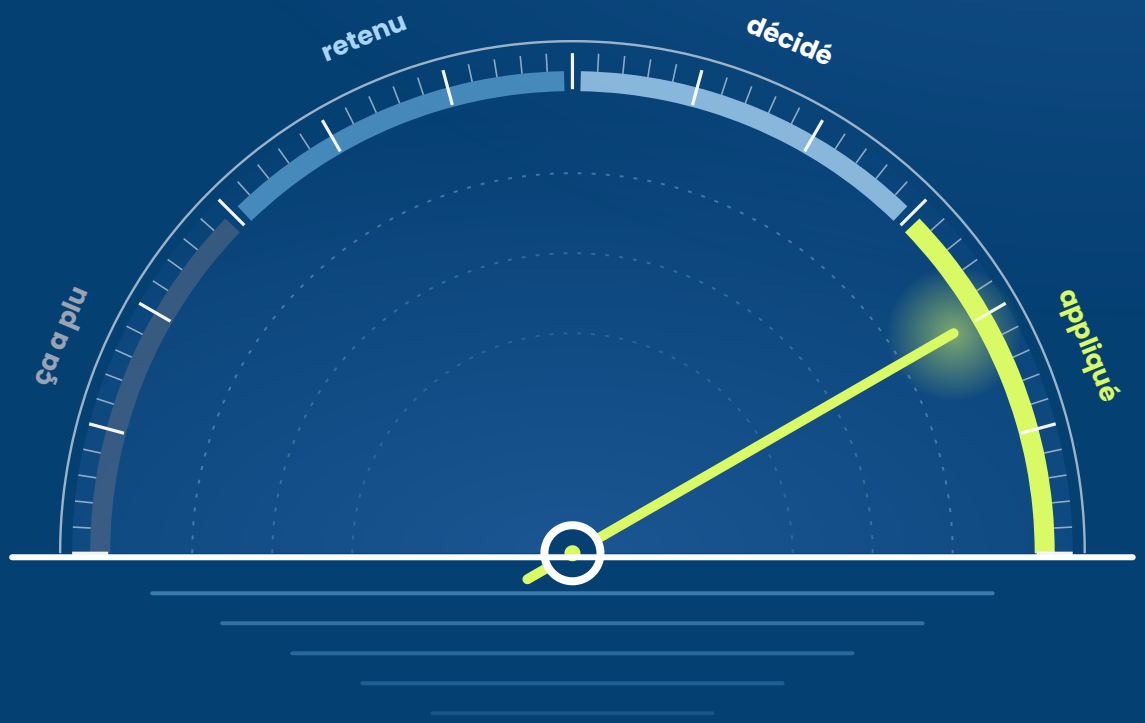
Cards

Édition novembre 2026

Guide pratique · Évaluation, ROI de la formation, pilotage

Mesurer l'ancrage sans promettre l'impossible

Ce qui se mesure honnêtement à J+30 et J+90, et comment le présenter à un CODIR



Ancrer les bons réflexes. Activer les bons savoirs.

AVANT DE COMMENCER

La question qui revient toujours



Dans presque tous les rendez-vous que nous avons avec des responsables formation, une question arrive tôt ou tard. Parfois c'est eux qui la posent, parfois c'est leur direction qui la leur a posée : « Comment est-ce qu'on mesure l'impact ? »

La question est légitime. La formation coûte du temps et de l'argent, et une direction a raison de vouloir savoir ce qu'elle produit. Le problème vient des réponses qu'on y apporte. D'un côté, des taux de satisfaction à 92 % qui ne disent rien de ce qui reste. De l'autre, des retours sur investissement spectaculaires, calculés à partir d'hypothèses que personne ne pourrait défendre devant un contrôleur de gestion.

Entre les 2, il existe un espace beaucoup plus solide. On peut mesurer, sans tricher, ce que les personnes ont retenu, ce qu'elles savent décider en situation, ce qu'elles déclarent avoir appliqué, et ce que leurs managers observent. On peut documenter des cas concrets où la formation a changé quelque chose. On peut raisonner sur le coût de ce qu'on oublie. Et on peut dire clairement ce qu'on ne peut pas mesurer.

Ce guide défend une idée simple : une mesure honnête et limitée convainc mieux qu'une mesure spectaculaire et fragile. Un CODIR expérimenté fait toujours la différence. Il s'appuie sur la recherche en évaluation de la formation, souvent mal connue en entreprise, et sur ce que nous voyons fonctionner sur le terrain.

Matthieu THOMAS
Cofondateur de Cards

Sommaire

	Avant de commencer	2
01.	**Ce que mesurent vraiment vos indicateurs habituels**	4
02.	**Kirkpatrick, et ce qui est venu après**	6
03.	**Ce qui se mesure honnêtement à J+30 et J+90**	8
04.	**La question du ROI**	10
05.	**Les chiffres qu'il vaut mieux ne plus citer**	12
06.	**Ce que l'IA change dans l'évaluation**	13
07.	**Le tableau de bord en 1 page**	16
08.	**3 routines de mesure, prêtes à adapter**	17
09.	**3 cas pratiques**	20
10.	**Sur le terrain, les erreurs que l'on voit le plus**	24
11.	**Construire votre dispositif de mesure en 6 semaines**	25
	Le tableau de bord en 1 page, prêt à remplir	26
	Comment nous le faisons chez Cards	27

Les sources sont citées en notes de bas de page. Vérifiées le 13 novembre 2026.

01.

