



Communication

LE BON ENSEIGNEMENT DE L'HISTOIRE: FONDEMENTS, RESULTATS DES REFLEXIONS ET RECOMMANDATIONS

Peter GAUTSCHI



LE BON ENSEIGNEMENT DE L'HISTOIRE: FONDEMENTS, RESULTATS DES REFLEXIONS ET RECOMMANDATIONS

Peter GAUTSCHI

En Suisse, il existe un enseignement scolaire de l'histoire au quotidien qui est réputé «bon». C'est là une constatation positive, en dépit de toutes les crises qui sont censées affecter l'enseignement de l'histoire.

Les historiens américains Paxton/Wineburg (2000, p. 855) introduisent par exemple leur exposé sur la didactique de l'histoire par un extrait de film sur l'enseignement de l'histoire, pour évoquer une image généralement répandue dans le grand public: des questions absurdes sur des détails du passé sans aucune importance, une traversée insensée à travers des faits tombés dans l'oubli. Wilson (2002, p. 530) décrit également l'enseignement de l'histoire comme une manifestation sans grande ambition intellectuelle, où se fait rare une pensée critique tendant à résoudre des problèmes.

Les conclusions des chercheurs de la zone alémanique sont tout aussi désabusées. Déjà en 1976, Klaus Bergmann avait précisé: «Il ressort de toutes les études empiriques que la plupart des élèves trouve l'histoire ennuyeuse, dépourvue d'intérêt, superflue - une matière qu'ils n'aiment pas» (Bergmann 1976, p. 3). Bodo von Borries (2004e) constate quelques années plus tard de graves *insuffisances* des jeunes dans leur approche de l'histoire.

Dans le présent exposé, j'essaie de démontrer qu'il existe un «bon enseignement de l'histoire». L'étude est focalisée sur ce qui «*a réussi*». Lors de mes nombreuses visites dans les écoles, j'ai pu voir des élèves intéressés, des enseignants engagés, un bon enseignement de l'histoire. Ces expériences contrastent avec les observations ci-dessus, faites par des chercheurs et des inspecteurs d'écoles et tendant à démontrer l'échec apparent de l'enseignement de l'histoire. L'étude présentée ici ne tend pas à ajouter des critiques adressées aux enseignants, mais à apporter les preuves d'un bon enseignement courant de l'histoire.

Pourtant, il me paraissait évident qu'il suffirait de se référer à la documentation sur l'échec de l'enseignement de l'histoire ou faire une comparaison entre les mauvaises et les bonnes leçons pour en tirer des conclusions, d'autant plus que la présente étude a ouvert l'accès à des travaux de recherche et qu'il aurait été possible de caractériser les leçons qui sont qualifiées «d'échecs» sous différents aspects. Toutefois, c'est parce que

les enregistrements vidéo de l'enseignement permettent d'identifier les participants, qu'il m'a paru nécessaire de me concentrer sur ce qui est «bon».

La présente contribution est répartie en quatre chapitres:

- en premier lieu, une présentation des interrogations par cette étude et de quelques réflexions fondamentales sur la théorie du bon enseignement de l'histoire (*fondements théoriques*),
- deuxièmement, une présentation du *concept de recherche*,
- troisièmement, une discussion sur certains *résultats sélectionnés*
- enfin, *trois questions* concernant la démarche future.

1. Interrogations faisant l'objet des recherches et fondements théoriques

Les objectifs de l'étude ont visé l'identification, la description et l'analyse d'un bon enseignement de l'histoire. C'est dans ce but que l'étude s'est référée aux travaux «L'enseignement de l'histoire aujourd'hui» (Gautschi/Moser/Reusser/Wiher, 2007) et a utilisé *trois sources de données* provenant des matériaux de recherche du projet «L'histoire et la politique dans l'enseignement»: premièrement, les enregistrements vidéos des leçons d'histoire, deuxièmement, les appréciations des élèves et troisièmement les appréciations des enseignants sur les enregistrements vidéos des leçons d'histoire. De plus, une *quatrième source de données* a été exploitée: les jugements qu'ont portés des experts externes sur ces enregistrements vidéo de leçons d'histoire.

Ce qui permet de répondre entre autres à deux questions majeures:

1. Existe-t-il dans l'enseignement au quotidien de l'histoire des leçons qui, dans la perception des enseignants, des élèves et des experts externes, sont jugées «bonnes»?
2. Quelles sont les caractéristiques et les facteurs clés des leçons d'histoire reconnues «bonnes»?

