

STANDARDS D'ÉVALUATION

DE LA SOCIÉTÉ SUISSE D'ÉVALUATION (STANDARDS SEVAL)

Thomas Widmer, Charles Landert et Nicole Bachmann

5 décembre 2000

INTRODUCTION

Objectifs des standards d'évaluation

Les standards d'évaluation de la Société suisse d'évaluation, appelées ci-après standards SEVAL, doivent contribuer à la professionnalisation de l'évaluation en Suisse. Le respect des standards SEVAL accroît la crédibilité, la compétence et la fiabilité des évaluations. Seule la collaboration de tous les intéressés, à savoir

- les évaluateurs et évaluatrices,
- les mandantes et mandants et
- les autres personnes intéressées à l'évaluation,

crée la base d'évaluations de haute qualité. C'est pourquoi les standards SEVAL formulent des critères qui engagent *toutes* les personnes participant aux évaluations.

Les standards SEVAL définissent les exigences qui doivent être posées aux évaluations, mais pas les instruments pour leur mise en œuvre. Dans l'idéal, une évaluation doit répondre à toutes ces exigences. Dans le cas concret, toutefois, il n'est pas toujours possible de prendre en compte de la même manière chaque standard. Il s'agit plutôt d'adapter le standard SEVAL à la situation spécifique. Il peut arriver que, pour différentes raisons, certains standards soient jugés peu importants dans un cas d'espèce, alors qu'on attribue à d'autres un poids particulièrement grand. Toutefois, les adaptations aux conditions spécifiques d'une évaluation devraient être mûrement réfléchies, exposées sans détour et explicitement motivées. Elles devraient être discutées et définies en commun, au début de l'évaluation, par tous les intéressés (mandantes et mandants, évaluateurs et évaluatrices, destinataires et, le cas échéant, autres personnes intéressées).

Lors de la formulation des standards SEVAL, on a veillé avec un soin particulier à ce qu'ils se prêtent aux évaluations de toutes sortes (à l'exception des évaluations de personnel). De ce fait, il incombe à l'applicatrice ou applicateur de les concrétiser en fonction de sa situation spécifique. Cela ne concerne pas seulement l'opération d'adaptation

exposée plus haut, mais aussi la concrétisation du contenu des standards SEVAL. Ce processus doit intervenir afin d'assurer la transparence et la possibilité de reconstitution.

Champ d'application des standards SEVAL

Les standards SEVAL s'appliquent aux évaluations en général (à l'exception des évaluations de personnel), indépendamment du contexte institutionnel, de la démarche choisie (p.ex. évaluation interne ou externe) et du thème spécifique. L'expression «objet d'évaluation» se réfère indifféremment, dans les standards SEVAL, à des programmes, mesures, projets, organisations, institutions, politiques, produits, documents ou autres objets d'évaluation, à l'exclusion des personnes.

Destinataires des standards SEVAL

Les standards SEVAL sont destinés à toutes les personnes qui participent à des évaluations ou les influencent. En conséquence, ils s'adressent notamment aux évaluatrices et évaluateurs, mandantes et mandants, mais aussi aux autres personnes susceptibles de soutenir la prise en compte des standards SEVAL (telles les personnes qui s'occupent de formation et de perfectionnement en matière d'évaluation).

Genèse des standards SEVAL

Les standards SEVAL s'inspirent des «Program Evaluation Standards» du «Joint Committee on Standards for Educational Evaluation» (Joint Committee 1994 et 2000). Le groupe de travail «Standards d'évaluation» de la SEVAL s'est longuement penché sur ce thème. Le groupe de travail, qui réunit des représentantes et représentants de la Confédération et des cantons, de la pratique de l'évaluation et des milieux scientifiques, a constitué un comité chargé de la formulation des standards SEVAL. Les membres de ce comité ont établi les présents standards en consultation avec les membres du groupe de travail. Les documents d'accompagnement contiennent des indications sur la procédure suivie, ainsi qu'une liste des personnes ayant participé à l'élaboration des standards.

Rôle de la SEVAL

La SEVAL considère les standards SEVAL comme un instrument destiné à garantir et à promouvoir la qualité. Elle ne lie pas à cette fonction l'intention de voir les standards SEVAL employés dans le cadre de procédés d'accréditation ou de certification, procédés auxquels les standards SEVAL ne se prêtent pas.

La SEVAL promeut l'emploi des standards SEVAL dans la pratique de l'évaluation suisse. Elle contribue à l'évolution de ces standards en créant un comité qui suit régulièrement les développements actuels et qui, si besoin est, propose à l'assemblée SEVAL des changements appropriés des standards SEVAL. Ce comité n'a ni la mission de veiller sur le respect des standards SEVAL, ni les compétences pour sanctionner un manquement à leur application.

Aperçu général

Les standards SEVAL sont divisés en quatre groupes: utilité, faisabilité, déontologie et précision. Ces quatre qualités sont à chaque fois décrites dans une phrase d'introduction.

On trouve ensuite les différents standards relevant de chaque groupe. Ils comprennent une désignation et une description en une phrase et sont toujours accompagnés d'explications qui servent à en éclaircir le contenu.

Tant à propos des standards que pour les quatre groupes (utilité, faisabilité, déontologie et précision), on a renoncé à une pondération attribuant une importance particulière à tel standard ou tel groupe de standards. Cette approche tient au fait que l'importance tant des standards que des quatre groupes varie d'une évaluation à l'autre.

Après l'exposé des standards, on trouve une série de documents d'accompagnement, à savoir

- un aperçu - appelé «aperçu fonctionnel» - indiquant les standards qui occupent une place prioritaire lors des différentes activités du processus d'évaluation et auxquelles il convient donc de prêter une attention particulière. Cet aperçu permet de distinguer facilement les standards les plus utiles dans des situations spécifiques;
- un texte expliquant comment on a tiré les standards de la SEVAL des «Program Evaluation Standards» du «Joint Committee on Standards for Educational Evaluation» (1994 et 2000) et exposant les motifs des écarts par rapport au texte original;
- un tableau de transformation pour les désignations abrégées permettant de mettre en relation les désignations abrégées employées par les standards SEVAL et celles employées par la version américaine des « Program Evaluation Standards » (Joint Committee 1994).
- une liste des membres du groupe de travail «Standards d'évaluation» de la SEVAL;
- une liste des responsables de la traduction française;
- une bibliographie.

