d'une évaluation ne se pose qu'à la lumière du *type de décision qu'elle vise à éclairer*. Autrement dit, si l'on ne sait pas quel type de décision on veut éclairer, la question de l'évaluation ne se pose pas. Donc, chaque fois qu'on nous parle d'évaluation, la première interrogation qui surgit est : quelle est la question ?

Beaucoup de questions sont intéressantes, légitimes, mais ce ne sont pas les mêmes : le ministre du travail peut se demander, dans l'espoir de réduire les TMS de 20 %, s'il faut subventionner largement les interventions ergonomiques (il s'agit d'une question fictive mais qui pourrait se poser). Le contrôleur de gestion de l'entreprise peut se demander si l'intervention coûte moins cher que les risques et dysfonctionnements qu'elle vise à éviter. Les chefs d'établissement vont plutôt se poser la question de savoir si la prochaine fois ils continueront à faire appel à un ergonome. Le CHSCT va se demander si l'évaluation de cette intervention incite à revendiquer la présence d'un ergonome dans les futurs projets. L'ergonome aussi va bien sûr se poser des questions : qu'est-ce j'ai appris ? Comment ma pratique va-t-elle évoluer à la lumière de cette intervention ? Et pour ceux qui sont chargés de formaliser et de transmettre le métier, la question qui va se poser est de savoir quelles sont les leçons relativement généralisables que l'on peut tirer de cette intervention. Toutes ces questions sont légitimes, mais ne donneront pas lieu au même regard sur la question de l'évaluation.

Par ailleurs, nous avons été échaudés à de nombreuses reprises sur les questions d'évaluation, et donc nous avons tendance à nous raidir. La question de l'évaluation est une vieille histoire de chercheurs fondateurs de l'ergonomie, car, dès 1934, Laugier – qui était responsable du laboratoire qui deviendra le laboratoire d'ergonomie du CNAM – et Piéron ont commencé à travailler sur la notation scolaire, ce qui a donné naissance à une discipline qui s'appelle la

¹ Contandriopoulos A.P., Champagne F., Denis J.L., Pineault R. (1992). *L'évaluation dans le domaine de la santé*. Montréal : Groupe de recherche interdisciplinaire en santé.

docimologie. Ils ont montré en travaillant sur les notes du bac que le degré de corrélation entre deux correcteurs était faible, de l'ordre de 0,58. L'évaluation faite par les enseignants manquait totalement de fiabilité. Depuis, un certain nombre de choses ont été faites, même si elles n'ont pas résolu tous les problèmes. Ces recherches ont conduit à faire passer progressivement le projecteur de *l'évaluation des résultats* à celle des *processus*, pour comprendre par quels mécanismes on arrive à telle ou telle différenciation des notes.

Nous avons aussi appris à nous méfier de certaines formes d'évaluation en entreprise, qui soulèvent beaucoup de questions, notamment l'évaluation des savoir-être, qui fait tant souffrir certains travailleurs. Il y a actuellement, vous le savez, une grosse inquiétude sur la mise en place d'une évaluation des risques de délinquance chez les enfants de 3 ans. Certain d'entre vous se souviennent peut-être aussi du programme « Technologie – Travail - Emploi », un remarquable programme du ministère de la recherche dans les années 1980, qui permettait de financer des recherches dans le domaine du travail comme jamais auparavant. Un jour, un ministre a voulu tuer ce programme qui fonctionnait très bien, et ce fut assez simple : il a suffi de nommer un expert évaluateur européen. Plusieurs d'entre nous, dès que nous avons connu le nom de l'expert, savions que le programme était mort. Si vous voulez tuer un programme, il suffit de choisir le bon évaluateur et vous êtes sûr du résultat.

Pour autant, nous aussi sommes, à certains moments, demandeurs d'évaluation. Notamment quand nous voyons le déficit d'évaluation de la conception réalisée par des cabinets d'architectes, ou par les concepteurs en général. Dans certain pays, il y a une obligation légale que l'architecte revienne 6 mois après l'ouverture pour évaluer l'usage des installations. En France, cette tradition de l'évaluation d'usage n'existe guère. Vous vous souvenez peut être du principal d'un lycée de banlieue parisienne qui avait dit : « A ma connaissance, l'architecte n'a pas été condamné à revenir sur les lieux ».

Nous sommes aussi confrontés à un phénomène paradoxal. Nous savons que chacune de nos interventions a des résultats « plutôt positifs » – bien sûr –, mais quand nous intervenons après un collègue dans une même entreprise, le bilan de la situation, telle que nous la racontent les acteurs de l'entreprise, est souvent plutôt négatif. C'est le même phénomène auquel les médecins généralistes sont confrontés. Comme les gens ne changent de médecin que lorsqu'ils ne sont pas contents, les médecins généralistes voient passer beaucoup de patients qui ne sont pas contents de leur médecin précédent. Ils ont donc souvent une opinion déplorable de leurs confrères, alors qu'il existe un « léger » effet de biais dans la sélection des patients qu'ils voient passer. Il est probable que des collègues voient aussi après nous des clients mécontents que nous nous ne voyons pas.

2. Equivalent général, ou voix multiples ?

Dès que l'on parle d'évaluation, il y a un risque que cette dernière soit complètement « technocratique », que l'on ramène tout à une sorte d'équivalent général, à l'aune duquel on pourrait tout mesurer, et qui permettrait « d'optimiser » les décisions toutes choses égales par ailleurs. Or, c'est bien le problème, la notion d'*optimisation* ne se pose que s'il y a une certaine commensurabilité, c'est-à-dire si ce que l'on veut prendre en compte peut effectivement se ramener à une unité commune. Mais, si évaluer c'est se référer à de la valeur, il y a différents types de valeurs.