La *question de la qualité* se trouve donc au centre de la présente étude: ce discours n'est pas nouveau, même si jusqu'alors il a été tenu d'une manière **non** spécifique et sans tenir compte de la matière elle-même. Ce n'est que depuis peu de temps qu'on tend à réfléchir sur la qualité de l'enseignement au regard de la matière et du contenu.

L'étude faisant l'objet de cet exposé et qui est focalisée sur les processus d'enseignement, se réfère à un modèle d'offre et d'utilisation. A la différence de nombreuses autres études, la qualité n'est pas définie en fonction du produit. En effet, le but est d'identifier et de décrire la qualité dans le processus d'enseignement lui-même et d'analyser ensuite le processus d'enseignement. Ce n'est pas une référence externe comme par exemple le résultat ('outcome') qui est pertinente pour valider un bon enseignement de l'histoire, mais la validité d'apparence ou «*face validity*» obtenue par le biais d'une triangulation.

Les modèles visant l'offre et l'utilisation permettent de comprendre que *l'enseignement proprement dit* est marqué par des *conditions préalables* et qu'il produit des *conséquences*. Les conditions préalables et les conséquences se distinguent en fonction

des différents échelons, à savoir «*la société*», «*les individus*» concernés et «*l'univers historique*». Les conséquences de l'enseignement peuvent se produire à trois niveaux: premièrement par un enrichissement de l'univers historique par des nouvelles représentations; deuxièmement, par un nouveau savoir et de nouveaux acquis qui permettront aux individus d'adopter des attitudes et des comportements nouveaux; troisièmement en donnant une nouvelle dimension à la culture historique.

Le processus d'enseignement proprement dit porte sur trois domaines: l'offre comprend «*la matière d'enseignement* » et la «*structure du processus* » - en quelque sorte il s'agit du «*quoi*» et «*comment*» de l'enseignement de l'histoire -, ces deux domaines permettant «*l'utilisation*» à celui qui apprend. Ces trois domaines, le «*sujet de l'enseignement*», la «*structure du processus*» et l'«*utilisation*», sont présentés dans cette étude selon les critères qualitatifs énumérés ci-après et orientés sur la théorie :

En ce qui concerne la matière d'enseignement :

1. un thème important
2. réflexions sur le comportement de l'homme dans un environnement social
3. réflexions sur les changements dans le temps et des liens entre différentes évolutions
4. l'exactitude factuelle, perspectives multiples, controverses
5. les liens entre le sujet et la situation de l'élève
6. représentation de l'histoire à titre d'exemple et adaptée au groupe cible

En ce qui concerne la structure du processus :

7. assurer l'efficacité de l'enseignement en classe et de l'emploi du temps
8. développer un climat favorable dans la classe
9. garantir des opportunités d'un apprentissage stimulant, vivant et approprié
10. veiller à la clarté et à une bonne structuration

En ce qui concerne l'utilisation:

11. perception des témoignages historiques et des changements dans le temps
12. examen, vérification et présentation d'analyses factuelles historiques
13. interprétation (analyse et explication) de l'histoire
14. orientation, symbolisation de l'expérience du temps passé et examen visant un jugement de valeur sur l'expérience du passé
15. Appropriation et restitution du savoir du passé et compréhension de l'histoire

S'agissant de la matière enseignée, les critères de qualité résultent du système catégoriel de Mayer/Pandel, de celui de Barricelli/Sauer, ce en tenant compte des «National Standards for History» du «National Center for History in the Schools».

S'agissant de la structure du processus, les critères de qualité résultent d'une comparaison des listes sur le bon enseignement établis par Brophy (2000), Meyer (2004), Helmke (2006) et sont en priorité basés sur 'la Structure des critères' de Clausen/Reusser/Klieme (2003).

S'agissant de l'utilisation, les critères de qualité résultent du modèle de compétence qui a été développé pour cette étude.

Sur le plan théorique, on obtient au total 15 critères de qualité pour l'enseignement de l'histoire dans trois domaines. A l'aide de ces critères de qualité sélectionnés, il devient possible de définir théoriquement ce qui est un bon enseignement de l'histoire:

L'enseignement de l'histoire est bon lorsque les élèves acquièrent des compétences fondamentales permettant de s'approprier des connaissances historiques et à apprendre l'histoire par le biais de contenus et de thèmes significatifs en la matière, et ce grâce à un processus d'enseignement adapté aux exigences des sciences concernées.