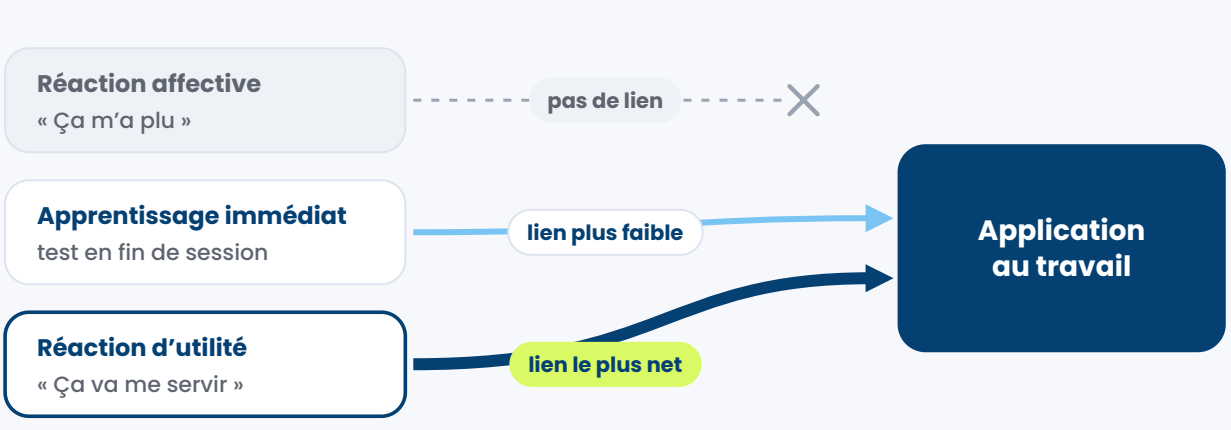
Ce que mesurent vraiment vos indicateurs habituels

Le questionnaire de satisfaction mesure surtout le plaisir. Et le plaisir prédit mal ce qui se passera ensuite.

L'indicateur le plus répandu en formation est le questionnaire à chaud, rempli à la fin d'une session. Il porte en général sur l'animateur, le contenu, le rythme, l'organisation, parfois le repas. Il est facile à collecter, facile à présenter, et presque toujours positif.

En 1997, George Alliger, Scott Tannenbaum et leurs collègues[1] ont publié dans Personnel Psychology une méta-analyse de 34 études et 115 corrélations entre les différents critères d'évaluation de la formation. Ils ont distingué 2 types de réactions. Les réactions affectives : est-ce que la formation m'a plu ? Les réactions d'utilité : est-ce que la formation me sera utile dans mon travail ? Leur résultat est net. Les réactions d'utilité sont nettement mieux reliées à l'application au travail que les réactions affectives, et même mieux que les mesures d'apprentissage immédiat[2].

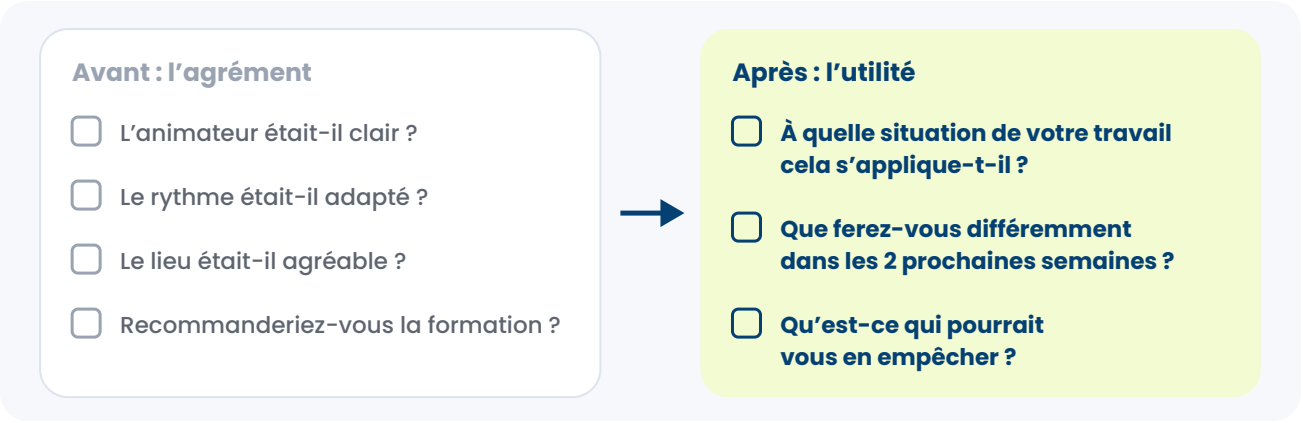
La méta-analyse de Brian Blume et de ses collègues sur le transfert de la formation[3], publiée en 2010, a confirmé ce point : les réactions d'utilité ont un lien modeste avec le transfert, les réactions affectives n'en ont pas.



Épaisseur du trait : force du lien avec l'application, d'après les méta-analyses citées.

1. Alliger, G. M., Tannenbaum, S. I., Bennett, W. Jr., Traver, H. et Shotland, A. (1997). A meta-analysis of the relations among training criteria. Personnel Psychology, 50, 341-358.
2. Voir aussi Sitzmann, T., Brown, K. G., Casper, W. J., Ely, K. et Zimmerman, R. D. (2008). A review and meta-analysis of the nomological network of trainee reactions. Journal of Applied Psychology, 93(2), 280-295.
3. Blume, B. D., Ford, J. K., Baldwin, T. T. et Huang, J. L. (2010). Transfer of training: A meta-analytic review. Journal of Management, 36(4), 1065-1105. doi.org/10.1177/0149206308321554

La conséquence pratique est immédiate, et elle ne coûte rien. Réécrivez votre questionnaire à chaud. Remplacez la moitié des questions sur l'agrément par des questions sur l'utilité : à quelle situation de votre travail ce que vous avez appris s'applique-t-il ? Qu'allez-vous faire différemment dans les 2 prochaines semaines ? Qu'est-ce qui pourrait vous en empêcher ? Ces questions sont plus utiles pour vous, et elles préparent déjà l'application.



Le même questionnaire, réécrit autour de l'utilité.

L'autre indicateur courant est le taux de complétion. Il dit qu'une personne est allée au bout d'un module. Il ne dit rien de ce qu'elle en a retenu. Un module de 20 minutes terminé en 6 minutes compte comme un module terminé.

> **Un taux de satisfaction à 92 % dit que la formation a plu. Il ne dit presque rien de ce qu'il en restera dans 3 mois.**

02.

Kirkpatrick, et ce qui est venu après

Le modèle le plus connu a 4 niveaux. Un modèle plus récent en a 8, et il dit mieux où se situe la vraie difficulté.