Pour reprendre les termes d'Yves Schwartz, il y a d'un côté des valeurs marchandes, et il est normal que l'ergonomie soit évaluée en termes de valeurs marchandes, et de l'autre, il y a des « valeurs sans dimension », par exemple l'amélioration de la relation entre la maîtrise et les

opérateurs, l'amélioration du fonctionnement d'un collectif, selon des « valeurs de bien commun », (par exemple la solidarité), qui ne peuvent pas être mesurées avec les mêmes outils.

Nous allons en permanence être confrontés à l'opposition entre deux regards sur l'évaluation : l'évaluation à large bande, où il s'agit de mobiliser une série d'acteurs pour débattre de la valeur d'une intervention, par opposition à une évaluation à bande étroite, où quelqu'un se donne la légitimité d'avoir un point de vue surplombant, et de porter un avis d'expert et d'autorité sur l'intervention.

3. Les ingrédients de l'évaluation d'une intervention

Rassemblons maintenant les ingrédients nécessaires à une forme d'évaluation des interventions qui soit utile au métier (fig. 1).

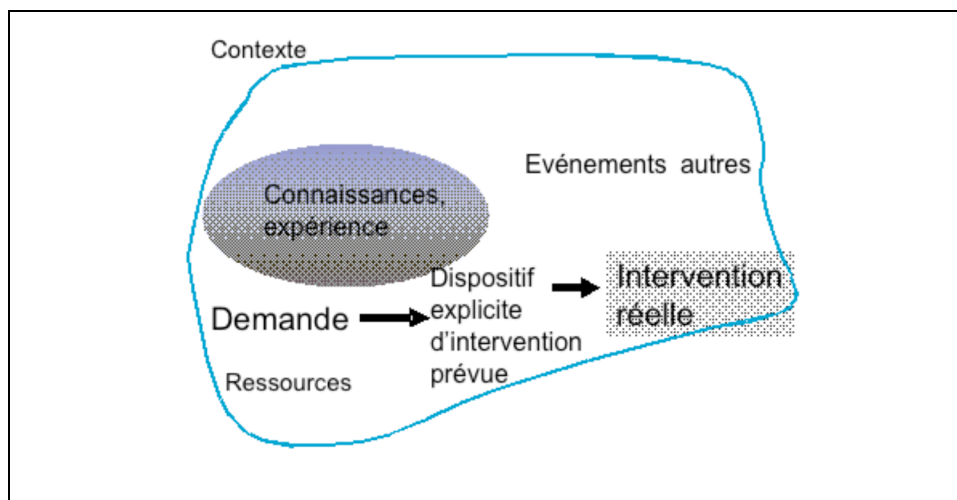


Figure 1

Nous intervenons en réponse à une demande qui se situe dans un contexte. Après avoir évalué les marges de manœuvre, et compte tenu de nos connaissances, de notre expérience, et des modèles de l'intervention dont nous sommes porteurs, nous allons faire une proposition, c'est-à-dire expliciter le dispositif de ce qui va être l'intervention. Ensuite, se déroule l'intervention réelle, qui n'est pas une simple application de la proposition initiale : différents événements vont conduire l'ergonome à ajuster le tir par rapport à ce qui était prévu initialement.

Tout ceci va produire un certain nombre d'effets, de beaucoup de sortes (fig.2). Il y a d'abord les objectifs que l'on poursuivait et qui ont été atteints : le bâtiment devait pouvoir bénéficier du label « haute qualité environnementale », et ce label a été obtenu. Il y a aussi des effets secondaires qui ont pu être anticipés : lorsque nous intervenons dans un atelier pour diminuer les TMS, les TMS augmentent dans les ateliers où nous n'avons pas pu intervenir (personnellement, je préviens le chef d'entreprise de ce phénomène classique). Il y a aussi, malheureusement des choses que nous avons essayé de faire et qui n'ont pas abouti, des résultats qui ne sont pas ceux que nous attendions. Enfin, il y a de nombreuses choses que nous n'avons pas prévues et qui se passent, qui peuvent être positives ou négatives. Le moins, si l'on veut évaluer, est donc de se doter d'une « photographie » de tout cela.

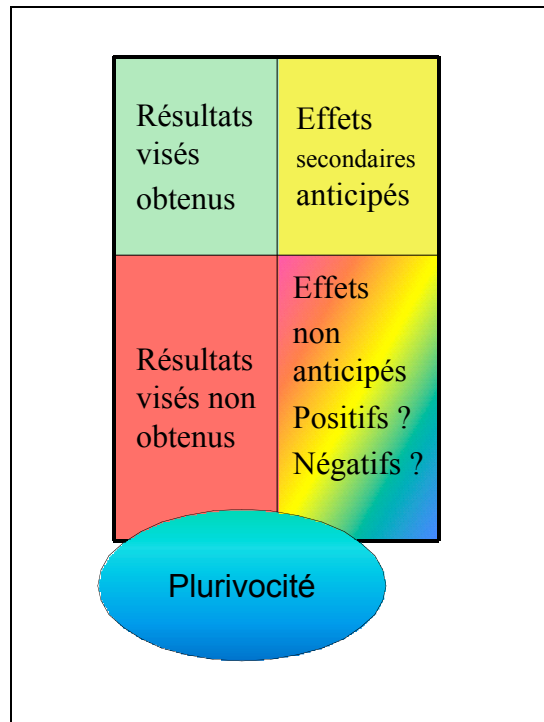


Figure 2 : les divers effets de l'intervention

Comme souligné précédemment, plusieurs acteurs sont concernés par la construction et la lecture de cette photographie. Le mot « plurivocité », équivalent français du terme « *multivoicedness* » utilisé par Engeström, renvoie à cette diversité des voix qui ont quelque chose à dire. Un même résultat peut être passé inaperçu pour certains, être très positif pour d'autres et négatif pour d'autres encore. Une évaluation de l'intervention ergonomique doit écouter ces différentes voix.