2. Concept de recherche

La qualité de l'enseignement de l'histoire ne pouvant pas se mesurer ni être déterminée directement, il convient donc de se pencher sur les jugements exprimés par les élèves, les enseignants et les experts. Dans la présente étude, cet examen se fait au moyen d'une triangulation des trois points de vue (à savoir de l'élève, de l'enseignant et des experts).

L'objet de recherche de l'étude «Un bon enseignement de l'histoire: fondements, résultats des réflexions, recommandations» porte sur 39 leçons d'histoire au quotidien qui dans le cadre du projet «L'histoire et la politique dans l'enseignement» ont été sélectionnées par échantillonnage dans les cantons suisses d'Argovie, de Berne et de Zurich.

Ces leçons ont fait l'objet d'enregistrements vidéo (1ère étape).

Après les leçons, les élèves et les enseignants ont été interrogés (2^e et 3^e étapes).

Les vidéos ont en plus été documentées grâce à un dépouillement général selon Legewie (1994) (4^e étape).

En parallèle, les experts sélectionnés à l'aide d'un Convenience Sampling ont regardé les enregistrements vidéo de l'enseignement et ont rempli un questionnaire (5^e étape).

Cette base de données a permis de calculer pour chacun des angles de vue un jeu de leçons qui peuvent être qualifiées comme étant «bonnes» sous l'angle de chaque groupe (6^e – 9^e étapes).

Je voudrais éclairer cette démarche en me penchant à titre d'exemple sur les élèves:

Dans une *première étape*, il convient de retenir dans le questionnaire les items qui d'un point de vue didactique sont des indicateurs de la qualité de l'enseignement vécu. Les questions sélectionnées sont attribuées au «processus» et au «résultat» en tenant compte du modèle cadre de l'enseignement de l'histoire.

Dans une *deuxième étape*, il est vérifié au niveau de chacun des élèves et à l'aide d'une analyse factorielle et de fiabilité si l'on peut établir une échelle de qualité unidimensionnelle pour chacun des deux domaines, le «processus» et le «résultat». Pour le domaine «*processus*», on obtient les résultats représentés dans le tableau 1.

	Faktor- ladung
LP4 Lernprozess: In dieser Geschichtsstunde habe ich darauf geachtet, dass ich die Sache wirklich verstehe.	.790
LP1 Lernprozess: In dieser Geschichtsstunde hat mich die Sache so fasziniert, dass ich mich voll einsetzte.	.752
LP5 Lernprozess: In dieser Geschichtsstunde habe ich genau aufgepasst, damit ich das Wichtigste behalte.	.699
LP2 Lernprozess: In dieser Geschichtsstunde wollte ich den Stoff wirklich verstehen.	.681
LP3 Lernprozess: In dieser Geschichtsstunde bin ich die ganze Zeit über gut mitgekommen.	.631
LP6 Lernprozess: In dieser Geschichtsstunde überlegte ich mir, was zu tun ist und wie ich genau vorgehen soll.	.563
LP7 Lernprozess: In dieser Geschichtsstunde habe ich sofort erkannt, was ich tun oder antworten muss.	.511

Tableau 1 les coefficients de chargement obtenus pour le domaine «processus»

Le tableau indique les coefficients de chargement obtenus de manière empirique et qui serviront dans la 3^e étape à calculer l'indice global pondéré des différents élèves. L'item «Lors de cette leçon d'histoire, j'ai fait attention à tout bien comprendre» présente le chargement le plus important avec .79 réponses affirmatives. La variance expliquée est de 44,5%. Pour ces 7 items, l'alpha de Cronbach est de $\alpha = .79$. La création d'une échelle unidimensionnelle avec les items sélectionnés est donc justifiée. Par ailleurs, tous les 7 items peuvent être retenus pour apprécier la qualité; en effet, la suppression d'un seul item entraînerait une baisse de l'alpha de Cronbach.