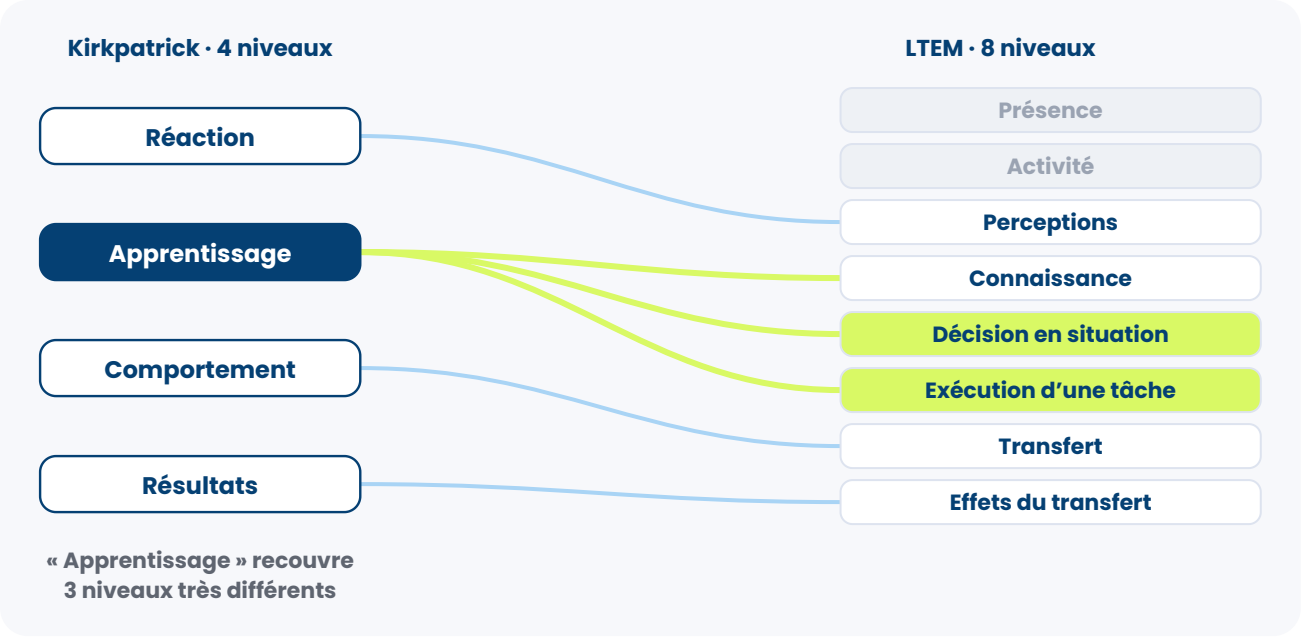
Le modèle de Donald Kirkpatrick[4], formulé à la fin des années 1950, structure l'évaluation en 4 niveaux : réaction, apprentissage, comportement, résultats. Il reste la référence de la plupart des entreprises, et il a un grand mérite : rappeler que la satisfaction n'est que le premier étage. Sa limite est qu'il regroupe sous un même mot des choses très différentes. « Apprentissage » peut désigner la capacité à réciter une définition comme la capacité à prendre la bonne décision dans une situation complexe.

En 2018, Will Thalheimer[5], consultant et chercheur américain en apprentissage, a proposé un modèle plus détaillé, le Learning-Transfer Evaluation Model, ou LTEM. Il compte 8 niveaux. Les 2 premiers, la présence et l'activité, sont explicitement jugés insuffisants pour évaluer l'apprentissage. Viennent ensuite les perceptions des apprenants, puis la connaissance, puis la compétence de décision, puis la compétence d'exécution d'une tâche, puis le transfert au travail, et enfin les effets de ce transfert.



Les 8 niveaux du LTEM, d'après Will Thalheimer (2018).

4. Kirkpatrick, D. L. et Kirkpatrick, J. D. (2006). Evaluating Training Programs: The Four Levels, 3e édition. Berrett-Koehler.
5. Thalheimer, W. (2018). The Learning-Transfer Evaluation Model (LTEM). Work-Learning Research. worklearning.com/ltem



Où se répartissent les 4 niveaux de Kirkpatrick dans les 8 niveaux du LTEM.

L'apport principal du modèle de Thalheimer, pour un responsable formation, est la distinction entre la connaissance et la compétence de décision. Savoir réciter la règle et savoir l'appliquer quand une situation se présente sont 2 niveaux différents. Le second se mesure par des mises en situation réalistes, que l'on peut administrer à distance, à J+30 ou J+90, sans observer chaque personne au travail. C'est le niveau le plus haut que l'on puisse mesurer régulièrement, à coût raisonnable, pour toute une population.

Thalheimer distingue aussi la connaissance mesurée juste après la formation et celle qui est retenue plusieurs jours plus tard. La première surestime toujours la seconde. Un test passé en fin de session mesure surtout la mémoire de travail du moment. Un test passé à J+30 mesure ce qui reste.

03.

Ce qui se mesure honnêtement à J+30 et J+90

5 mesures, toutes réalisables sans dispositif lourd, et ce que chacune permet de dire.

La rétention se mesure par des questions posées plusieurs semaines après la formation, sur ce qui a été enseigné. Elle dit ce qui reste. Elle est simple, peu coûteuse, et c'est la mesure qui révèle le plus vite les contenus qui ne tiennent pas.

La compétence de décision se mesure par des mises en situation : une situation de travail réaliste, plusieurs options, une seule bonne. Elle dit si la personne sait appliquer, dans un contexte proche du réel. Elle demande plus d'effort de conception, mais c'est la mesure la plus parlante pour un manager ou une direction.