Une fois la « photographie » produite, il va se poser la question de savoir si tout ce que l'on constate est lié à l'intervention ou pas, et dans quelle proportion. Bien sûr, pendant que l'ergonome intervenait il s'est passé beaucoup de choses (changement de direction, de gamme de produits, ...). En évaluation, on parle de la question de *l'imputabilité*, c'est-à-dire de la possibilité d'affecter des résultats à des causes. Cette question, me semble-t-il, ne se pose que si l'on est capable de disposer d'une description fine de l'intervention. Si l'intervention reste dans une large mesure décrite comme une boîte noire, il est très difficile d'affecter des résultats à des causes. Pour autant, on ne peut pas décrire intégralement une intervention, de même que l'on ne peut pas décrire intégralement une activité. Ce que l'on va utiliser n'est pas donc une description mais une *modélisation*, c'est-à-dire que l'on va retenir certains caractères de l'intervention, que l'on va mettre en avant comme étant ses composantes principales. Et c'est par rapport à ces composantes de l'intervention que l'on va se poser la question : en quoi les effets que l'on observe sont-ils liés soit à l'intervention elle-même, soit aux variations du contexte, soit aux autres événements qui sont survenus, soit à une combinaison de ces éléments.

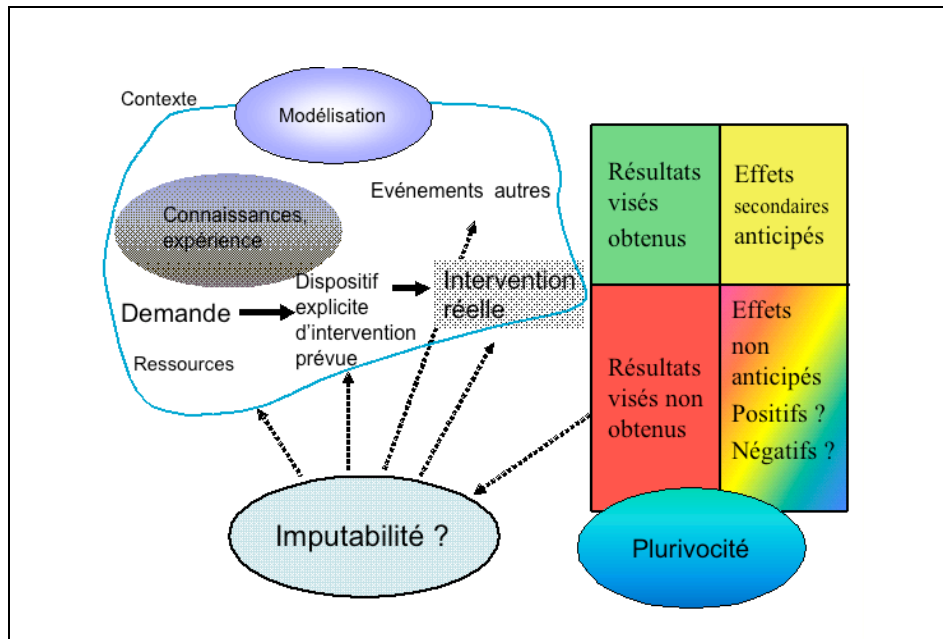


Figure 3 : les ingrédients minimaux de l'évaluation

Il me semble que ce schéma représente l'unité élémentaire d'évaluation de l'intervention. On ne peut pas raisonnablement descendre en dessous de cela, mais c'est déjà beaucoup, et cela ne peut être fait à chaque intervention. Il est possible de choisir de mettre en place une telle évaluation sur certaines interventions, pour essayer de comprendre s'il y a des régularités, des invariants dans le fait que certaines composantes de l'intervention produisent certains effets. Si ces effets sont positifs on va chercher à les généraliser, s'ils sont négatifs on va tenter d'y remédier.

Reprenons quelques-unes de ces composantes de l'intervention qu'il est nécessaire de prendre en compte dans l'évaluation, et arrêtons-nous sur quelques précautions indispensables à chaque étape.

4. Des précautions à toutes les étapes

Comment, à partir de quoi décrire l'intervention que l'on veut évaluer ?

4.1. La proposition d'intervention

Bien évidemment, le dispositif explicite prévu, notre proposition, ne décrit pas l'intervention réelle. Quand des évaluateurs externes évaluent des interventions en n'en prenant comme description que la proposition initiale du consultant, je me permets d'avoir beaucoup de doutes sur ce qu'ils vont évaluer.

La première raison, c'est que seulement une toute petite partie de nos actions est écrite dans la proposition. La proposition n'est pas un protocole expérimental dans lequel on décrirait intégralement tout ce que l'on va faire. La proposition affiche un certain nombre d'étapes de l'intervention, mais il y a beaucoup de choses qui ne sont pas dans la proposition. Une des raisons, que Jean-François Thibault avait développée à un congrès de la SELF, c'est que

lorsque nous écrivons notre proposition d'intervention, nous sommes notre propre prescripteur. Et nous n'avons pas intérêt à nous prescrire un cadre qui serait trop figé.