LES STANDARDS D'ÉVALUATION

Utilité

Les standards d'utilité garantissent qu'une évaluation est orientée sur les besoins d'information des utilisatrices et utilisateurs de l'évaluation prévus.

U1 Identification des parties prenantes et groupes concernés

On identifie les personnes participant à une évaluation de même que les personnes touchées par elle, de manière à prendre en compte leurs intérêts et leurs besoins.

Le cercle des personnes qui devrait être pris en compte dans le contexte d'une évaluation comprend avant tout les protagonistes suivants:

- le groupe qui va décider de l'avenir de l'objet d'évaluation (généralement, l'instance qui a la compétence financière),
- le groupe responsable de la conception / de la forme de l'objet d'évaluation,
- le groupe participant à la mise en œuvre pratique de l'objet à examiner (projet, programme, loi, produit, etc.),
- les groupes qui sont ou devraient être directement ou indirectement touchés par l'objet d'évaluation (les groupes cibles et leur environnement social),
- les autres milieux intéressés aux résultats de l'évaluation (p.ex. décideurs chargés de planifier des projets semblables, évaluatrices et évaluateurs, public).

Ces personnes, groupes et institutions sont appelés parties prenantes et groupes concernés («stakeholders»).

U2 Clarification des buts d'évaluation

Toute personne prenant part à une évaluation se préoccupe de clarifier les buts des évaluations pour chaque participant et chaque personne concernée.

Pour qu'une évaluation se déroule avec succès, il est primordial d'informer clairement tous les participants et toutes les personnes concernées sur les buts de l'évaluation. Il est attendu des acteurs prenant part à l'évaluation de se mettre d'accord entre eux quant à ses buts. Ceci évite les attentes exagérées vis-à-vis de l'évaluation (venant en particulier du demandeur de l'étude). En contrepartie, les acteurs responsables de la réalisation de l'évaluation s'engagent à poursuivre des buts d'évaluation bien précis. Cette clarification des buts aide à éviter des malentendus qui peuvent se présenter lors du processus d'évaluation.

La clarification des buts de l'évaluation est aussi d'une importance fondamentale dans l'objectif de promouvoir un procédé dirigé par le but. Il n'est souvent pas possible de clarifier définitivement les buts d'une évaluation dès son commencement. Une telle clarification demande un processus plus long qui représente lui-même un élément central du processus d'évaluation.

U3 Crédibilité

Le groupe qui exécute l'évaluation est aussi digne de confiance que compétent, afin que les résultats de l'évaluation bénéficient d'un maximum de crédibilité et d'acceptation.

La fiabilité des évaluatrices et évaluateurs influence de manière décisive la faisabilité et l'efficacité de l'évaluation. Pour être jugés dignes de confiance par les différents groupes intéressés, les évaluateurs doivent manifester avant tout les qualités suivantes: intégrité, indépendance, compétences sociales et de communication.

U4 Ampleur et choix des informations

L'ampleur et le choix des informations recueillies permettent de traiter les questions pertinentes par rapport à l'objet d'évaluation et prennent en considération en même temps les intérêts et les besoins du mandant et des autres parties prenantes et groupes concernés.

Dans la planification d'un projet d'évaluation, il faut distinguer quelles informations sont indispensables pour répondre aux questions d'évaluation et quelles informations, même si elles sont intéressantes et souhaitables, sont sans importance pour la réponse aux questions principales. A ce propos, il faut veiller à ce que les ressources disponibles pour la collecte des données soient investies conformément aux priorités qu'implique la réponse aux questions principales et aux besoins des groupes cibles les plus importants de l'évaluation.

U5 Transparence de l'appréciation

Les perspectives et raisonnements sur lesquels se fonde l'interprétation des résultats sont décrits de telle sorte que les bases des jugements de valeur apparaissent clairement.

L'interprétation des informations recueillies et des résultats constitue l'un des points les plus importants et les plus critiques d'un processus d'évaluation. Dans cette tâche d'interprétation, les évaluatrices et évaluateurs s'orientent sur une représentation théorique et un système de valeurs. Pour rendre cette démarche convaincante et reconstituable et faire en sorte qu'on puisse la juger, il est indispensable de présenter les bases des jugements de valeur de manière transparente.

U6 Caractère complet et clair du rapport

Les rapports d'évaluation décrivent l'objet d'évaluation et son contexte, de même que les objectifs, questions, procédures et résultats de l'évaluation, dans le souci de mettre à disposition les informations essentielles sous une forme facilement compréhensible.

Le caractère complet et la clarté du rapport (ou de tout autre moyen de présentation de l'évaluation) sont une condition à la communication convaincante des résultats de l'évaluation. Le langage choisi doit être précis (p.ex. définition des concepts importants et utilisation cohérente de ceux-ci) et compréhensible pour les lectrices et lecteurs. La présentation des résultats les plus importants sous forme de tableaux ou de graphiques facilite une compréhension rapide. Dans l'idéal, la forme et le genre du rapport sont conçus dans la perspective d'une perception optimale par le groupe cible visé. A cet égard, un rapport final volumineux ne représente pas la forme optimale de transmission des informations à chaque type de destinataires. Selon le destinataire et la situation, on suscite plus d'attention par des exposés, des ateliers ou d'autres formes de communication.

U7 Présentation du rapport en temps utile

Les résultats intermédiaires importants et les rapports finaux doivent être communiqués aux utilisateurs prévus de telle sorte que ceux-ci puissent en faire usage en temps utile.

Un projet d'évaluation perd une grande partie de l'effet visé si sa présentation ne correspond pas au calendrier des décisions que doivent prendre les destinataires. A ce propos, il ne faut pas oublier de prévoir, dans de nombreux cas (p.ex. mandat donné par l'administration publique), un délai important avant la présentation, car le rapport d'évaluation doit encore faire l'objet d'un traitement interne (p.ex. procédure de rapport conjoint, mise en consultation) avant que des décisions puissent être prises. Dans de nombreux projets d'évaluation, il apparaît également utile de communiquer des résultats provisoires et rapports intermédiaires au cours de l'enquête

déjà, avant tout quand les informations fournissent un motif d'action aux destinataires. Il est avantageux d'en tenir compte dès la conception de l'évaluation et de prévoir pour cela des ressources correspondantes.

U8 Effet de l'évaluation

La planification, l'exécution et la présentation d'une évaluation encouragent les parties prenantes et groupes concernés à suivre le processus d'évaluation et à tirer parti de l'opération.