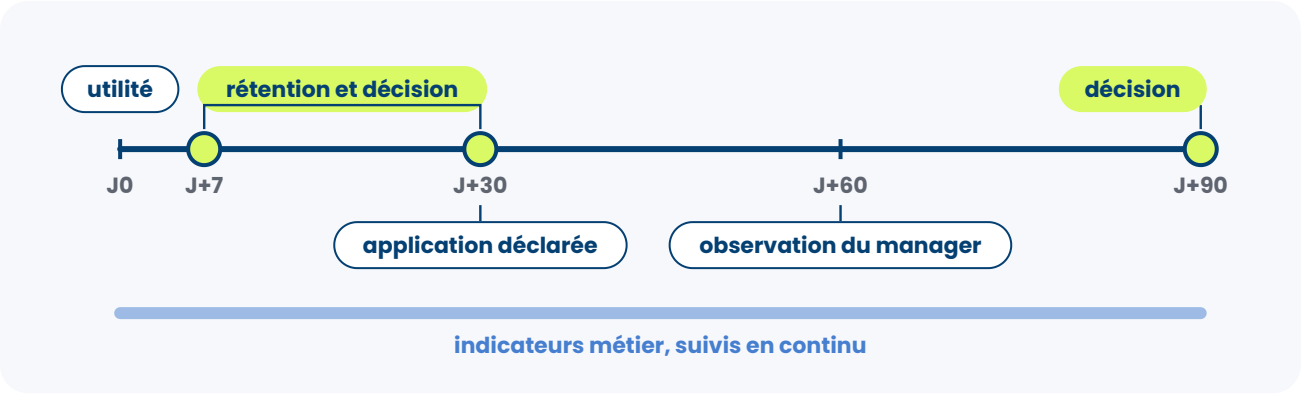
L'application déclarée se mesure par une question directe, à J+30 : avez-vous eu l'occasion d'appliquer tel point ? Si oui, qu'est-ce que ça a donné ? Si non, qu'est-ce qui vous en a empêché ? Elle est subjective, mais les réponses aux 2 dernières questions sont souvent les plus riches de tout le dispositif, parce qu'elles révèlent les obstacles du terrain.

	effort	conviction	ce que la mesure permet de dire
Rétention	● ○ ○	● ● ○	ce qui reste
Décision en situation	● ● ○	● ● ●	savoir appliquer
Application déclarée	● ○ ○	● ● ○	obstacles du terrain
Observation du manager	● ● ○	● ● ○	comportements visibles
Indicateurs métier	● ● ●	● ● ○	une direction, pas une preuve

Appréciation indicative de l'effort et du pouvoir de conviction de chaque mesure.

L'observation du manager se mesure par 2 ou 3 questions qui lui sont posées à J+60 sur des comportements observables, pas sur une impression générale. Elle est imparfaite, mais elle implique le manager, et la recherche sur le transfert montre que son soutien est l'un des facteurs les plus déterminants.

Les indicateurs métier se choisissent avant la formation, pas après. Taux d'erreur, réclamations, délai de traitement, presqu'accidents, taux de transformation : ceux que la formation est censée faire évoluer. Ils ne prouvent pas à eux seuls que la formation est la cause d'une évolution, mais suivis dans la durée, et comparés quand c'est possible entre équipes, ils donnent une direction.



Les 5 mesures, à leur moment.

Ce qui ne se mesure pas, et qu'il faut écrire

Dans la plupart des entreprises, il est impossible d'attribuer de façon rigoureuse l'évolution d'un indicateur métier à une formation. Trop de facteurs bougent en même temps : marché, recrutements, organisation, saison. Il est aussi impossible de mesurer le comportement réel de chaque personne, en permanence, sans un dispositif d'observation que personne ne veut. Ces limites ne sont pas une faiblesse du dispositif. Les écrire dans le rapport est ce qui rend le reste crédible.

04.

La question du ROI

Le calcul rigoureux existe, mais il demande des conditions rarement réunies. Il y a des alternatives plus honnêtes, et souvent plus convaincantes.

Jack Phillips[6] a ajouté un cinquième niveau au modèle de Kirkpatrick : le retour sur investissement, qui compare la valeur monétaire des bénéfices au coût complet de la formation. Sa méthode est sérieuse, et elle insiste sur un point que beaucoup de calculs de ROI ignorent : il faut isoler l'effet de la formation de tous les autres facteurs. Les techniques proposées vont du groupe de contrôle à l'analyse de tendance, jusqu'aux estimations faites par les participants et corrigées par leur degré de confiance.



Schéma de principe : la répartition réelle reste inconnue tant que l'effet n'est pas isolé.

En pratique, peu d'entreprises peuvent réunir ces conditions. Un groupe de contrôle suppose de priver une partie des équipes de la formation pendant un temps. Une analyse de tendance suppose des données historiques stables. Les estimations corrigées sont honnêtes, mais elles restent des estimations. Un ROI calculé sans ces précautions n'est pas un ROI : c'est un argument commercial habillé en chiffre.

6. Phillips, J. J. et Phillips, P. P. (2016). Handbook of Training Evaluation and Measurement Methods, 4e édition. Routledge.

L'alternative des cas de réussite

Robert Brinkerhoff[7], chercheur à l'université Western Michigan, a proposé dans les années 2000 une méthode qui part d'une autre question. Plutôt que de chercher l'effet moyen d'une formation sur tout le monde, il propose d'identifier les personnes qui en ont tiré le plus et celles qui en ont tiré le moins, puis de mener des entretiens pour comprendre ce qui a fait la différence. C'est la Success Case Method.

Ses avantages sont nombreux. Elle produit des cas concrets, vérifiables, avec une valeur que l'on peut documenter : une réclamation évitée, un client conservé, un incident évité, un délai réduit. Elle révèle aussi ce qui empêche les autres de réussir, souvent des obstacles d'organisation qui n'ont rien à voir avec la formation. Et elle se présente beaucoup mieux à un CODIR qu'un ratio que personne ne croit vraiment.



Raisonner sur le coût de l'oubli

Une troisième approche consiste à inverser la question. Au lieu de calculer ce que la formation rapporte, on estime ce que coûte ce qu'on oublie : les erreurs liées à une procédure mal appliquée, le temps passé à redemander, les reprises, les incidents. Le raisonnement reste fondé sur des hypothèses, mais elles sont plus faciles à poser et à discuter avec un contrôleur de gestion, parce qu'elles partent de coûts que l'entreprise connaît déjà. Le calculateur fourni avec ce guide propose ce raisonnement. Sa règle est la même que pour le reste : des hypothèses prudentes, affichées, et modifiables par celui qui lit le résultat.

> **Un ROI calculé sans isoler l'effet de la formation n'est pas un ROI. C'est un argument commercial habillé en chiffre.**

7. Brinkerhoff, R. O. (2003). The Success Case Method: Find Out Quickly What's Working and What's Not. Berrett-Koehler.

05.

Les chiffres qu'il vaut mieux ne plus citer

Certains chiffres circulent depuis des décennies dans les présentations sur la formation. Ils fragilisent ceux qui les citent.

Le plus répandu est celui de l'oubli : « 70 % de ce qu'on apprend est oublié en 24 heures », parfois attribué à Ebbinghaus. Les expériences d'Ebbinghaus portaient sur des syllabes sans signification, apprises par lui-même. Leur réplication moderne par Murre et Dros en 2015[8] confirme la forme générale de la courbe, mais ses pourcentages ne se transposent pas à une procédure métier qui a du sens et qu'on pratique.