Mais surtout, l'intervention est une activité, pas seulement un ensemble de tâches. Nous avons incorporé de nombreuses compétences professionnelles que nous serions bien incapables d'explicitier par écrit de façon exhaustive, comme tout travailleur serait incapable de décrire en détail son activité.

De plus, au cours de l'intervention, il va y avoir des ajustements. Soit parce que nous nous rendons compte que ce que nous avons prévu de faire produit des effets qu'il faut gérer en mettant en place des boucles de régulation, soit parce qu'il y a des éléments du contexte qui varient. Un projet ne se déroule jamais de la demande jusqu'au démarrage comme un long fleuve tranquille, il y a toujours des changements – par exemple de financement –, des retards, etc. L'intervention n'est pas le déroulement d'un plan initial. L'action humaine n'est d'ailleurs jamais la mise en œuvre d'un plan initial qui serait ensuite mis en pratique.

De plus, dans une intervention, les objectifs se construisent au fur et à mesure, et quand on a gagné quelque chose avec nos interlocuteurs, on va essayer de gagner l'étape suivante. Une intervention en conception est, par construction, une intervention « à état final non connu ». On ne peut donc pas mettre dans la proposition tout ce qui va s'y passer, car l'état final et le processus pour y arriver ne sont pas complètement connus.

4.2. Comprendre la théorie sous-jacente

On ne peut pas chercher à comparer des interventions si l'on ne sait pas « ce que les intervenants ont en tête », quelles sont leurs hypothèses, sur quels paramètres ils essaient de jouer, quelle est la vision qu'ils ont des mécanismes de l'intervention ergonomique. C'est la question de *la prise en compte de la « théorie sous-jacente »* (Berthelette). Bien sûr nous avons des points communs entre ergonomes, mais il y a des aspects sur lesquels nous ne sommes pas toujours d'accord sur les cibles fondamentales de notre action. Par exemple, le mécanisme central de l'information est-il de fournir à tous les acteurs, à partir de la connaissance de l'activité, des informations qui vont nourrir le débat social ? Ou bien est-il de se positionner aux bons aiguillages dans les processus de décision pour faire passer le train de tel ou tel côté ? Il ne s'agit pas du tout de la même théorie sous-jacente. Non pas qu'il en existe une bonne et une mauvaise, mais lorsque l'on construit son intervention avec l'une en tête, ce n'est pas pareil que si l'on avait l'autre.

Pour accéder à cette « théorie sous-jacente », il est nécessaire de faire expliciter par les intervenants les objectifs qu'ils visent en début d'intervention, les leviers sur lesquels ils entendent jouer, etc.

4.3. Les réajustements de l'intervention

Il faut aussi comprendre la négociation de la proposition, comment les objectifs de l'intervenant se sont heurtés à des questions de ressources, de compréhension, de rapports de pouvoir.

On a besoin de décrire les actes et les opérations qui permettent la construction de l'intervention, mais une partie de ces actes d'intervention sont des actes publics – qui font l'objet de comptes rendus (intervention au comité de pilotage, au CHSCT, analyse de l'activité) – tandis que d'autres sont privés, (une discussion à la machine à café ou dans le bureau du directeur), pour lesquels il n'y a aucune trace. Or ces derniers sont probablement

tout aussi déterminants dans le déroulement de l'intervention que l'ensemble des autres facteurs.

Nous passons une partie de notre temps en réflexion sur l'action, et nous réajustons nos méthodologies en cours d'intervention. Si l'on veut évaluer une intervention, il est extrêmement intéressant d'avoir des traces de ces moments où les ergonomes ont délibéré, se sont dit : « compte tenu des évolutions du contexte on va changer quelque chose dans l'intervention » (Jackson).

4.4. Composantes formelles et composantes réelles de l'intervention

On l'a dit précédemment, on ne peut pas réfléchir sur l'évaluation d'une intervention en restreignant la description de celle-ci à quelques composantes formelles qui seraient supposées en rendre compte. Par exemple, il existe des travaux sur : « est-ce que les interventions participatives sont plus efficaces pour la prévention des TMS que les interventions non participatives ? » Je ne suis pas sûr que, posée de cette façon, la question ait beaucoup de sens, parce que pour « faire participer » des opérateurs, l'ergonome a travaillé des jours pour construire les conditions sociales et techniques de cette participation d'une façon qui puisse la rendre pertinente. Ce n'est pas en soi parce qu'on a fait participer des opérateurs que cela signifie quoi que ce soit.

Il y a beaucoup de questions comme celle là : « est-ce qu'une intervention est plus pertinente quand il y a un comité de pilotage ou pas ? » Comme si l'existence d'un comité de pilotage était indépendante de tous les colloques singuliers, de tout le temps que l'ergonome a passé en privé avec telle ou telle personne du comité de pilotage pour la préparer, de la dynamique sociale dans l'entreprise, etc. On peut dire la même chose des groupes de travail, des formations-actions, etc.