Le degré d'application des résultats et des recommandations de l'évaluation dépend étroitement de la mesure dans laquelle les parties prenantes et groupes concernés attendent que l'évaluation ait une utilité pour eux-mêmes et pour leurs projets. Une condition importante pour encourager et satisfaire cette attente consiste à inclure les différents protagonistes dans la planification et l'organisation du projet d'évaluation. Le fait de fournir des informations en permanence et de manière transparente au cours de l'enquête exerce également un effet positif.

Faisabilité

Les standards de faisabilité garantissent qu'une évaluation est exécutée de manière réaliste et mûrement réfléchie, avec diplomatie et dans la conscience des coûts.

F1 Procédure pratique

La procédure d'évaluation permet de recueillir les informations nécessaires sans déranger inutilement l'objet d'évaluation ou l'évaluation elle-même.

Dans la planification et l'exécution de l'évaluation, il ne faut pas seulement viser à utiliser la procédure idéale du point de vue scientifique. Il est tout aussi important de faire en sorte que la méthode et la procédure choisies soient applicables de manière aussi pratique que possible et qu'elles ne perturbent pas inutilement l'objet d'évaluation ni les personnes touchées par l'investigation. Dans le cadre des projets d'évaluation, il n'est souvent pas possible d'utiliser les méthodes les plus pertinentes du point de vue scientifique, car elles apparaissent trop coûteuses, trop longues ou inacceptables du point de vue éthique dans la situation spécifique. A cet égard, il est important d'établir de manière transparente, lors de la planification du projet d'évaluation, les avantages et les inconvénients ainsi que la pertinence de la procédure choisie et d'en discuter avec les mandantes et mandants, ainsi que les groupes de destinataires.

F2 Acceptabilité politique

Les évaluations sont planifiées et exécutées en tenant compte des positions différentes des groupes d'intérêts en cause, afin d'obtenir leur coopération et d'éviter d'éventuelles tentatives de limiter les activités de l'évaluation, de déformer des résultats ou d'en faire un usage abusif.

Si on ne veut pas être surpris par des réactions négatives au projet d'évaluation, il importe de distinguer autant que possible tous les groupes d'intérêts. En dehors du cercle étroit des intéressés, il peut s'agir aussi de personnes qui se trouvent dans l'environnement plus général de l'objet d'évaluation (p.ex. aussi des personnes qui offrent un produit en quelque sorte concurrent de l'objet examiné). Si on parvient à cerner les besoins de ces divers groupes d'intérêts et, le cas échéant, à les prendre en compte, les réactions négatives à l'évaluation sont plus facilement prévisibles et, selon les circonstances, ne se manifestent même pas. Outre les intérêts évidents de ces groupes, des «agendas cachés» peuvent aussi avoir de l'importance. Dans ces

considérations sur les besoins explicites et implicites des différents protagonistes par rapport aux résultats de l'évaluation, il convient de ne pas exclure non plus l'instance mandante.

F3 Rapport coûts/avantages

Les évaluations fournissent des informations dont la valeur justifie les moyens engagés.

Une évaluation est efficiente quand le profit à attendre est égal ou supérieur au coût. Le *coût* se réfère à la valeur de toutes les ressources engagées; il comprend donc aussi le temps nécessaire à l'accompagnement de l'évaluation ou les frais supportés par d'autres institutions. Dans cette perspective, le coût représente la totalité de la valeur (sociale et monétaire) de toutes les ressources nécessaires à l'exécution de l'évaluation (coût global). L'*avantage* désigne la somme de toutes les valeurs générées par l'évaluation (optimisation des effets, possibilités d'économies, connaissance du degré d'acceptation d'un programme, etc.).

Dans une évaluation, le rapport coût/avantage doit être aussi optimal que possible. C'est ainsi que, parmi plusieurs modalités d'enquête apportant un avantage identique, on choisit celle qui implique les frais les plus bas. Dans le cas de modalités de valeur égale, on favorise celle dont on attend le plus grand avantage. Si le coût demeure dans tous les cas plus élevé que l'avantage à attendre, mieux vaut renoncer à l'évaluation.

Déontologie

Les standards de déontologie garantissent qu'une évaluation est exécutée correctement du point de vue juridique et éthique et qu'elle accorde à l'intégrité des parties prenantes et groupes concernés l'attention qui leur est due.

D1 Convention formelle

Les obligations des parties contractantes d'une évaluation (ce qui doit être fait, comment, par qui, quand) sont consignées par écrit, de sorte que les parties soient engagées à remplir toutes les conditions de la convention ou à les renégocier.

Au début d'une évaluation, les rapports entre les mandants ou mandantes et les évaluateurs ou évaluatrices sont marqués le plus souvent par le respect et la confiance mutuels. C'est là un environnement optimal pour fixer toutes les règles et obligations importantes des deux parties dans une convention écrite (contrat, confirmation de mandat, etc.). La convention formelle doit préciser au moins les aspects suivants: financement, durée, personnes, publication, contenu, méthodologie et démarche. On consigne de manière aussi précise que possible, en particulier, les droits et obligations des personnes intéressées. Si, au cours du temps, des changements apparaissent nécessaires, on a la faculté de renégocier les conditions. Une convention formelle écrite réduit le risque de malentendus entre les parties contractantes ou, le cas échéant, facilite la dissipation de ceux-ci.

D2 Protection des droits individuels

Les évaluations sont planifiées et exécutées de telle sorte que les droits et l'intégrité des personnes sont respectés et protégés.

Les individus disposent de droits qui s'appuient sur la loi ou reposent sur des règles éthiques, le «bon sens» ou la bienséance. Dans la planification et l'exécution d'une évaluation, les droits et l'intégrité des personnes ne doivent pas être touchés. Les personnes impliquées dans une évaluation doivent être informées de ce fait et les effets prévisibles de l'évaluation doivent être exposés de manière détaillée. Les mandants ou mandantes devraient se garder d'obliger les évaluateurs ou évaluatrices à prendre des décisions susceptibles d'enfreindre les lignes directrices

indiquées. Si une évaluation conduit à des conclusions qui mettent en danger l'intégrité de certaines personnes, il est impératif d'apprécier avec prudence la mesure dans laquelle la diffusion de ces conclusions est justifiée.

D3 Interaction empreinte d'humanité

Les évaluations sont conçues de telle sorte que les contacts entre les personnes intéressées se fondent sur le respect mutuel.