Le deuxième est la pyramide de l'apprentissage, qui affirme qu'on retient 10 % de ce qu'on lit, 20 % de ce qu'on entend, et jusqu'à 90 % de ce qu'on enseigne aux autres. Ces pourcentages n'ont pas de source scientifique identifiable. Le troisième est le ROI sans méthode, du type « chaque euro investi en formation en rapporte 4 ».

~~70 %~~

oubliés en 24 heures

ce qui est établi

L'oubli est rapide au début, puis ralentit. Son ampleur dépend du contenu et de l'usage.

~~10 % à 90 %~~

la pyramide de l'apprentissage

ce qui est établi

On apprend mieux en étant actif. Les pourcentages n'ont pas de source.

~~1 € = 4 €~~

le ROI sans méthode

ce qui est établi

Sans population, période et isolement de l'effet, le ratio ne dit rien.

Citer ces chiffres devant un public averti, un DAF, un auditeur, un chercheur, c'est perdre en une phrase la crédibilité de tout le reste. Il vaut mieux s'en tenir à ce qui est établi, et le dire avec ses nuances.

8. Murre, J. M. J. et Dros, J. (2015). Replication and analysis of Ebbinghaus' forgetting curve. PLOS ONE. doi.org/10.1371/journal.pone.0120644

06.

Ce que l'IA change dans l'évaluation

Elle rend exploitables des données qu'on ne lisait pas. Elle ne doit pas servir à juger les personnes.

L'évaluation produit beaucoup de texte libre : commentaires des questionnaires, réponses aux questions d'application, retours des managers. Ces textes sont souvent les plus riches du dispositif, et ce sont ceux qu'on ne lit pas, faute de temps. Sur 400 réponses, on en parcourt 20.

L'IA générative change cette situation. Elle peut regrouper des centaines de réponses ouvertes par thèmes, repérer les obstacles cités le plus souvent, distinguer ce qui relève de la formation de ce qui relève de l'organisation, et faire remonter les citations les plus représentatives. Elle peut aussi aider à préparer les entretiens de la méthode des cas de réussite, et à en synthétiser les comptes rendus.



Exemple fictif de thèmes issus des réponses ouvertes.

Les limites à poser

La première limite est la fidélité. Une synthèse produite par une IA peut lisser, généraliser ou surinterpréter. Chaque thème qu'elle propose doit être accompagné des citations d'origine, et quelqu'un doit en relire un échantillon.

La deuxième limite est l'usage. L'analyse doit porter sur des tendances collectives, pas sur des personnes. Utiliser l'IA pour évaluer individuellement les salariés à partir de leurs réponses changerait la nature du dispositif, poserait des questions juridiques sérieuses au regard du RGPD et du Code du travail, et ruinerait la confiance nécessaire pour que les réponses restent sincères. Les réponses doivent être anonymisées avant analyse.

Un prompt de départ

À utiliser avec un export anonymisé des réponses ouvertes :

> Voici les réponses anonymisées à la question « Avez-vous appliqué ce que vous avez appris ? Qu'est-ce qui vous a aidé ou empêché ? ».
>
> Regroupe-les en 5 à 8 thèmes au maximum.
>
> Pour chaque thème : une phrase de description, le nombre approximatif de réponses concernées, et 2 citations exactes représentatives.
>
> Distingue les obstacles liés à la formation elle-même de ceux liés à l'organisation du travail.
>
> N'invente aucun thème qui ne soit pas appuyé par au moins 3 réponses.



Exemple fictif de restitution d'un thème.

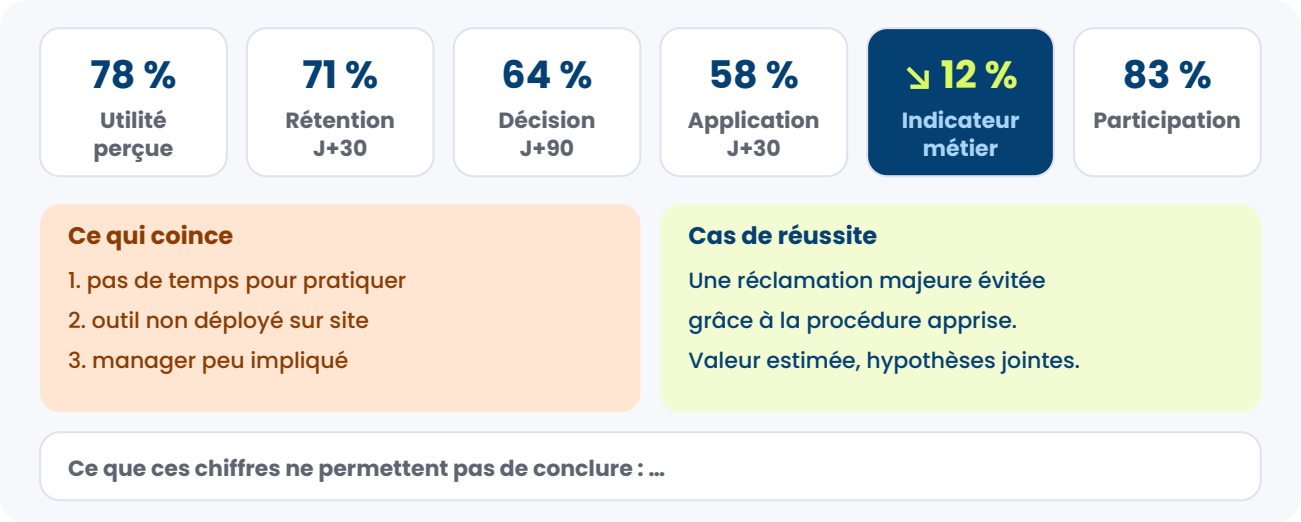
07.

Le tableau de bord en 1 page

6 indicateurs suffisent pour piloter et pour rendre compte. Au-delà, personne ne lit.

Le tableau de bord idéal tient sur 1 page et répond à 3 questions : qu'est-ce qui reste, qu'est-ce qui est appliqué, et qu'est-ce qui coince. Nous proposons 6 indicateurs.