Si l'on ne rentre pas dans le contenu de chacune de ces « boîtes noires », je trouve que la question de l'évaluation n'a simplement pas de sens, et qu'il ne faut pas accepter de la discuter. Par contre si l'on est entre professionnels, et que l'on veuille répondre à ces questions, alors il faut ouvrir les boîtes noires, et du coup la discussion devient possible – c'est ce que nous essayons de faire ici. C'est la préoccupation qui, dans les travaux sur l'évaluation, notamment ceux de Diane Berthelette, est décrite sous le terme : « *décrire la stratégie d'implantation* ». Dire qu'il y a telle ou telle composante n'a pas de sens, si l'on ne dit pas ce qu'il a fallu faire pour qu'il y ait un comité de pilotage, un groupe de travail, etc.

4.5. Expérimentation et expérience vécue

Si on veut se poser la question d'évaluer des interventions, il faut prendre en compte le fait que toute intervention a deux dimensions² : d'une part, une intervention est dans une certaine mesure une *expérimentation contrôlée*, c'est-à-dire que nous avons des connaissances, des modèles, des méthodes, qui nous permettent de mettre en place quelque chose qui pour partie ressemble à un protocole (planning, nombre de jours, ...). D'autre part, c'est une expérience, au sens de l'expérience humaine, de *l'expérience vécue*. C'est-à-dire que l'on va rencontrer des personnes, qu'elles vont être plus ou moins vite convaincues, que les éléments d'information qu'on leur apporte sur le travail vont modifier leur façon de réfléchir,

² Schwartz (1994[2000]), Conférence inaugurale du 29^e congrès de la SELF, pp. 537-548, in Y. Schwartz, *Le paradigme ergologique ou un métier de philosophe*, Toulouse : Octarès Editions.

d'interagir entre elles. Il va donc y avoir des effets de l'intervention sur les personnes, sur les relations entre les personnes, qui ne peuvent pas du tout faire l'objet d'un protocole.

Pour la première partie, ce que l'on peut mettre sous forme de protocole, la question de l'évaluation est assez simple. On sait ce que l'on va faire, on peut donc mettre en place un certain nombre de capteurs pour voir ce que cela va donner. Une petite partie des objectifs initiaux du projet est soit quantifiable, soit « critériable », pour utiliser le jargon de la docimologie. Par exemple, un objectif d'amélioration du rendement matière est quantifiable. On peut dire qu'un objectif est « critériable » lorsqu'il y a des critères qui doivent être remplis, par exemple « obtenir le label HQE ». L'atteinte de tels objectifs est parfois facile à évaluer.

Mais pour la deuxième dimension, les effets que l'on produit sur les personnes, sur les collectifs, sur les relations entre les personnes, l'évaluation est plus difficile. Ce qui ne veut pas dire que l'on ne doit pas essayer, mais l'on est sûr d'avance qu'il est impossible d'être exhaustif, d'écrire tous les effets que l'on a produits sur telle ou telle personne. Si nous nous fixons par exemple un objectif d'amélioration des relations entre la maîtrise et les opérateurs, faut-il vraiment mettre en place un protocole lourd pour évaluer le résultat ?

4.6. La difficulté d'évaluer certains effets

Et puis il y a des domaines où la mesure de ce que l'on veut évaluer paraît simple, mais où dans les faits l'évaluation est extrêmement difficile. Un exemple est l'objectif d'une diminution des TMS. Il existe des moyens de mesurer les TMS, mais ce que l'on sait maintenant, notamment grâce aux travaux d'Agnès Aublet-Cuvelier et d'Yves Roquelaure, c'est qu'il s'agit d'une pathologie à la fois cyclique et évolutive, c'est-à-dire qu'elle varie en intensité, et peut se déplacer sur les différentes localisations du membre supérieur. Suivant les outils que l'on emploie, le nombre de TMS que l'on comptabilise avant/après, et les effets évalués de l'intervention seront complètement différents.

Par exemple, si vous évaluez les TMS avec une question du type : « avez-vous eu mal aux doigts au cours des 6 derniers mois ? », quelle que soit l'intervention, il sera extrêmement difficile de faire baisser la proportion de gens qui répondront positivement à cette question, du moins dans des secteurs où il y a des risques de TMS. Il y a toujours un bruit de fond que l'on n'arrive pas à supprimer. Par contre, si vous posez la question, « avez-vous eu mal au cours des 7 derniers jours ? », si l'intervention s'est bien passée, on pourra mettre en évidence des effets très significatifs sur le nombre de gens qui ont eu des douleurs pendant cette période. Même sur quelque chose qui paraît aussi quantifiable que des pathologies, on est confronté à des questions théoriques et méthodologiques majeures. Quelqu'un qui ne se serait pas penché sur cette difficulté de l'évaluation des TMS peut mettre en place des mesures qui produiront des résultats au mieux ininterprétables et au pire pernicieux.

4.7. Mettre l'intervention sous contrôle ?

On pourrait souhaiter, pour les besoins de l'évaluation, mettre en place en début d'intervention de nombreux indicateurs, qui permettraient de mettre sous observation l'intervention, de voir comment ça se passe, et d'en tirer de leçons. Le problème est que le travail à fournir pour définir les indicateurs pertinents est à peu près le travail de toute l'intervention. C'est en fin d'intervention que l'on sait quels indicateurs il aurait fallu mettre en place pour avoir une idée précise de tel ou tel effet que l'on n'avait pas anticipé. C'est un résultat de l'intervention que d'aider les acteurs de l'entreprise à se doter de nouveaux indicateurs, mais il aura fallu tout ce temps, tout ce parcours avec eux pour qu'ils soient en

situation de se dire que cela en vaut la peine. En début d'intervention, on n'aurait eu aucune chance de l'obtenir.