Les évaluatrices et évaluateurs devraient éviter de blesser, sous quelque forme que ce soit, la dignité et l'estime de soi des personnes avec lesquelles ils entrent en contact au cours de l'évaluation. Par un comportement approprié, on évitera que des sentiments d'hostilité ne se développent à l'encontre de l'évaluation, cela non seulement pour répondre à un impératif de dignité humaine, mais aussi pour des raisons pratiques: le fait que des personnes se sentent atteintes dans leur dignité ou leur estime de soi n'a pas seulement pour effet de leur faire perdre leur potentiel créatif. Elles peuvent aussi essayer d'interférer dans les activités de l'évaluation. C'est pourquoi il est nécessaire de connaître et de comprendre les valeurs culturelles et sociales des personnes impliquées et de distinguer l'importance que les différentes personnes accordent à l'évaluation.

D4 Appréciation complète et honnête

Les évaluations sont complètes et honnêtes dans l'examen et dans la représentation des forces et des faiblesses de l'objet d'évaluation, de manière à renforcer les points positifs et à faciliter la solution des problèmes.

La représentation équilibrée des forces et des faiblesses d'un objet d'évaluation vise à en permettre une appréciation complète et honnête. Même si l'objectif premier d'une évaluation est souvent la mise en évidence des points faibles, cela ne la dégage pas du devoir de rechercher les aspects positifs et de les mentionner. Il est souvent possible, en effet, de corriger les faiblesses existantes en faisant valoir des forces également existantes. D'autre part, il faut considérer que le fait de corriger les faiblesses de l'objet d'évaluation peut conduire à affaiblir des points forts. C'est pourquoi il est utile de soumettre le rapport, avant la fin de l'évaluation, à des personnes externes qui auront peut-être d'autres réactions à l'égard de la présentation des points positifs et négatifs. Si, pour diverses raisons (p.ex. limitations de temps ou restrictions financières), il n'est pas possible de collecter certaines données, ces lacunes doivent être exposées clairement. Les mandants et mandantes devraient s'abstenir d'interventions susceptibles de mettre en question la présentation d'un rapport équilibré.

D5 Publication des résultats

Les parties contractantes d'une évaluation garantissent que les résultats de l'évaluation sont accessibles aux personnes concernées, de même qu'à toutes celles qui sont en droit d'attendre une telle communication.

Lors de la diffusion du rapport, on veille à ce que toutes les personnes qui, sous une forme ou sous une autre, sont parties prenantes à l'évaluation ou concernées par elle aient accès au rapport. Etant donné que le cercle des parties prenantes et groupes concernés est souvent très large, un rapport doit, dans bien des cas, être rendu public. Les protagonistes d'une évaluation (en premier lieu les mandants ou mandantes et les évaluateurs ou évaluatrices, mais aussi d'autres personnes) portent ensemble la responsabilité de cette diffusion. En outre, le rapport devrait être rédigé sous une forme qui réponde aux besoins des destinataires. Cela implique souvent la nécessité de communiquer l'évaluation sous une forme adaptée au cercle des destinataires: par exemple, en préparant des résumés de rapports volumineux, des traductions ou en ajoutant une annexe contenant des indications méthodologiques.

D6 Identification des conflits d'intérêts

Les conflits d'intérêts sont abordés de manière ouverte et franche, afin qu'ils portent aussi peu atteinte que possible à la procédure et aux résultats de l'évaluation.

Dans de nombreuses situations, les évaluateurs et évaluatrices sont confrontés à des conflits d'intérêts; eux-mêmes ont aussi des intérêts propres qui, selon les circonstances, sont susceptibles de porter atteinte aux résultats d'une évaluation. Ainsi, les évaluatrices ou évaluateurs sont plus ou moins soucieux de continuer à recevoir des mandats, ont des conceptions spécifiques en matière philosophique, théorique, méthodologique ou politique et se situent dans un réseau de relations tant personnelles qu'organisationnelles. Des conflits entre des groupes d'intérêts qui se trouvent dans le même contexte peuvent aussi apparaître au cours de l'évaluation, ce qui risque de susciter une distorsion de l'évaluation dans son processus, ses résultats et ses interprétations. En principe, il faudrait éviter les conflits d'intérêts. Toutefois, étant donné la multiplicité mentionnée des conflits d'intérêts potentiels, il n'est souvent pas possible de les éviter entièrement. Dans ce cas, il importe de développer des moyens de les gérer sans que l'évaluation en souffre.

Précision

Les standards de précision garantissent qu'une évaluation fournit et transmet des informations valides et utilisables.

P1 Documentation de l'objet d'évaluation

L'objet d'évaluation est décrit et documenté de manière claire et précise, afin qu'il puisse être identifié sans ambiguïté.

L'objet d'une évaluation, qu'il s'agisse d'une mesure, d'un programme ou d'une organisation, doit être examiné avec précision. On notera que l'objet d'évaluation peut prendre des formes différentes dans des contextes de temps et d'espace divers. La description de l'objet d'évaluation devrait montrer avec clarté et précision ce qui a été examiné. Cela permet aussi aux destinataires de faire des comparaisons avec d'autres objets d'évaluation. Par ailleurs, un examen précis de l'objet d'évaluation aide à distinguer les rapports entre celui-ci et ses effets et à repérer des effets secondaires qui n'avaient pas été remarqués jusque-là. A ce propos, il convient de considérer avec attention dans quelle mesure des divergences se manifestent entre l'objet d'évaluation prévue au départ et la forme qu'il a prise dans la phase d'application.

P2 Analyse du contexte

Les influences du contexte sur l'objet d'évaluation sont identifiées.

Nous entendons par contexte de l'objet d'évaluation la combinaison de toutes les conditions cadres qui entourent l'objet à évaluer, comme par exemple l'intégration institutionnelle, le climat social et politique, les caractéristiques personnelles des parties prenantes et groupes concernés, la structure de l'arène politique, les activités étatiques (mais aussi privées) voisines (et concurrentes) ou les conditions cadres économiques. Ces facteurs contextuels et d'autres, doivent être examinés de manière suffisamment précise pour en tenir compte tout au long de la planification, de l'exécution et de la diffusion de l'évaluation. Cette connaissance est nécessaire pour qu'on puisse formuler l'évaluation de manière réaliste par rapport aux possibilités existantes et à son contexte, car les facteurs contextuels exercent souvent une influence déterminante sur les effets d'un objet d'évaluation. Une analyse fondée du contexte permet aussi d'apprécier les résultats d'une évaluation sous l'angle de leur transfert possible dans d'autres contextes. Il faut éviter de définir le contexte de manière trop étroite. Mais, d'autre part, il faut se garder aussi

d'effectuer une analyse du contexte si fouillée que l'analyse de l'objet proprement dit ne peut plus être faite de manière suffisamment approfondie.