Dans une *troisième étape*, c'est au niveau de chaque élève qu'un indice global pondéré est calculé pour les deux domaines, et ce à l'aide des coefficients déterminés lors de l'analyse factorielle. Les coefficients de chargement empiriques servent à la pondération. C'est ainsi que les valeurs des différentes échelles sont multipliées et additionnées pour tous les élèves.

Ensuite, dans une *4^e étape*, les deux échelles sont transformées selon le spectre de réponse des items initiaux, à savoir pour le domaine «processus» sur une échelle de 1 à 4 et pour le domaine «résultat» sur une échelle de 1 à 6. Ainsi, une valeur scalaire sur chacune des échelles est attribuée à chaque élève. Grâce à cette transformation, il devient possible d'appliquer la valence des observations de l'item initial au niveau de l'échelle nouvellement créée.

Dans une *cinquième étape*, les valeurs de tous les élèves d'une classe sont additionnées pour obtenir la valeur moyenne d'une classe.

Dans une *sixième étape*, il est procédé à une identification des leçons qui du point de vue de l'élève sont reconnues «bonnes». En fonction de la construction de l'item initial, on considère qu'une leçon est «bonne» si la moyenne arithmétique est supérieure ou égale à 3 dans le domaine «processus» et supérieure ou égale à 4 dans le domaine «résultat». Les leçons qui remplissent les deux conditions sont marquées en gris sur le tableau 2.

Lektionsnummer	Lernprozess	Lernergebnis
	Gewichteter Summenindex der Items LP1-LP7 [Minimum = 1; Maximum = 4]	Gewichteter Summenindex der Items LE1-LE4 [Minimum = 1; Maximum = 6]
1	3.21	4.64
2	2.54	3.87
3	2.60	3.66
4	2.62	3.44
6	2.88	4.25
8	3.02	4.41
9	3.18	4.53
10	2.98	4.26

Tableau 2 L'appréciation par les élèves des leçons enregistrées qui porte sur le processus d'apprentissage et les résultats de l'enseignement (valeurs moyennes d'une classe)

L'extrait du tableau montre qu'il existe des *différences considérables* du point de vue des élèves en ce qui concerne les processus d'apprentissage et les résultats de l'enseignement. Parmi les 10 leçons représentées ici, les élèves estiment que la leçon 1 est la meilleure tant en ce qui concerne les processus d'apprentissage que les résultats de l'enseignement. Les leçons 8 et 9 obtiennent également des valeurs médianes qui sont supérieures à la moyenne au regard des deux dimensions.

Pour la triangulation, ces leçons sont qualifiées comme étant «bonnes» pour la triangulation du point de vue des élèves et sont identifiées par le symbole binaire «1». Le «0» est attribué aux leçons qui ne sont pas jugées «bonnes».

Le calcul de contrôle à l'aide d'un tableau croisé bivarié avec test révèle d'ailleurs qu'entre les deux domaines, à savoir «l'appréciation du processus d'apprentissage par les élèves» et «l'appréciation du résultat d'enseignement par les élèves», il existe une *cohérence statistique certaine*: si les élèves ont une opinion négative du processus d'apprentissage, ils estiment souvent que le résultat de l'enseignement est également négatif; il en est de même pour l'appréciation positive.

Le point de vue des enseignants et des experts a été examiné de la même manière. Ainsi, trois jeux de «bonnes» leçons d'histoire ont pu être identifiés.

Ces trois jeux ont été mis en comparaison à l'aide d'une triangulation des données, ce qui a permis d'identifier 5 leçons parmi les 39 leçons d'histoire qui tant du point de vue des élèves que de celui des enseignants et des experts ont été qualifiées comme étant «bonnes».

Cette identification des 5 leçons a été faite à l'aide d'une triangulation. La triangulation dans la première phase exploratoire de l'étude a servi à la vérification croisée des résultats. C'est une triangulation qui a permis d'identifier parmi toutes les données les leçons qui sont reconnues «bonnes».

Dans la deuxième phase, la triangulation a servi à renforcer les conclusions et a aidé à éclaircir pour quelle raison les leçons sélectionnées ont été qualifiées comme étant «bonnes». Cette démarche a contribué à enrichir les critères de qualité pour l'enseignement de l'histoire au niveau des banques de données.

