

L'outil en ligne.

1. Structure du questionnaire.

La structure, les thèmes et les questions de l'outil reposent sur la roue de Deming, en relation avec les préconisations de l'UE sur la qualité dans la formation professionnelle (<http://www.eqavet.eu>). Une première version de l'outil a été mise au point et testée à partir du « cycle de la qualité » (Quality Circle) issu de la démarche PDCA et d'autres outils existants (SPEAK, Eduqua). Sa structure est présentée dans le schéma ci-dessous. Les différentes couleurs indiquent quels groupes doivent répondre aux questions des différents thèmes (voir légendes).

Légende :

Prestataires, intervenants et apprenants.

Prestataires et intervenants.

Prestataires.

A partir de l'expérience tirée de la phase pilote, des retours des personnes qui ont répondu au questionnaire et d'une première analyse des questions, le questionnaire a fait l'objet de discussions et de modifications minutieuses lors de la réunion de Durham de mars 2012. Le thème « Financement et impacts de la formation » a été supprimé. Par ailleurs, de nombreuses questions ont été reformulées. Il a fallu résoudre des problèmes de langue et de concepts, mais surtout adapter les questions relatives aux thèmes traités par plusieurs groupes de sorte qu'elles correspondent à la situation de tous. Ces modifications répondent à des critiques qui apparaissaient fréquemment dans les retours. Elles apportent par ailleurs une nouvelle possibilité d'évaluation basée sur le feedback à 360 degrés (en l'occurrence, il s'agit plutôt d'un feedback à 270 degrés).

Une fois modifié, l'outil a de nouveau été testé auprès de 120 prestataires, 51 intervenants et 58 apprenants. Ses caractéristiques sont décrites ci-dessous. Dans un premier temps, nous avons analysé les différentes questions dans le but d'évaluer la cohérence interne des thèmes et nous avons mené une analyse factorielle dans le but de contrôler la structure des thèmes (voir schéma ci-dessus). Dans un second temps, nous avons réalisé une synthèse des commentaires des personnes interrogées.

2. Caractéristiques de l'outil.

2.1 Analyse des questions.

L'analyse des questions repose sur la théorie classique des tests telle qu'elle est mise en œuvre dans la procédure « Fiabilité » du programme SPSS. Le tableau ci-dessous présente le coefficient de fiabilité (coefficient alpha de Cronbach et le nombre de questionnaires pris en compte dans le calcul (entre parenthèses) pour chacun des trois groupes interrogés (prestataires, intervenants et apprenants).

La première chose qui frappe est l'importante différence entre le nombre de personnes qui ont participé à l'enquête (côté prestataires par exemple, 120 personnes) et le nombre de questionnaires qui ont pu être pris en compte dans l'évaluation (entre 47 et 68). Cette différence tient notamment au fait que certaines personnes n'ont répondu que très succinctement au questionnaire et que seuls

ont été pris en compte pour l'évaluation les questionnaires qui comportaient au maximum une question sans réponse pour le thème pris en considération. Nous avons attribué à cette question sans réponse une valeur équivalente à la valeur moyenne des autres questions du thème. Par convention, un coefficient de fiabilité ayant une valeur de 0,6 est jugé (juste) satisfaisant ; au-delà de 0,8, il correspond à une évaluation bonne à très bonne. Le tableau montre que les intervenants et les apprenants évaluent les questions comme bonnes à très bonnes. Les prestataires, en revanche, estiment que le thème « Évaluation des connaissances » n'est pas satisfaisant. On observe que sans la question « Des procédures sont prévues pour analyser les difficultés, identifier les axes de progrès et préconiser des mesures d'amélioration au stagiaire ? », le coefficient pourrait atteindre un niveau satisfaisant, avec une valeur de 0,601. Cela étant, il semble tout à fait compréhensible qu'une telle question suscite un tel résultat. Pour le thème « L'Offre de formation répond-elle aux priorités nationales et/ou locales identifiées sur le marché du travail ? », le coefficient est de 0,678. Comme le thème « Adaptation à l'environnement » ne comporte que deux questions, il n'est pas possible d'améliorer son coefficient en en supprimant une. Du reste, le contenu des deux questions semblant important, il semble (tout juste) pertinent de conserver ce thème dans le questionnaire.

Nous n'avons pas réalisé d'analyse factorielle sur l'ensemble des questions ; nous nous sommes limités aux thèmes et aux questions correspondants aux éléments du cycle de la qualité. Ce choix tient avant tout au petit nombre de personnes interrogées. Dans l'ensemble, entre 60 et 80 pour cent de la variance a pu être reproduite par des facteurs en fonction des éléments du cycle et des groupes cibles. Cela étant, au vu des résultats de l'analyse, nous n'avons pas toujours attribué aux questions les facteurs qui semblaient cohérents d'un point de vue théorique. On note certaines différences entre les groupes cibles. Ainsi par exemple, pour l'analyse de l'élément « Evaluation » du point de vue du prestataire, l'analyse factorielle prend en compte quatre facteurs. Les deux premiers (« Titres et qualifications requises pour les formateurs » et « Veille professionnelle et formation des formateurs ») sont reproduits. Le troisième (« Evaluation de l'action de formation ») est scindé en deux : respect des standards de qualité et feedback à l'intervenant. L'évaluation des mêmes questions au vu des réponses des intervenants donne deux facteurs, ce qui tient au fait que le premier thème ne comporte dans ce cas que deux questions.