P3 Description des objectifs et de la démarche

Les objectifs, les questions et la démarche de l'évaluation sont documentés avec suffisamment de précision pour qu'on puisse les identifier et les apprécier.

Les objectifs visés par une évaluation, les questions à traiter et la procédure choisie doivent être documentés avec précision au cours du processus d'évaluation et communiqués de manière reconstituable pour les destinataires de l'évaluation, lors de la présentation du rapport. Ce standard vise à donner une forme transparente au processus d'évaluation. Dans la description des objectifs et des questions de l'évaluation, on veillera spécialement à prendre en compte aussi les manières de voir divergentes. La documentation et la description de la démarche comprennent la description détaillée de l'organisation, de la collecte, de la mise au point, de la saisie et de l'exploitation des données et de la présentation du rapport. A propos de la procédure suivie, il faut noter que celle-ci peut changer au cours du processus d'évaluation et que, de ce fait, il peut y avoir des différences entre la démarche prévue et le déroulement effectif des opérations. Il importe de présenter clairement vis-à-vis de l'extérieur les écarts qui se sont produits et pour quelles raisons. Ces écarts doivent être motivés explicitement, de même que la démarche en général. En renonçant à exposer les objectifs, les questions et la démarche, on peut immuniser à tort une évaluation contre la critique justifiée, mais aussi laisser surgir des objections hors de propos.

P4 Sources d'information sûres

Les sources d'information utilisées dans une évaluation sont décrites avec suffisamment de précision pour qu'on puisse apprécier la justesse des informations.

La description des sources d'information utilisées permet aux parties prenantes et groupes concernés de se faire une opinion personnelle sur la qualité des informations issues de ces sources. Les sources d'information des évaluations peuvent être des individus ou des groupes, des documents, des matériels audiovisuels, des statistiques, etc. Le recours à différentes sources d'information permet une comparaison des informations obtenues. La crédibilité d'une évaluation peut être mise en question par l'absence ou l'insuffisance de la description des sources d'information. Outre la description des sources d'information, les informations acquises de ces sources devraient aussi être qualifiées. Le degré de sûreté des informations doit être pris en considération lors de l'interprétation des résultats.

P5 Informations valides et fiables

La procédure de récolte d'informations est choisie ou développée puis appliquée de manière à garantir la validité et la fiabilité des interprétations dans la perspective de l'objectif donné.

La saisie empirique d'un objet est exposée jusqu'à un certain point à des sources d'erreur. La validité et la fiabilité sont deux qualités nécessaires du processus de saisie. Ces qualités ne peuvent être appréciées que dans le contexte concret de l'évaluation et dans la perspective de l'objectif spécifique de la saisie empirique.

On entend par validité le caractère valable de la saisie, c'est-à-dire la mesure dans laquelle les instruments engagés saisissent effectivement l'aspect intéressant qu'il s'agit d'examiner. La fiabilité d'un instrument d'enquête, à son tour, est fonction de la cohérence de la saisie par différents instruments et personnes, à différents moments, c'est-à-dire qu'elle dépend de la fiabilité et de la stabilité de la saisie empirique. La validité et la fiabilité d'une approche sont étroitement liées.

Dans le choix et l'évaluation des instruments d'enquête, il convient de prendre en considération ces deux aspects de manière égale.

P6 Vérification systématique des informations

Les informations recueillies, compilées et présentées dans une évaluation sont systématiquement contrôlées afin de détecter d'éventuelles erreurs.

Dans la collecte, la saisie, l'exploitation et l'interprétation des informations, il y a de multiples possibilités de commettre des erreurs. Il peut s'agir de simples fautes de frappe dans la saisie des données, mais aussi de mauvaises interprétations des données recueillies. Pour cette raison, il est impératif de veiller, dans le processus d'évaluation, à éviter autant que possible les sources d'erreurs potentielles. D'autre part, les informations recueillies doivent être contrôlées sous l'aspect des erreurs possibles, par des méthodes appropriées (test de plausibilité, saisie parallèle, validation communicative, etc.). Les erreurs possibles et leurs conséquences doivent être discutées de manière transparente dans le cadre de la présentation du rapport d'évaluation. Si on ne parvient pas à éviter les erreurs, il peut en résulter des interprétations et des conclusions erronées. En outre, des indications incorrectes - même matériellement insignifiantes - peuvent discrediter l'ensemble de l'évaluation.

P7 Analyse des informations qualitatives et quantitatives

Les informations qualitatives et quantitatives d'une évaluation sont analysées de manière appropriée et systématique, afin de répondre effectivement aux questions de l'évaluation.

Dans le cadre des évaluations, on tire des conclusions sur la base d'analyses d'informations qualitatives et quantitatives. L'exploitation des informations recueillies devrait se dérouler de manière systématique, conformément aux règles de la méthodologie qualitative et quantitative. La plupart du temps, il est judicieux et approprié d'inclure dans les évaluations des informations aussi bien qualitatives que quantitatives. Pour le choix des informations comme pour celui de la méthode d'analyse, la question posée et la situation des données sont déterminantes. Ces décisions ne devraient pas être influencées par d'autres facteurs tels que les préférences des personnes intéressées. Les décisions sur le choix des méthodes doivent être communiquées de manière transparente et leurs conséquences soumises à un examen critique. Enfin, on exposera explicitement la valeur des méthodes, de même que leurs limites.

P8 Conclusions motivées

Les conclusions tirées dans une évaluation sont expressément motivées, de manière que les parties prenantes et groupes concernés puissent les reconstituer et les apprécier.

Les conclusions d'une évaluation doivent être explicitement motivées et présentées en indiquant de manière transparente les hypothèses qui leur servent de base et la procédure d'évaluation utilisée. On mentionnera aussi clairement le champ de validité des conclusions. De même, il devrait y avoir dans le cadre du rapport une discussion d'autres interprétations possibles, expliquant les raisons pour lesquelles ces interprétations ont été rejetées. En définissant les prémisses servant de base aux conclusions, on veillera à éviter des hypothèses qui ne sont pas partagées par les protagonistes directs. Le respect de ce standard permet aux utilisatrices et utilisateurs des résultats de l'évaluation d'apprécier le caractère probant des conclusions et accroît ainsi la force de persuasion des conclusions.