Dans cette étude, les triangulations ont donc rempli deux fonctions, à savoir la validation des résultats et le renforcement de la crédibilité des conclusions. Tandis que les appréciations des enseignants et des élèves ont été recueillies immédiatement après l'enseignement vécu à l'aide d'un questionnaire comportant des questions fermées et ouvertes, lequel a été dépouillé ensuite en vue d'une validation quantitative et d'un renforcement qualitatif des conclusions, les appréciations des experts ont été recueillies à l'aide de l'enregistrement vidéo de l'enseignement et d'un questionnaire qui a également été dépouillé en vue d'une validation quantitative et d'un renforcement des conclusions qualitatives. C'est ainsi que des méthodes qualitatives et quantitatives sont utilisées dans une approche complémentaire et non pas en tant que «camps rivaux» (Jick 1979, p. 135). Cette démarche reflète les différents efforts tendant à créer un lien entre la recherche qualitative et quantitative.

3. Résultats sélectionnés

Les résultats des triangulations ont servi à réaliser des tests de Cochran et d'équirépartition. Bien que les 39 cas étudiés ne présentent qu'un faible nombre et qu'il convient donc d'interpréter avec prudence les résultats des tests, les résultats sont néanmoins intéressants:

Premièrement, le test de Cochran révèle de manière statistiquement significative que les appréciations des trois groupes (élèves, enseignants, experts) diffèrent.

Deuxièmement, les résultats des triangulations ont été utilisés pour réaliser le test d'équirépartition (Lautsch/von Weber 1995, p. 85-87), qui a permis de comparer la fréquence observée sur la base de l'équirépartition. Le test d'équirépartition révèle de manière statistiquement significative une grande homogénéité des jugements négatifs. En d'autres termes: les élèves, les enseignants et les experts jugent différemment, mais ils sont d'accord quand une leçon d'histoire n'est pas «bonne». Il existe manifestement un consensus quand une leçon d'histoire n'est pas bonne, mais il n'y a pas de consensus

sur la question de savoir ce qu'il convient d'entendre par un bon enseignement de l'histoire.

L'analyse des observations des enseignants, des élèves et des experts portant sur les leçons reconnues bonne révèlent par ailleurs que les critères de qualité déduits de la théorie sont fiables. Dans le modèle cadre présenté, il est possible de localiser les appréciations des experts et les caractérisations des élèves et des enseignants et de leur attribuer des critères de qualité.

D'une manière générale, il ressort de l'analyse comparative des 5 leçons d'histoire reconnues «bonnes» que «la garantie d'une opportunité d'un apprentissage stimulant, vivant et approprié» est un critère clé pour le bon enseignement de l'histoire. Ce critère de qualité et celui visant «l'adaptation du sujet à la situation de l'élève» (critère de qualité 5) sont les deux seuls critères qui ont été appréciés favorablement pour les leçons d'histoire reconnues «bonnes», dont le contenu a d'ailleurs été fort différent.

Après une analyse des 5 leçons, on peut constater en résumé et avec prudence ce qui suit: *Ce sont des tâches d'apprentissage ponctuelles, spécifiques et significatives qui sont déterminantes pour des bonnes leçons d'histoire. Elles doivent comporter un lien avec les élèves et les stimuler à raconter des événements historiques. Le bon enseignement de l'histoire est donc basé sur les tâches, comporte un lien avec les élèves et est orienté vers la compétence.*

Toutefois, les observations des personnes interrogées dans cette étude révèlent également que les conditions individuelles des enseignants et des élèves ont une influence considérable sur le processus d'enseignement et sur sa qualité: c'est la personne de l'enseignant qui est déterminant au niveau de l'offre (le sujet d'apprentissage et la structure du processus), tandis que son utilisation est commandée par l'élève.

4. Questions concernant la démarche future

En ce qui concerne l'importance des conditions individuelles, il est intéressant de se pencher sur les résultats d'un test de réussite, qui a été réalisé dans toutes les classes ayant fait l'objet d'un enregistrement vidéo, et de comparer ces résultats avec l'appréciation des leçons d'histoire. On peut tirer les conclusions suivantes si l'on compare les résultats du test de réussite avec les appréciations sous trois angles différents:

1. Les classes dont les leçons d'histoire sont reconnues «bonnes» selon les trois critères ont également révélé d'excellents résultats lors du test de connaissances par comparaison à d'autres classes du même type d'école.
2. Les réussites de ces classes sont comparables à d'autres classes dont l'enseignement n'a été reconnu «bon» sous aucun des trois angles de vue.
3. Tandis que l'appréciation d'une «bonne leçon d'histoire» permet de toute évidence d'en tirer des conclusions sur la capacité de performance, il ressort de cette étude qu'il n'en est pas de même dans le sens inverse, à savoir que de bonnes performances ne permettent pas de conclure que les leçons d'histoire sont «bonnes»: la meilleure

moyenne d'une classe d'un type d'école dont les exigences sont élevées a été constatée dans une classe dont l'enseignement n'a été reconnu «bon» sous aucun des angles.