Par contre, nous apprenons de chaque intervention et nous partageons entre nous, ce qui permet d'améliorer, d'une fois sur l'autre, le type d'indicateurs que l'on va mettre en place en début d'intervention, sachant qu'il y a bien sûr des différences entre les situations et que tout n'est pas transposable.

4.8. Les indicateurs : une focalisation excessive de la compréhension ?

Une autre difficulté est que la mise en place d'un élément de mesure peut conduire les décideurs à limiter le phénomène à ce que l'instrument de mesure prend en compte. Il s'agit d'un phénomène classique. En matière d'accidents du travail par exemple, vous connaissez les effets d'instruments de mesure du genre « jours sans arrêt », avec des concours sur le nombre de jours sans arrêt entre les établissements d'un même groupe, qui conduisent souvent à ce que les gens ayant subi un accident soient maintenus à ne rien faire dans l'entreprise plutôt qu'à déclarer l'accident.

Les dernières statistiques de la CRAM font état de 485 décès résultant de maladies professionnelles, alors que l'Institut de veille sanitaire estime pour les seuls cancers professionnels le nombre de décès à 7500 par an. Si l'on considère que ce qui est important, c'est ce que l'outil de la CRAM mesure, il y a 485 décès par maladies professionnelles en France, donc pas un problème majeur. Si l'on considère que l'outil intéressant est celui de l'InVS, on va passer à un niveau de priorité supérieur à celui des accidents de la route et très supérieur à celui des homicides.

4.9. Les limites de la comparaison avant/après

Il faut aussi s'arrêter sur l'illusion du avant/après. C'est une question complexe : qu'est-ce que « l'avant-intervention » ? Le mieux que nous puissions faire, c'est une photographie de la situation quand nous arrivons. Mais quand nous sommes déjà arrivés, ce n'est pas avant l'intervention : des acteurs se sont battus pour obtenir notre présence, il y a eu des négociations et le fait qu'ils aient gagné une intervention ergonomique, c'est déjà un soulagement, un mieux. Probablement, ils vont déjà mieux que ce que nous aurions constaté si, par une quelconque opération de magie, nous avions pu faire la photographie 6 mois plus tôt. Il y a des situations où, par chance, d'excellents médecins du travail ont pu faire cette photographie. Mais c'est un cas assez rare.

Le « après » est, lui, à décliner de façon précise. Il y a l'évaluation au moment du démarrage. Des ergonomes ont contribué à la conception de nouvelles installations, tout le monde est très content de ce qui c'est passé durant la conception, mais quoi qu'ils aient fait, le démarrage reste un moment difficile. Les opérateurs prennent en main les nouvelles installations, doivent acquérir un certain nombre de compétences gestuelles par exemple. Ils ont eu une formation, mais ce n'est pas la même chose de l'appliquer sous cadence. Même dans un projet extrêmement bien réussi, il est possible qu'il y ait un jour, une semaine, un mois difficile pour les opérateurs. Si on ne fait que l'évaluation du démarrage, on va avoir une vision très pessimiste des effets de l'intervention, parce que les opérateurs font état de douleurs, par exemple.

Si on revient quelques mois après, les effets de l'intervention sont encore visibles, mais l'appropriation est faite, les opérateurs ont habité leur installation, l'ont mise à leur main.

Probablement, on va alors voir des résultats plus significatifs des effets globaux de l'intervention – ce qui n'empêche pas que les opérateurs ont eu mal pendant le démarrage.

Puis, si l'on revient 12 ou 18 mois après, on arrive dans quelque chose qui n'a plus rien à voir avec le contexte dans lequel on est intervenu : les directeurs, les actionnaires, les produits, le nom de l'entreprise peuvent avoir changé. On constatera peut-être de nombreux problèmes, mais ils n'ont plus grand-chose à voir avec l'intervention. Cela montre seulement notre incapacité à assurer des résultats pérennes dans un contexte mouvant.

La comparaison avant/après, de mon point de vue, n'a pas de sens si elle ne porte pas sur une série de plusieurs points, et doit comporter beaucoup de précautions pour ce qui est de « l'avant ».

4.10. Des effets ailleurs

Mais il y a aussi des effets « ailleurs », liés notamment au turn-over de l'encadrement. Un des effets négatifs, c'est que le chef de projet s'en aille au bout de 3 ans et que tout ce qui avait été construit avec lui disparaisse du site sur lequel on est intervenu. Mais il peut aussi y avoir éventuellement un effet positif, c'est qu'il emmène avec lui un intérêt pour l'ergonomie qui va produire des effets ailleurs. Il serait, de mon point de vue, intéressant de prendre en compte ces effets dans l'évaluation, mais on est la plupart du temps bien incapable de mesurer le périmètre et la durée de ce phénomène.

4.11. Prédire et convaincre

Il faut aussi avoir en tête qu'il y a deux choses qui n'ont rien à voir : notre capacité, notamment à travers les simulations, à *prédire* tel ou tel incident, dysfonctionnement ou risque, et notre capacité à *convaincre* les acteurs de l'entreprise. On peut très bien avoir une simulation qui montre qu'il y a un risque d'incident grave et que le chef de projet soit convaincu que ce problème sera réglé et ne se posera pas.

Évaluer l'intervention, ce n'est pas seulement évaluer le fait que l'on ait eu raison d'alerter, – car malheureusement ce n'est pas très utile d'avoir raison dans ce cas de figure –, mais c'est aussi évaluer dans quelle mesure on a su faire partager notre pronostic sur l'activité future, et convaincre nos interlocuteurs.