P9 Présentation impartiale du rapport

La présentation du rapport doit être à l'abri des distorsions suscitées par les sentiments personnels et les préférences de toute partie intéressée à l'évaluation, afin que le rapport d'évaluation communique honnêtement les résultats obtenus.

L'environnement d'une évaluation est caractérisé par des perspectives multiples. Les parties prenantes et groupes concernés («stakeholders») d'une évaluation ont souvent des visions divergentes de l'objet d'évaluation. L'évaluation court alors le risque d'être détournée par un groupe déterminé, ou d'en devenir l'instrument. Une évaluation ne devrait jamais refléter une seule manière de voir spécifique, mais viser à représenter équitablement tous les intérêts en jeu. C'est pourquoi il faut veiller à ce que l'évaluation repose sur une base aussi indépendante que possible. Ainsi, il faut éviter un rapport trop étroit avec les mandantes ou mandants, mais aussi par exemple avec les personnes responsables de l'objet d'évaluation. Le rapport entre les responsables de l'évaluation et les mandantes ou mandants, ou d'autres groupes intéressés, doit être éclairci dès le début du processus d'évaluation. Cet éclaircissement des rôles implique également qu'on précise les conditions dans lesquelles s'exerce le droit de publier des rapports d'évaluation.

P10 Méta-évaluation

L'évaluation elle-même est évaluée, sur la base des présents standards ou d'autres standards importants, afin qu'elle soit réalisée en les respectant et que les parties prenantes et groupes concernés puissent, au terme de l'évaluation, en apprécier les forces et les faiblesses.

Les évaluations non réussies peuvent susciter de mauvaises décisions. Les évaluations peuvent être exposées (à tort) à de violentes critiques. Pour éviter de telles situations, l'évaluation elle-même devrait être contrôlée sous l'angle de sa qualité. Dans le cadre d'une méta-évaluation (évaluation d'une évaluation), on peut par exemple utiliser les standards définis ici pour apprécier une évaluation. Selon les circonstances, une méta-évaluation peut être conçue comme récapitulative (établissant un bilan) ou endoformative (base d'amélioration). Elle peut être effectuée par les soins de l'équipe d'évaluation (donc dans le cadre interne), ou par des personnes extérieures (cadre externe). Comme dans le cas des évaluations, une méta-évaluation peut être plus ou moins étendue. Tandis qu'une méta-évaluation approfondie et complète n'est indiquée que dans certains cas, il ne faut jamais renoncer à une simple auto-évaluation de l'évaluation par les personnes participant à l'évaluation. Les ressources (généralement modestes) nécessaires à la méta-évaluation doivent être prévues dans la planification de l'évaluation.

DOCUMENTS D'ACCOMPAGNEMENT

Aperçu fonctionnel des standards SEVAL

L'aperçu ci-après montre quels sont les standards auxquelles il convient de prêter une attention particulière lors des principales activités d'évaluation.

Décision sur l'exécution d'une évaluation

- U1 Identification des parties prenantes et groupes concernés
- U2 Clarification des buts d'évaluation
- U3 Crédibilité
- U8 Effet de l'évaluation
- F2 Acceptabilité politique
- F3 Rapport coûts/avantages
- D1 Convention formelle
- D6 Identification des conflits d'intérêts
- P1 Documentation de l'objet d'évaluation
- P2 Analyse du contexte
- P10 Méta-évaluation

Définition du problème d'évaluation

- U1 Identification des parties prenantes et groupes concernés
- U2 Clarification des buts d'évaluation
- P1 Documentation de l'objet d'évaluation
- P2 Analyse du contexte
- P3 Description des objectifs et de la démarche
- P10 Méta-évaluation

Planification de l'évaluation

- U1 Identification des parties prenantes et groupes concernés
- U2 Clarification des buts d'évaluation
- U4 Ampleur et choix des informations
- U5 Transparence de l'appréciation
- F1 Procédure pratique
- D1 Convention formelle
- D4 Appréciation complète et honnête
- P1 Documentation de l'objet d'évaluation
- P3 Description des objectifs et de la démarche
- P4 Sources d'information sûres
- P5 Informations valides et fiables
- P7 Analyse des informations qualitatives et quantitatives
- P8 Conclusions motivées
- P9 Présentation impartiale du rapport
- P10 Méta-évaluation

Collecte des informations

- U3 Crédibilité
- U4 Ampleur et choix des informations
- U5 Transparence de l'appréciation

- F1** Procédure pratique
- F2** Acceptabilité politique
- D1** Convention formelle
- D2** Protection des droits individuels
- D3** Interaction empreinte d'humanité
- D4** Appréciation complète et honnête
- P1** Documentation de l'objet d'évaluation
- P2** Analyse du contexte
- P3** Description des objectifs et de la démarche
- P4** Sources d'information sûres
- P5** Informations valides et fiables
- P6** Vérification systématique des informations
- P10** Méta-évaluation

Exploitation des informations

- U5** Transparence de l'appréciation
- F1** Procédure pratique
- P1** Documentation de l'objet d'évaluation
- P2** Analyse du contexte
- P7** Analyse des informations qualitatives et quantitatives
- P8** Conclusions motivées
- P10** Méta-évaluation

Présentation du rapport d'évaluation

- U1** Identification des parties prenantes et groupes concernés
- U4** Ampleur et choix des informations
- U5** Transparence de l'appréciation
- U6** Caractère complet et clair du rapport
- U7** Présentation du rapport en temps utile
- U8** Effet de l'évaluation
- D2** Protection des droits individuels
- D4** Appréciation complète et honnête
- D5** Publication des résultats
- P1** Documentation de l'objet d'évaluation
- P2** Analyse du contexte
- P3** Description des objectifs et de la démarche
- P4** Sources d'information sûres
- P8** Conclusions motivées
- P9** Présentation impartiale du rapport
- P10** Méta-évaluation

Etablissement du budget de l'évaluation

- U2** Clarification des buts d'évaluation
- U4** Ampleur et choix des informations
- F3** Rapport coûts/avantages
- D1** Convention formelle
- P1** Documentation de l'objet d'évaluation
- P3** Description des objectifs et de la démarche
- P10** Méta-évaluation

Conclusion d'un contrat d'évaluation

- U1 Identification des parties prenantes et groupes concernés
- U2 Clarification des buts d'évaluation
- U3 Crédibilité
- U4 Ampleur et choix des informations
- U7 Présentation du rapport en temps utile
- F2 Acceptabilité politique
- D1 Convention formelle
- D2 Protection des droits individuels
- D5 Publication des résultats
- D6 Identification des conflits d'intérêts
- P1 Documentation de l'objet d'évaluation
- P3 Description des objectifs et de la démarche
- P10 Méta-évaluation