Le taux d'utilité perçue, à chaud : la part des participants qui voient une situation précise de leur travail où appliquer la formation. La rétention à J+30 : le taux de bonnes réponses aux questions posées un mois après. La compétence de décision à J+90 : le taux de réussite aux mises en situation, 3 mois après. L'application déclarée à J+30 : la part des participants qui disent avoir appliqué au moins un point. Les 3 obstacles les plus cités, issus des réponses ouvertes. Et l'indicateur métier choisi au départ, avec sa tendance et une phrase honnête sur ce qu'il permet ou non de conclure. À ces 6 indicateurs s'ajoute une rubrique qualitative : 1 ou 2 cas de réussite documentés, et ce que les cas d'échec ont appris.



Exemple fictif. Le gabarit vierge est en fin de guide.

08.

3 routines de mesure, prêtes à adapter

Une routine de mesure est une suite d'étapes dans le temps, avec ses conditions. Elle évite que l'évaluation se résume au questionnaire de fin de session.

Routine 1 · L'évaluation d'une formation, sur 4 mois

Évaluation d'une formation, de J0 à J+120

J0	Questionnaire de fin de session, réécrit autour de l'utilité, avec 1 question ouverte : « à quelle situation de votre travail cela s'applique-t-il ? ».
J+7	3 questions de rétention sur les points clés.
J+30	3 questions de rétention, 2 mises en situation, et la question « avez-vous appliqué ? qu'est-ce qui vous a aidé ou empêché ? ».
Condition	Moins de 2 bonnes réponses sur 3 en rétention, et un contenu de reprise ciblé avant le rendez-vous suivant.
J+60	3 questions au manager sur des comportements observables.
J+90	4 mises en situation nouvelles.
J+100	Analyse des réponses ouvertes, identification des participants qui ont le plus et le moins appliqué.
J+120	5 à 8 entretiens courts de cas de réussite et d'échec, puis le tableau de bord en 1 page.

Routine 2 · Le pilote comparé

Comparer 2 groupes, sur 3 mois

Avant	2 groupes comparables (2 régions, 2 sites, 2 équipes), et l'indicateur métier choisi, mesuré sur les 3 mois précédents.
Mois 0	La même formation initiale pour les 2 groupes.
Mois 1 à 3	Le groupe A suit une routine d'ancrage (rétention, mises en situation, réactivation espacée). Le groupe B ne reçoit rien de plus.
J+30, J+90	Les mêmes mises en situation pour les 2 groupes.
Mois 3	Comparaison des résultats aux mises en situation, et de l'évolution de l'indicateur métier, avec la liste de tout ce qui a pu différer entre les 2 groupes.
Ensuite	Le groupe B reçoit à son tour la routine, ce qui règle la question d'équité.



Les 2 groupes suivent la même formation, puis se rejoignent au moment de la comparaison.

Routine 3 · La revue trimestrielle

Revue du dispositif, chaque trimestre

Semaine 1	Export des résultats question par question.
Semaine 1	Les questions réussies par plus de 95 % des participants sont ré-écrites : elles n'apprennent plus rien. Celles réussies par moins de 40 % sont vérifiées : elles sont peut-être ambiguës.
Semaine 2	Analyse des réponses ouvertes du trimestre avec l'IA, et relecture d'un échantillon.
Semaine 2	Le tableau de bord en 1 page, présenté au comité qui pilote la formation.
Semaine 3	1 décision par trimestre, au minimum, sur ce qui change dans les contenus ou l'organisation.



Exemple fictif : la répartition des taux de réussite par question montre ce qu'il faut réécrire et ce qu'il faut vérifier.

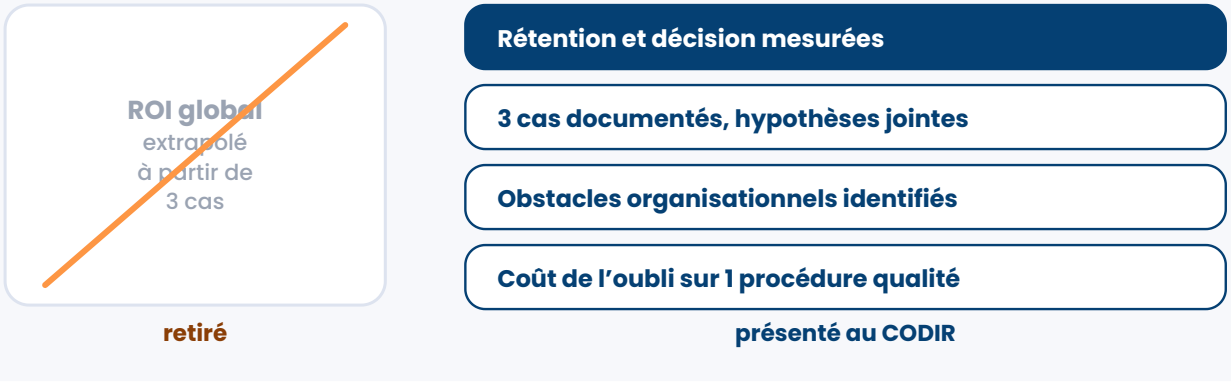
09.

3 cas pratiques

Les cas qui suivent sont des cas composites, construits à partir de situations fréquentes pour illustrer la démarche. Les chiffres sont des ordres de grandeur, pas des résultats mesurés.

Cas 1 Une DRH qui doit défendre son budget au CODIR

Le problème	Dans une ETI industrielle de 1 500 personnes, le budget formation est discuté chaque année. Cette année, le directeur financier demande un ROI. La DRH dispose de taux de satisfaction élevés et d'heures de formation réalisées, rien d'autre.
Le diagnostic	Calculer un ROI rigoureux est impossible dans le délai : pas de groupe de contrôle, pas d'indicateurs choisis à l'avance. Présenter un ROI approximatif serait risqué face à un directeur financier qui en vérifierait les hypothèses.
Le dispositif	Sur les 2 formations les plus coûteuses, la DRH lance la routine 1 en urgence, en partant de J+30 pour les dernières sessions. En parallèle, elle mène 8 entretiens de cas de réussite et d'échec. Elle présente au CODIR la rétention et la compétence de décision mesurées, 3 cas documentés avec leur valeur estimée et ses hypothèses, les obstacles organisationnels identifiés, et une estimation du coût de l'oubli sur une procédure qualité à partir des non-conformités connues.
Ce qui n'a pas marché au début	La première version de la présentation contenait un ROI global calculé à partir des cas de réussite, extrapolés à tous les participants. Le directeur financier l'aurait démonté en une question. Il a été retiré au profit des cas eux-mêmes.
Ce qu'on en retient	Le budget a été maintenu, avec une condition : choisir les indicateurs métier avant les formations de l'année suivante. C'est exactement ce qu'il fallait obtenir.