4.12. Évaluer ce que ça demande

Il faudrait rajouter que pour évaluer l'intervention à des fins de transmission du métier, il faudrait décrire non seulement ce que l'on a fait, mais aussi ce que ça nous a coûté en tant que travailleurs. Quels sont les moments où nous étions en difficulté, où nous avons fait des cauchemars, les moments où nous hésitions ? Là, on commencerait à avoir une vision de l'intervention vraiment utile pour la formation des nouveaux professionnels. Il n'est pas possible de faire ça à chaque fois, mais je crois beaucoup à l'intérêt de le faire une fois de temps en temps. Pour une partie des interventions, soit des intervenants praticiens, soit des binômes entre des praticiens et des chercheurs, peuvent faire des monographies détaillées sur leur déroulement.

Ceci suppose de garder des traces des interventions, car bien sûr notre souvenir n'a aucune valeur de vérité. Marçal Jackson l'a montré dans sa thèse³, le récit *a posteriori* d'une intervention n'a pas grand-chose à voir avec ce qui s'est vraiment passé : pas le bon ordre, pas les bonnes personnes, pas la bonne raison, etc. On reconstitue *a posteriori* un récit qui a une certaine cohérence mais qui comporte des raccourcis considérables. La mémoire est une reconstruction et pas une trace fidèle.

On ne peut donc pas évaluer le contenu réel d'une intervention à partir de ce que les gens disent *a posteriori*. Ce qu'ils disent *a posteriori* est intéressant, parce que c'est un ressenti, qui fournit des informations sur comment ils ont vécu l'histoire, mais ça ne fournit pas d'information sur comment l'histoire s'est réellement déroulée. Si on veut reconstituer les épisodes d'une intervention, il faut partir de traces d'époque : les agendas, les courriers électroniques, les comptes rendus de réunions, ... Dans certains cas, il est intéressant, quand on veut mettre une intervention sous observation, d'essayer de travailler avec des *journaux de bord*. Nos collègues québécoises ont beaucoup développé cette idée, et Aurélie Landry est en train de le faire sur des interventions sur les TMS ici. C'est extrêmement lourd, mais à la fin les intervenants eux-mêmes en sont contents, parce que c'est un outil de communication entre les différents membres de l'équipe d'intervention. C'est aussi un outil très utile ensuite pour prendre du recul.

A partir de ces traces, on peut enfin essayer d'écrire *un récit de l'intervention*. On met en lien des éléments factuels, dans un récit qui essaie d'établir des relations de causes à effets. Ces liens sont évidemment hypothétiques. Ils peuvent être plus ou moins corroborés par l'opinion d'un plus ou moins grand nombre d'acteurs. C'est ce récit qui peut servir de base à l'évaluation.

5. L'évaluation est souvent nuancée

Quand on fait une évaluation, il faut s'attendre à ce qu'elle soit très nuancée. Une évaluation de démarrage est presque toujours lourdement négative, et une évaluation quelques semaines après est souvent fortement nuancée. Il y a des choses qui vont bien, des choses pour lesquelles les opérateurs sont vraiment contents, qui ont vraiment amélioré leur situation et puis il y a des choses qui ne vont pas bien. On n'a pas eu ni la compétence, ni les ressources pour transformer fondamentalement la situation. Par exemple, on a beaucoup amélioré une chaîne, mais une chaîne reste une chaîne. Les gens ont beaucoup moins mal, mais ils continuent à travailler à la chaîne.

Dans les évaluations, le plus intéressant c'est ce à quoi on ne s'attendait pas. Une fois, des collègues qui sont intervenus sur une maison de retraite m'ont demandé de participer à l'évaluation de cette intervention, à laquelle je n'avais pas contribué. Une ancienne maison de retraite dans un parc était vouée à être démolie et une nouvelle avait été construite à côté, avec l'intervention des ergonomes dans tout le projet. Les nouvelles personnes recrutées dans le nouvel établissement étaient ravies, mais en revanche plusieurs aides-soignantes anciennes ont fait des dépressions nerveuses au moment de la destruction physique de l'ancienne maison de retraite. Nous n'avions pas mesuré que le regard positif qui était porté sur la nouvelle

³ Jackson J.M. (1998). *Entre situations de gestion et situations de délibération : l'action de l'ergonome dans les projets industriels*. Thèse de Doctorat. Éditions du Laboratoire d'Ergonomie des Systèmes Complexe de l'Université Victor Segalen Bordeaux2.

maison de retraite était comme une sorte de mépris pour l'ancienne. Cette maison de retraite, on la détruit, mais c'était celle où elles avaient vécu pendant 15 ans, qu'elles avaient fait fonctionner. Le fait de valoriser quelque chose de nouveau qui va mieux, peut être considéré comme une forme de mépris par rapport à ce qu'il avait fallu faire pour faire fonctionner une installation – certes problématique –, pour que des personnes *puissent y vivre*.

L'enjeu c'est finalement que l'évaluation soit l'occasion de nouveaux développements. C'est-à-dire que pour nous ce soit l'occasion de faire (encore) mieux la prochaine fois, et que pour les acteurs de l'entreprise ce soit l'occasion de se parler de façon nouvelle, de continuer à travailler ensemble, d'avoir des idées nouvelles pour les projets à venir, de laisser des traces dans l'organisation du dialogue et de la négociation de l'entreprise, etc.