Conduite / gestion de l'évaluation

- U1 Identification des parties prenantes et groupes concernés
- U2 Clarification des buts d'évaluation
- U3 Crédibilité
- U7 Présentation du rapport en temps utile
- F2 Acceptabilité politique
- F3 Rapport coûts/avantages
- D1 Convention formelle
- D2 Protection des droits individuels
- D3 Interaction empreinte d'humanité
- D6 Identification des conflits d'intérêts
- P3 Description des objectifs et de la démarche
- P6 Vérification systématique des informations
- P10 Méta-évaluation

Personnel engagé dans l'évaluation

- U3 Crédibilité
- F2 Acceptabilité politique
- D6 Identification des conflits d'intérêts
- P9 Présentation impartiale du rapport
- P10 Méta-évaluation

Commentaires explicatifs sur la procédure suivie dans la formulation des standards SEVAL

Les standards SEVAL s'inspirent des «Program Evaluation Standards» du «Joint Committee on Standards for Educational Evaluation» (Joint Committee 1994), dans la version traduite en allemand par Wolfgang Beywl et Thomas Widmer (Joint Committee 2000). Les membres du groupe de travail «Standards d'évaluation» de la SEVAL ont été invités à donner leur avis sur cette version des standards. Sur la base des prises de position enregistrées, un comité du groupe de travail a effectué un travail de remaniement. Le texte remanié a été à nouveau discuté dans le cadre du groupe de travail. Le comité a adapté le texte en fonction des résultats de la discussion. Cette version a été soumise à la discussion lors du congrès intitulé « Assurer la qualité en matière d'évaluation » tenu à Berne le 26 mai 2000, auquel ont participé environ quatre-vingt personnes. C'est sur la base des expériences ainsi acquises que le groupe de travail et le comité ont encore une fois remanié le texte. La présente version est le résultat de ces travaux.

Parmi les changements effectués, il convient de signaler que, de manière générale, les termes «programme» et «évaluation de programme» ont été remplacés par «objet d'évaluation» et «évaluation». Cela n'a pas de conséquences sur le fond, mais répond au souci d'éviter que le terme «programme» tel que l'entend l'original américain ne prête à confusion. Dans un but de simplification linguistique, les affirmations construites avec le verbe «devoir» (soll) sont maintenant énoncées avec le verbe «être» (ist). Outre un remaniement linguistique général, nous mentionnerons encore les changements suivants (désignations abrégées de Joint Committee 2000):

- réunion des standards P5 et P6 et des standards P8 et P9,
- reformulation des standards U7, F1, F3, D3, D4, D6, D7, P2, P7, P10 et P12,
- suppression sans remplacement des standards D1 et D8.
- complément d'un nouveau standard : « clarification des buts d'évaluation » (nouveau : U2)

Les différents changements sont expliqués et motivés ci-après.

Standards réunis

Les *standards P5 et P6* traitent de deux critères de qualité du travail des sciences sociales, la validité et la fiabilité. Etant donné que l'appréciation séparée de ces aspects n'a guère de sens en raison de leur interdépendance (cf. à ce propos Widmer 1996: 296), ces deux standards ont été réunis en un seul.

Compte tenu du fait qu'en règle générale les évaluations devraient s'appuyer sur des informations tant qualitatives que quantitatives, les critères *P8 sur l'analyse des informations quantitatives et P9 sur l'analyse des informations qualitatives* ont été réunis en un seul standard.

Standards reformulés

U7 Effet de l'évaluation: une autre formulation a été choisie dans un but de simplification linguistique, sans toutefois exercer un effet sur le contenu.

F1 Procédure pratique: là aussi, il s'agit d'une simplification purement linguistique.

F3 Rapport coûts/avantages: le standard a été raccourci en supprimant l'exigence d'efficacité, car celle-ci est déjà contenue dans le reste du standard et par conséquent superflue.

D3 Protection des droits individuels: le titre original de ce standard (protection des droits individuels de la personne humaine) ne correspond pas entièrement au contenu du standard, car celle-ci va au-delà des droits de la personne humaine au sens strict.

D4 Interaction empreinte d'humanité: le standard a été reformulé, car la formulation originale demandant qu'on respecte «la dignité et la valeur des êtres humains», qui faisait intervenir une asymétrie (entre évaluatrices et évaluateurs et autres personnes intéressées), n'était pas convaincante.

D6 Publication des résultats: l'insertion de «y compris les restrictions importantes» qui figurait à l'origine après «résultats de l'évaluation» a été supprimée, étant donné qu'on ne comprend pas très clairement de quelles restrictions il s'agit et que ce point est déjà traité dans d'autres standards (voir U4, U5, D5 et P4).

D7 Identification des conflits d'intérêts: étant donné que les préjudices suscités par les conflits d'intérêts ne peuvent être toujours entièrement évités, la formulation absolue «qu'ils ne portent pas atteinte» a été relativisée par l'expression «qu'ils portent aussi peu atteinte que possible».

P2 Analyse du contexte: le standard a été formulé plus simplement, sans qu'il en résulte de changement dans le contenu.

P7 Vérification systématique des informations: on a renoncé à la phrase complémentaire demandant que toutes les erreurs détectées soient corrigées, car la mention explicite du besoin de correction a été jugée inutile.

P10 Conclusions motivées: le standard a été complété par l'exigence de la possibilité de reconstitution qui constitue une condition préalable importante à la formulation d'une appréciation solide.

P12 Méta-évaluation: ce standard a été légèrement adapté sous deux aspects. Premièrement, dans le souci d'une meilleure compréhension, on a renoncé aux termes techniques «récapitulatif» et «endoformatif». Deuxièmement, la formulation «examiner minutieusement» a été remplacée par «apprécier», de manière à permettre, avant tout pour les projets d'évaluation de petites dimensions, une méta-évaluation appropriée.

Standards biffés sans remplacement

D1 Soutien à l'orientation sur les prestations de services

L'évaluation est planifiée de manière à aider les organisations à tenir compte des intérêts et des besoins de tout l'éventail des groupes cibles et à orienter leurs activités dans ce sens.