Cas 1 : ce qui a été retiré, et ce qui a été présenté.

Cas 2 Un programme managérial dont les résultats sont insaisissables

Le problème

Un groupe de services a formé 200 managers de proximité sur 6 mois. Le programme est apprécié, mais personne ne sait dire s'il a changé quelque chose. Les indicateurs RH, turnover et absentéisme, bougent pour bien d'autres raisons.

Le diagnostic

Les résultats d'un programme managérial sont lents, diffus et influencés par tout le reste. Chercher à les attribuer directement à la formation est voué à l'échec. En revanche, la compétence de décision et l'application déclarée sont mesurables.

Le dispositif

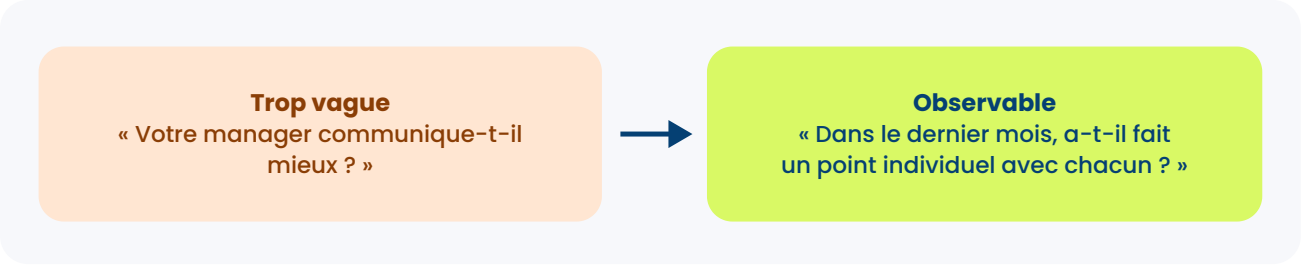
Des mises en situation managériales, construites à partir de situations réelles remontées par les RH : un entretien de recadrage, un conflit dans l'équipe, une demande de télétravail contestée. Elles sont posées à J+30 et J+90. À J+60, 3 questions aux N+1 des managers sur des comportements observables. À J+120, 10 entretiens de cas de réussite. L'enquête d'engagement annuelle est suivie, sans lui faire porter plus que ce qu'elle peut dire.

Ce qui n'a pas marché au début

Les questions posées aux N+1 étaient trop générales. Elles ont été remplacées par des comportements précis.

Ce qu'on en retient

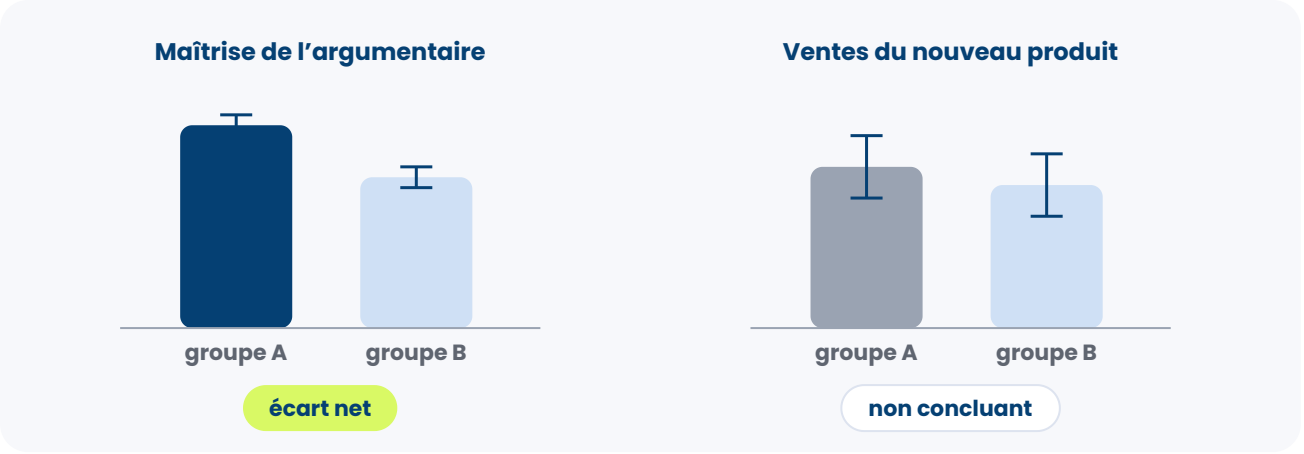
Pour un programme managérial, la valeur de la mesure est surtout d'identifier les situations que les managers ne maîtrisent toujours pas. Elle a conduit à refaire 2 modules sur 6.



Cas 2 : la question posée aux N+1, avant et après.

Cas 3 Une formation commerciale et un pilote comparé

Le problème	Un fabricant lance un nouveau produit et forme ses 120 commerciaux. La direction commerciale veut savoir si une routine d'ancrage après la formation améliore les ventes du nouveau produit.
Le diagnostic	C'est l'un des rares cas où une comparaison est possible : 2 régions de taille et de profil proches, un indicateur métier clair, une période courte.
Le dispositif	La routine 2, avec la région Nord en groupe A et la région Sud en groupe B. Les mises en situation portent sur les objections clients et l'argumentaire du nouveau produit. Les ventes du nouveau produit sont suivies sur 3 mois.
Ce qui n'a pas marché au début	En cours de période, la région Sud a perdu un gros client pour une raison sans rapport avec la formation. La comparaison des ventes est devenue difficile à interpréter. Les résultats aux mises en situation, eux, sont restés comparables, et ont montré un écart net en faveur du groupe A sur la maîtrise de l'argumentaire.
Ce qu'on en retient	Même un pilote bien conçu est exposé aux aléas du terrain. Le rapport final a présenté la différence de maîtrise comme solide, et la différence de ventes comme encourageante mais non concluante. La direction commerciale a jugé cette honnêteté plus convaincante qu'un chiffre de croissance.