Cette comparaison entre des appréciations de l'enseignement et des résultats du test de connaissances permet d'en déduire qu'il y a lieu de faire avancer intensément la recherche en matière de qualité par une approche de *l'analyse de l'offre et de l'utilisation* de l'enseignement - tel que cela a été fait dans la présente étude - et par une approche visant l'impact. Il y a donc lieu de se demander si les résultats des tests de réussite sont un indicateur fiable qui permet d'apprécier la qualité du processus des leçons d'histoire.

Après cette étude, il convient de poser la question de savoir si le modèle de compétence déduit de la théorie et qui s'est révélé constituer un outil approprié dans cette étude, est susceptible d'une validation empirique.

Ce qui me préoccupe le plus, c'est la question de savoir si les leçons d'histoire reconnues «bonnes» donneraient les mêmes résultats avec d'autres enseignants et d'autres élèves. Ou bien, pour poser la question d'une autre manière: est-ce qu'une standardisation des opportunités d'apprentissage permettrait d'améliorer l'acquisition de concepts et de compétences chez les adolescents? C'est la recherche intervention qui pourrait éventuellement répondre à cette question.

La recherche intervention effectuée à la suite de la présente étude a visé en premier lieu la composition, l'expérimentation et l'amélioration des tâches d'apprentissage et l'étude des impacts. Ces travaux ont certainement rempli l'ensemble des cinq *fonctions* de la recherche dans le domaine de la formation des enseignants et de la didactique que Fritz Oser a mentionné dans son discours lorsque le prix de reconnaissance lui a été décerné par la fondation Aebli Näf en mars 2004 en récompense pour ses travaux pionniers. La recherche devrait en premier lieu ériger un mur de protection contre des espoirs irréalistes, deuxièmement reconnaître ou rejeter sans ambiguïté des solutions alternatives à ce qui existe déjà, troisièmement exiger ce qui est nécessaire, quatrièmement «encourager» l'action et la réflexion et cinquièmement procéder à une évaluation de la qualité.

Ces cinq objectifs ont tous été poursuivis par la présente étude. Cette étude n'aura eu un sens que si elle permet de stimuler la curiosité pour l'enseignement quotidien de l'histoire et de renforcer la certitude qu'il existe un bon enseignement de l'histoire et que sa diffusion est possible.

Bibliographie

- Barricelli, Michele / Sauer, Michael (2006): «Was ist guter Geschichtsunterricht?» Fachdidaktische Kategorien zur Beobachtung und Analyse von Geschichtsunterricht. In: Geschichte in Wissenschaft und Unterricht, Jg. 57, Heft 1, S. 4-26.
- Bergmann, Klaus (1976): Warum sollen Schüler Geschichte lernen? In: Geschichtsdidaktik, Jg. 1, Heft 1, S. 3-14.