Motifs de la suppression

L'un des principes des standards devrait être qu'ils sont valables autant que possible pour tous les champs d'application prévus (tels qu'ils sont définis dans l'introduction). Or tous les objets d'évaluation ne sont pas des organisations prestataires de services (p.ex. évaluations dans le secteur de l'économie privée). Il n'est pas non plus évident d'emblée que la prise en compte des besoins de *tout* l'éventail des groupes cibles soit judicieuse. Il peut y avoir des cas où il est important de faire précisément l'inverse en se concentrant sur un segment déterminé.

D8 Responsabilité financière

La gestion des ressources (affectation et dépenses) par l'évaluatrice ou l'évaluateur est attestée par une comptabilité soigneusement tenue, rationnelle et conforme à l'éthique, afin que les dépenses témoignent de la conscience des responsabilités et apparaissent appropriées.

Motifs de la suppression

- Le standard mélange des préoccupations diverses, même si elles sont parfois connexes (comptabilité soigneusement tenue, affectation éthiquement responsable des ressources, documentation des dépenses, gestion «dans tous les cas» avisée des ressources (affectation et dépenses), dépenses effectuées dans la conscience des responsabilités et de manière appropriée...).
- L'essentiel du standard D8 est déjà contenu dans le standard F3. Dès le moment de la conclusion du contrat, la relation entre l'évaluatrice ou l'évaluateur et la mandante ou le mandant constitue une relation commerciale normale, réglée dans le contrat et dans le CO. Les événements qui se déroulent avant la conclusion du contrat sont/seraient beaucoup plus importants (pratique des appels d'offres par les pouvoirs publics et les organisations privées). Dans cette perspective, les intérêts des mandantes et mandants seraient aussi mieux protégés par des formulations correspondantes dans les règlements des soumissions.
- Dans le secteur des services, un très grand nombre de mandats prévoient le paiement de montants forfaitaires fondés sur une liste de prestations. Dès le moment où un contrat est signé, le droit de regard dans la comptabilité s'éteint. De ce point de vue, le standard est peu réaliste et naïf.
- Le standard oublie le principe de la symétrie au sens où on ne trouve nulle part dans les standards le devoir, pour les mandantes ou mandants, de mettre à disposition les ressources nécessaires pour des évaluations «éthiquement responsables» et des dépenses «responsables et appropriées».

Nouveau standard**U2 Clarification des buts d'évaluation**

Toute personne prenant part à une évaluation se préoccupe de clarifier les buts des évaluations pour chaque participant et chaque personne concernée.

Ce standard a été ajouté aux standards SEVAL pour combler une sérieuse lacune perçue par le groupe de travail au niveau des standards du Joint Committee. Malgré le fait que les standards abordent à plusieurs endroits le thème des buts de l'évaluation aussi bien de manière implicite qu'explicite (voir, par exemple, U1 ou P3), le comité de travail est d'avis qu'il est justifié d'introduire un nouveau standard explicitement à ce propos, étant donnée la grande importance accordée à la clarification des buts d'évaluation.

Tableau de transformation pour les désignations abrégées

A cause des changements mentionnés avant, il en résultent, pour les désignations abrégées, les transformations les suivantes entre les standards SEVAL et les «Program Evaluation Standards» (dans la version allemande; voir Joint Committee 2000):

<i>JC</i>	<i>SEVAL</i>	<i>JC</i>	<i>SEVAL</i>	<i>JC</i>	<i>SEVAL</i>	<i>JC</i>	<i>SEVAL</i>
N1	U1	D1	F1	K6	D5	G7	P6
absent	U2	D2	F2	K7	D6	G8/G9	P7
N2	U3	D3	F3	K8	supprimé	G10	P8
N3	U4	K1	supprimé	G1	P1	G11	P9
N4	U5	K2	D1	G2	P2	G12	P10
N5	U6	K3	D2	G3	P3		
N6	U7	K4	D3	G4	P4		
N7	U8	K5	D4	G5/G6	P5		

Membres du groupe de travail «Standards d'évaluation» de la SEVAL

Dr. Paul Ammann, Gesundheits- und Fürsorgedirektion des Kantons Bern, Bern
 Dr. Nicole Bachmann, Gesundheits- und Fürsorgedirektion des Kantons Bern, Bern*
 Renate Cléménçon, Bundesamt für Justiz, Bern
 Valérie Eggimann-Lanz, Eidgenössische Finanzkontrolle (EFK), Bern
 Frohmüt W. Gerheuser, Büro für Politikberatung und Sozialforschung, Brugg (AG)
 Dr. Armin Kühne, Eidgenössische Finanzkontrolle (EFK), Bern
 Charles Landert, Landert Farago Davatz & Partner, Zürich*
 Dr. Marlène Läubli Loud, Office fédéral de la santé publique, Berne
 Dr. Tobias Ledergerber, Organisationsamt des Kantons Bern, Bern
 Dr. Urs Moser, Kompetenzzentrum Bildungsevaluation & Leistungsmessung, Universität Zürich
 Urs Schmidt, e&e Entwicklung & Evaluation im Sozialwesen, Zürich
 Prof. Dr. Christian Simon, Schweizerischer Wissenschafts- und Technologierat (SWTR), Bern
 Dr. Heini Sommer, Ecoplan, Altdorf (UR)
 Patrick Trees, Staatskanzlei des Kantons Bern, Bern
 Dr. Thomas Widmer, Institut für Politikwissenschaft, Universität Zürich, Zürich (président)*

*: membre du comité du groupe de travail

Responsables de la traduction française:

Françoise Nagy, Prangins (VD)
 Celina Ramjoué, Institut für Politikwissenschaft, Universität Zürich, Zürich
 Emmanuel Sangra, Commission externe d'évaluation des politiques publiques, Genève
 Dr Frédéric Varone, Département des sciences politiques et sociales, Université catholique de Louvain, Louvain-la-Neuve (Belgique)

Bibliographie

Joint Committee on Standards for Educational Evaluation (Ed.) (1994): The Program Evaluation Standards, 2nd Edition. Newbury Park: Sage.

Joint Committee on Standards for Educational Evaluation (Hrsg.) (2000): Handbuch der Evaluationsstandards. Die Standards des Joint Committee on Standards for Educational Evaluation. 2. Auflage. Opladen: Leske + Budrich.

Widmer, Thomas (1996): Meta-Evaluation. Kriterien zur Bewertung von Evaluationen. Bern: Haupt.

Widmer, Thomas (2000): « Qualität der Evaluation – Wenn Wissenschaft zur praktischen Kunst wird. » in Stockmann, Reinhard (Ed.): Evaluationsforschung. Opladen : Leske + Budrich : 77-102.