Les résultats concernant l'élément « Planification » du cycle de la qualité doivent être interprétés légèrement différemment. Pour des raisons avant tout pratiques, nous souhaitons respecter la structure définie sur des bases théoriques. Il sera toujours possible, en dernier recours, de supprimer une question pour ramener le coefficient de fiabilité d'un thème à un niveau satisfaisant. Rappelons ici que l'objectif de l'instrument n'est pas de tester des critères de qualité, mais, dans une optique d'assurance de la qualité, d'identifier des points nécessitant d'être améliorés. En conséquence, il n'est pas nécessaire d'appliquer les critères de qualités aussi rigoureusement que lors des tests.

2.1 Evaluation.

Les commentaires des participants à propos du questionnaire sont très enrichissants pour la poursuite du développement de l'outil. Nous vous présentons ci-dessous une synthèse de leurs retours (rappelons que dans la version test, nous n'avons pas effectué d'évaluation pour toutes les personnes qui ont rempli un questionnaire).

Dans l'ensemble, l'outil a été bien évalué. En revanche, la clarté des questions ainsi que leur pertinence et leur applicabilité le sont moins. Si l'on se penche de plus près sur les réponses – notamment celles des questionnaires dont les évaluations sont régulièrement négatives –, il apparaît clairement que les commentaires négatifs visent avant tout la langue : emploi de jargon et de tournures compliquées, phrases mal formulées. Ces critiques concernent avant tout la version anglaise du questionnaire. Elles sont de prime abord surprenantes, d'une part parce que les différentes versions du questionnaire ont été contrôlées et validées par des collaborateurs originaires des pays correspondants, d'autre part parce que lors des réunions de travail, nous nous sommes justement efforcés d'employer une langue compréhensible et concrète et d'axer les questions sur des faits observables. Une observation plus minutieuse des évaluations montre que ces critiques apparaissent avant tout lors des premiers tests de l'instrument. En revanche, elles ont disparu dans les évaluations ultérieures. Au contraire même, certains questionnaires soulignent la convivialité (facilité d'utilisation) de l'instrument. Du fait de la défaillance de notre partenaire en Irlande, les évaluations ont été réalisées très tardivement par d'autres personnes. Aussi, les problèmes de traduction qui sont apparus au début ont sans doute pu être résolus. Certains aspects temporels entrant en compte, il n'est pas possible d'effectuer une analyse pays par pays.

Les personnes qui ont utilisé l'outil ont également été interrogées sur son intérêt. Les réponses sur ce point sont très variables. Certaines personnes n'ont pas répondu ou n'ont répondu que par oui ou non, la plupart ont donné des réponses détaillées. Nous ne rendons compte ici que de ces dernières réponses circonstanciées. Environ la moitié des personnes interrogées (et non des personnes qui ont répondu !) estiment que l'instrument les a encouragées à s'impliquer davantage dans l'assurance qualité. Et ce, parce qu'il les a amenées à réfléchir à la question de la qualité, à apprécier davantage l'intérêt des feedbacks ou à prendre conscience de la possibilité d'utiliser plusieurs instruments, de comparer leurs résultats et d'obtenir de meilleures conclusions. De même, la majorité des personnes interrogées considèrent que l'instrument est utile à leur structure, qu'il permet d'impliquer activement les collaborateurs dans la réflexion sur la qualité, de mettre l'accent sur les bonnes pratiques, de réfléchir aux changements ou de collecter des informations importantes. Le plus souvent, les réponses mettent l'accent sur l'intérêt des exemples de bonnes pratiques. La troisième question, sur la façon dont les collaborateurs perçoivent le benchmarking, a suscité des réactions variées, les commentaires pouvant être très positifs (une occasion de participer et de s'impliquer) ou négatifs, avec des phrases comme « cela ne vas pas faire le bonheur de tout le monde » ou des remarques soulignant que même s'il existe une adhésion orale des collaborateurs, l'engagement réel risque d'être faible. En toute logique, les réponses à la question relative au soutien des collaborateurs sont également très hétérogènes, allant de l'approbation au franc rejet.

En conclusion, les personnes ont été interrogées sur la manière de poursuivre le développement de l'outil. Outre les critiques dont nous avons parlé sur les questions de langue, les réponses sont tantôt élogieuses (« Tous les aspects de l'évaluation sont traités » ou bien « C'est simple à utiliser ») ou tantôt critiques (par exemple, on ne voit pas suffisamment rapidement quand on a fini de répondre aux questions). Concernant le contenu, certaines personnes souhaiteraient par exemple qu'il soit possible de réaliser une comparaison entre l'état des choses et les objectifs ou de prioriser les

mesures d'assurance qualité. D'autres regrettent qu'il ne soit pas possible d'évaluer les données ; cette option n'a été mise en place qu'à la fin du projet (voir ci-dessous).

3. Possibilités d'analyse.

Les possibilités d'analyse sont basées sur les possibilités de comparaison de benchmarking en général et sur les particularités de l'outil en particulier. Les possibilités suivantes sont prévues :

1 Evaluation basée sur des modifications (analyse des séries temporelles).

2 Evaluation basée sur un groupe de référence.

3 Evaluation basée sur les correspondances (= comparaison du comportement de réponse général) entre les prestataires, les intervenants et les apprenants.

4 Evaluation à partir d'une norme (analyse de benchmarking).

Ces quatre possibilités d'évaluation ont été mises en place sur l'application en ligne. Par ailleurs, deux documents présentent les procédures d'évaluation et expliquent comment obtenir un aperçu des résultats. Pour de plus amples informations sur les possibilités d'évaluation du contenu et de la forme, nous vous invitons à vous reporter à ces deux documents disponibles sur le site Internet.