Cas 3, schéma illustratif : un écart net sur la maîtrise, un écart non concluant sur les ventes. Les barres verticales figurent l'incertitude.


Photo : Redd Francisco · Unsplash

10.

Sur le terrain, les erreurs que l'on voit le plus

01 **La première erreur est de choisir les indicateurs après la formation.** On trouve toujours un indicateur qui a progressé. C'est précisément pour cela qu'il faut le choisir avant.

02 **La deuxième est de mesurer la connaissance en fin de session.** Le résultat est flatteur et trompeur. Ce qui compte est ce qui reste plusieurs semaines plus tard.

03 **La troisième est de poser des questions de restitution au lieu de mises en situation.** Savoir réciter une règle et savoir l'appliquer sont 2 compétences différentes.

04 **La quatrième est de se passer du manager.** C'est le seul qui voit le travail réel, et son implication dans la mesure est déjà un facteur de transfert.

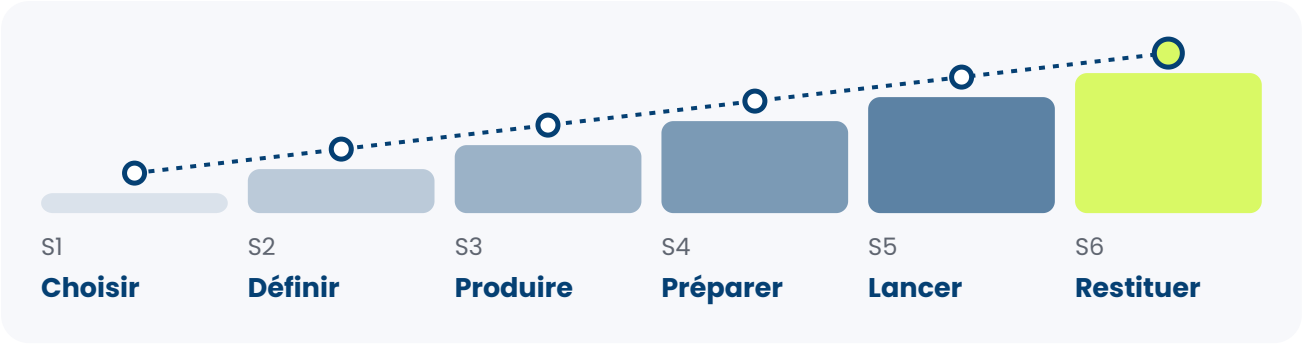
05 **La cinquième est de ne pas lire les réponses ouvertes.** Ce sont elles qui disent pourquoi les choses ne marchent pas, et elles sont aujourd'hui faciles à analyser.

06 **La sixième est d'attribuer à la formation tout ce qui bouge.** Un indicateur qui progresse pendant un programme n'est pas une preuve de son efficacité. Le dire soi-même évite que quelqu'un d'autre le dise.

07 **La septième est de vouloir tout mesurer.** Un tableau de bord de 25 indicateurs n'est lu par personne. 6 indicateurs bien choisis et 2 cas documentés valent mieux.

11.

Construire votre dispositif de mesure en 6 semaines



Semaine 1	**Choisir 1 formation pilote.** Une formation qui compte pour l'entreprise, avec une population d'au moins 50 personnes, et une session prévue dans les semaines qui viennent.
Semaine 2	**Définir ce qui doit changer.** Avec le commanditaire et 2 managers : quelles décisions les participants doivent-ils savoir prendre, quel indicateur métier devrait évoluer, et ce qu'on acceptera de conclure ou non. Mesurer l'indicateur sur les mois précédents.
Semaine 3	**Réécrire le questionnaire à chaud et produire les mises en situation.** Questions d'utilité, questions de rétention, mises en situation préparées avec l'IA à partir des contenus validés, puis relues par un expert.
Semaine 4	**Préparer les managers.** Expliquer leur rôle, les 3 questions qui leur seront posées à J+60, et pourquoi leur soutien compte.
Semaine 5	**Lancer** la routine 1 avec la prochaine session.
Semaine 6	**Préparer le tableau de bord.** Le gabarit en 1 page, rempli au fil des résultats, et la date de présentation au commanditaire fixée dès maintenant, à J+120.

Le tableau de bord en 1 page, prêt à remplir

À imprimer, ou à reproduire dans votre outil.

Formation _____ **Population** _____ **Dates** _____ **Commanditaire** _____

Utilité perçue à chaud
_____ %

Rétention à J+30
_____ %

Décision en situation à J+90
_____ %

Application déclarée à J+30
_____ %

Indicateur métier
_____ avant · après · tendance

Participation aux re-tests
_____ %

Ce qui coince

Les 3 obstacles les plus cités

_____

_____

_____

Cas de réussite

1 récit court, avec la valeur estimée et ses hypothèses

_____

_____

_____

Ce que nous changeons

1 à 3 décisions

_____

_____

Ce que ces chiffres ne permettent pas de conclure

_____

_____

Comment nous le faisons chez Cards

Cards a été conçu autour de la question de ce qui reste après une formation. Les questions de rétention et les mises en situation sont programmées dans le temps par Learning Routine®, et ajustées aux résultats de chacun. Les résultats sont suivis question par question dans le tableau de bord, ce qui permet de repérer les contenus qui ne tiennent pas et les questions à réécrire, et ils sont exportables pour vos propres analyses. Les questions d'application et les réponses ouvertes peuvent être intégrées aux routines. Les mises en situation peuvent être préparées avec l'assistant de conception intégré, ou depuis Claude, ChatGPT, Mistral ou Copilot grâce au connecteur MCP Cards, puis relues avant publication. Les données sont hébergées en France.

SNCF Université RH ancre avec Cards la formation managériale de 25 000 collaborateurs. Au-delà de la satisfaction, l'indicateur que le cas publié met en avant est celui qui compte vraiment.

96 %
de satisfaction

+40 %
de mise en pratique sur le terrain

POUR ALLER PLUS LOIN

cards-microlearning.com/formation/evaluation-a-froid-roi-formation

cards-microlearning.com/formation/evaluation-des-acquis

cards-microlearning.com/formation/cycle-dancrage

cards-microlearning.com/ressources/formation-sans-ancrage-cout