- Borries, Bodo von (2004) (unter Mitarbeit von Filser, Karl / Pandel, Hans-Jürgen / Schönemann, Bernd): Kerncurriculum Geschichte in der gymnasialen Oberstufe. In: Tenorth, Heinz-Elmar: Kerncurriculum Oberstufe II. Expertisen - im Auftrag der Ständigen Konferenz der Kultusminister. Weinheim: Beltz. S. 236–321.
- Brophy, Jere (2000): Teaching (Educational Practices Series, Vol. 1). Brussels: International Academy of Education & International Bureau of Education. www.ibe.unesco.org/publications/EducationalPracticesSeriesPdf/prac01e.pdf (aufgerufen am 27.12.2007).
- Clausen, Marten (2002): Unterrichtsqualität. Eine Frage der Perspektive? Empirische Analysen zur Übereinstimmung, Konstrukt- und Kriteriumsvalidität. Münster: Waxmann.
- Clausen, Marten/Reusser, Kurt/Klieme, Eckhard (2003): Unterrichtsqualität auf der Basis hoch-inferenter Unterrichtsbeurteilung. In: Unterrichtswissenschaft, Jg. 31, Heft 2, S. 122-141.
- Flick, Uwe (2004): Triangulation. Eine Einführung. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gautschi, Peter / Moser, Daniel V. / Reusser, Kurt u.a. (Hrsg.) (2007): Geschichtsunterricht heute. Eine empirische Analyse ausgewählter Aspekte. Bern: h.e.p. Verlag.
- Helmke, Andreas (2006): Was wissen wir über guten Unterricht? Über die Notwendigkeit einer Rückbesinnung auf den Unterricht als «Kerngeschäft» der Schule. In: Pädagogik, Heft 2/2006, S. 42-45.
- Jick, Thomas (1979): Mixing qualitative and quantitative methods: Triangulation in action. In: Van Maanen, John (Hrsg.): Qualitative Methodology. London: Thousand Oaks. S. 135–148.
- Körper, Andreas / Schreiber, Waltraud / Schöner, Alexander (Hrsg.) (2007): Kompetenzen historischen Denkens. Ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik. Neuried: Ars Una (Kompetenzen: Grundlagen - Entwicklung - Förderung; 2).
- Lautsch, Erwin / von Weber, Stefan (1995): Methoden und Anwendungen der Konfigurationsfrequenzanalyse. Weinheim: Psychologie Verlags Union.
- Mayer Ulrich / Pandel Hans-Jürgen (1976): Kategorien der Geschichtsdidaktik und Praxis der Unterrichtsanalyse. Zur empirischen Untersuchung fachspezifischer Kommunikation im historisch-politischen Unterricht. Stuttgart: Klett.
- Messner, Helmut (1998): Die kompetente Lernerin, der kompetente Lerner. In: forum schule heute, Jg. 12, Heft 4, S. 3-6.
- Meyer, Hilbert (2004): Was ist guter Unterricht? Berlin: Cornelsen Verlag Scriptor.
- NCHS (National Center for History in the Schools UCLA) (1996): National Standards for History (Revised Edition): <http://nchs.ucla.edu/standards/toc.html> (aufgerufen am 2.1.2008).

- Oser, Fritz (2004): Von der Lehrerbildungsforschung zur Forschung, die Lehrerinnen und Lehrer bildet. Einige Vergleichselemente Deutschland – Schweiz. In: Beiträge zur Lehrerbildung, Jg. 22, Heft 2, S. 242-254.
- Paxton, Richard J./Wineburg, Samuel S. (2000): Expertise and the Teaching of History. In: Moon, Bob/ Brown, Sally/Ben-Perez, Miriam (Hrsg.): Routledge International Companion to Education. London: Routledge. S. 855–864.
- Reusser, Kurt/Halbheer, Ueli (2008): Bildungsstandards als Ausgangspunkt für Unterrichtsentwicklung. In: Beiträge zur Lehrerbildung, Jg. 26, Heft 3, S. 304-317.
- Rüsen, Jörn (2008): Historisches Lernen. Grundlagen und Paradigmen. 2., überarbeitete und erweiterte Auflage. Schwalbach/Ts.: Wochenschau Verlag.
- Sauer, Michael (2007): «Historisches Denken» fördern. Kompetenzentwicklung im Geschichtsunterricht. In: Becker, Gerold / Feindt, Andreas / Meyer, Hilbert u.a. (Hrsg.): Guter Unterricht. Massstäbe / Merkmale – Wege / Werkzeuge. Jahresheft XXV / 2007 aller pädagogischen Zeitschriften des Friedrich Verlages. Seelze-Velber: Friedrich Verlag. S. 42–46.
- Schnell, Rainer/Hill Paul Bernhard/Esser Elke (2005): Methoden der empirischen Sozialforschung. 7. Auflage. München/Wien: Oldenbourg.
- Weinert, Franz E. (1996): «Der gute Lehrer», «die gute Lehrerin» im Spiegel der Wissenschaft. In: Beiträge zur Lehrerbildung, Jg. 14, Heft 2, S. 141–151.
- Wilson, Suzanne M. (2002): Research on History Teaching. In: Richardson, Virginia (Hrsg.): Handbook of Research on Teaching. 4. Auflage. Washington: American Educational Research Association. S. 527–544.