Il faut aussi avoir en tête que « l'unité de compte » dans notre relation avec les entreprises ce n'est pas l'intervention. Il y a des ergonomes dans cette salle qui sont intervenants externes depuis 10 ou 15 ans dans la même entreprise. Les relations avec une entreprise, ce sont des feuillets avec beaucoup d'épisodes, qui ont leur lot de succès et d'échecs, mais où il y a des rebonds successifs, et où finalement l'évaluation qu'on peut faire d'une seule intervention dans cet enchaînement là, peut paraître dérisoire.

Telle intervention ne nous a pas satisfaits, mais il s'est passé beaucoup de choses depuis qui ne se seraient pas passées si celle-ci n'avait pas eu lieu. Ou bien nous étions plutôt contents, mais l'entreprise ne nous a jamais recontactés et a fait appel à d'autres conseils depuis. Et ça ne se limite pas à l'entreprise où l'on intervient, ça diffuse dans le groupe, dans la branche, dans la profession, dans les institutions dans le meilleur des cas.

5.1. Réduire et conceptualiser pour enseigner

Ce que l'on produit ce sont des modèles de l'intervention, au sens d'une *réduction finalisée* (on ne conserve que certaines dimensions, avec un objectif donné), et ces modèles ont toutes les propriétés d'un *instrument* au sens de Rabardel⁴. C'est-à-dire d'une part que pour pouvoir les utiliser, chaque ergonome doit développer ses schèmes d'utilisation. Nous avons besoin de développer à l'intérieur de nous-mêmes des compétences pour pouvoir utiliser cet instrument. Et d'autre part, pour pouvoir l'utiliser, nous allons le mettre à notre main : personne *n'appliquera* un schéma d'intervention enseigné à l'université. On va modifier l'outil pour qu'il devienne un instrument.

Les modèles enseignés sont une cristallisation des interventions qui ont été faites dans le passé, ce sont des artefacts, des outils que les ergonomes vont s'approprier pour faire des instruments. Mais il faut que les modèles qui sont enseignés soient remis en permanence sur le métier. Ce travail sur les modèles est constitué d'allers-retours entre le singulier et la généralisation : le singulier n'est pas transmissible et la généralisation n'est jamais strictement applicable dans les situations singulières. C'est pourquoi Yves Schwartz formule autant d'alertes sur les précautions qui doivent être prises pour manipuler tout concept sur le travail : les concepts sont indispensables parce qu'ils évitent d'être paralysé par le caractère infiniment singulier d'une situation ; ils permettent de comprendre, de transmettre, de transformer. En même temps, dit-il, « Toute conceptualisation est aussi une forme d'appropriation, une décision de description de la réalité, une manière d'en instituer les significations et les valeurs, une manière d'arrêter la pensée sur ses contours décrétés pertinents. Parfois, un vrai

⁴ Rabardel P. (1995). *Les hommes et les technologies*. Paris : Armand Colin

"hold-up" sur un milieu de vie »⁵. Les concepts doivent toujours être remis à l'épreuve des situations vécues.

6. En conclusion

La question de l'évaluation de l'intervention ne peut pas être ignorée, puisque, qu'on le veuille ou non, nous y sommes constamment confrontés. Nous savons cependant combien certaines formes d'évaluation sont défavorables lorsqu'elles sont utilisées sans précaution pour juger de la valeur d'interventions ergonomiques. La solution est probablement de revendiquer la spécificité de la *production de connaissances à partir d'interventions*, la dynamique d'une généralisation raisonnablement maîtrisée de situations particulières sur lesquelles on produit des descriptions, des modélisations *et sur lesquelles on agit*. L'intervention est une source de production de connaissances, mais les formes sous lesquelles ces connaissances sont produites ne peuvent être évaluées à la même aune que les recherches expérimentales ou épidémiologiques.

« Le travail c'est quelque chose qui fonctionne comme la mauvaise conscience de la science » écrit C. Dejours⁶. Il dit cela à propos d'un patient. Ce monsieur était chargé de mesurer la résistance du béton qui allait servir à construire une enceinte pour protéger de possibles fuites dans des industries dangereuses. Ce patient est devenu gravement malade. Il savait d'expérience que pour que les résultats sur les éprouvettes qu'il testait soient bons, il fallait que le béton repose pendant 28 jours. Il avait fait des essais, qu'on ne lui demandait pas, pour savoir ce qui se passait au bout de 20 jours, 15 jours, ... et il savait que le béton était dans un état proche de celui qu'il aurait dans le futur au bout de 28 jours. Des réformes de budget du laboratoire où il travaillait ont imposé que les essais soient faits en 5 jours. Le patient de Dejours savait que ce que l'on retirait des essais au bout de 5 jours avait l'allure de la science, avec des capteurs parfaitement étalonnés, mais que le béton n'avait pas été travaillé comme il le fallait, et que donc on était en train d'affirmer avec autorité n'importe quoi sur la façon dont ce béton garantirait la tenue des enceintes devant contenir les fuites de gaz dans une industrie dangereuse. L'appareillage technique de l'évaluation ne garantit pas sa pertinence, si la dynamique de l'évaluation ignore tout le travail nécessaire à produire ce qu'il s'agit d'évaluer.

J'ai toujours cette histoire en tête lorsqu'on nous oppose des formes d'évaluation qui ont toute l'allure de la bonne science, mais qui sont trop simples par rapport à la complexité de toute intervention.

⁵ p. 615, Schwartz, Y., (2000), *Le paradigme ergologique ou un métier de philosophe*, Toulouse : Octarès Editions.

⁶ Dejours C., 2003, *L'évaluation du travail à l'épreuve du réel : critique des fondements de l'évaluation*. Paris : INRA Editions.