

Christof Arn

# Qualität in der Lehre

## Thesen und Überlegungen

Version 1.1, 28.1.2024

Der Text wurde angeregt durch ein Fachgespräch in der Kerngruppe von [adminus.ch](https://adminus.ch) (Esther Wydler, Hermann Knoll u. a.). Ich bedanke mich zudem für Feedbacks aus der Runde der Projektgruppe „Evaluation der Qualität in der Lehre“ der Hochschule Luzern (Projektlaufzeit 2016 bis 2017, Leitung Pia Gabriel und Nicole Grab).

Reaktionen erwünscht:  
[christof.arn@ethikprojekte.ch](mailto:christof.arn@ethikprojekte.ch)



# Grundlagen

## Grundfrage

Wollen wir ein  
Qualitätsmanagement  
oder  
Qualität?

Das Wort «oder» in diesem Satz ist ein nicht ausschliessendes «oder»: Man kann beides auch kombinieren. Wichtig ist dennoch die Einsicht, dass jedes dieser beiden je auch ohne das andere zu haben ist.

Entscheidet man sich dafür, dass man Qualität will, so wird deutlich, dass ein Qualitätsmanagement (zu dem auch eine systematische Anlage von Evaluationen gehören kann), nur ein *mögliches* Mittel ist – und in den bekannten Formen nicht in allen Fällen und Formen zwingend ein probates, wie die real existierende Welt zeigt (Besser 2004, 26–27, am Beispiel von Evaluationen, die nicht zur Qualität beitragen).

## Grundunterschied: Einfachheit und Komplexität

Bekannt ist der Zirkel «Qualität messen», dann «Massnahmen ergreifen», dann wieder «Qualität messen» usw. (vgl. den vierschrittigen «Demingkreis», Deming 1986 – 30 Jahre ist es her). Dieser Endloskreis entstammt der industriellen Produktion, eignet sich z.B. für das Qualitätsmanagement in der Schraubenproduktion – und gehört auch dort hin. Er funktioniert für einfache Qualität von einfachen Produkten, die einfach messbar ist.

	einfach	komplex	dynamisch
Produkt/ Ergebnis	Einfache Produkte sind «an einem Stück» und statisch: Schrauben, Plastikbeutel, usw.	Komplexe Produkte bestehen aus einem Zusammenspiel von Elementen: Fahrplan im öffentlichen Verkehr, Windkraftwerk, usw.	Dynamische Produkte bzw. Ergebnisse verändern sich, auch wenn der Produktions- bzw. Hervorbringungsvorgang nicht aktiv geändert wird: Openairkonzert, Gesundheit,

			Sozialraum usw.
Qualität	Einfache Qualität besteht in der Einhaltung von einfach feststellbaren Eigenschaften: Massgenauigkeit und Festigkeit von Schrauben, Bremswirkung und Verschleissbeständigkeit von Scheibenbremsen, usw.	Komplexe Qualität besteht dann, wenn sie sich aus verschiedenen Eigenschaften so zusammensetzt, dass die Zusammensetzung «mehr ist als die Summe ihrer Teile», sich also nicht auf das Vorhandensein bzw. das Mass der Teileigenschaften einfach reduzieren lässt: Sicherheit eines Verkehrssystems, Witterungsbeständigkeit eines Gebäudes, usw.	Dynamische Qualität besteht dort, wo der Versuch, Qualität zu verbessern bzw. zu halten, (oder auch andere Erfahrungen und Erkenntnisse) laufend zu Neu- und Umdefinition dessen führt, was Qualität in diesem Fall überhaupt ausmacht: Güte von Bildung (die mehr sein will als Humankapitalproduktion), Güte von Mediation, usw.
Messung	Einfache messbar sind Eigenschaften, für die spezialisierte und exakte Messinstrumente bestehen: Länge, Geschwindigkeit, Temperatur usw.	Komplex zu messen sind Eigenschaften, für die kein direkt darauf ausgerichtetes, exaktes Messinstrument besteht, so dass man versucht, Teileigenschaften oder Indikatoren zu messen, von denen man annimmt, sie seien aussagekräftig: Bequemheit eines Sessels, Ästhetik eines Gebäudes, Schädigungspotenzial einer Kernschmelze in einem Atomkraftwerk	Dynamisch zu messen sind Eigenschaften, die auf die man Einfluss nimmt, wenn man sie misst: Position von Elektronen, Fleiss, usw.

Tabelle 1: einfach – komplex – dynamisch. (Die Erstellung einer solchen Tabelle erweist sich als komplex. Sie ist als unfertig und primär als Markierung von notwendigen Unterscheidungen zu verstehen – also als dynamisch, somit als Einladung zur Mitwirkung an verbesserten Versionen.)

Im Bildungssystem sind die «Produkte» komplex – ja mehr als das: Sie sind nicht nur komplex, sondern dazu noch dynamisch. Ein Unterrichtstag, ein Ausbildungsmodul, ein Weiterbildungskurs ist «lebendig», eine Schraube «tot». Verhält sich die Schraubenproduktionsreihe nach einem halben Jahr gleich wie vorher, entstehen gleiche Schrauben. Verhält sich eine Ausbildungsdozentin im nächsten Semester gleich wie im letzten, kann das Modul gleich gut laufen, viel schlechter oder viel besser. Das Modul selbst ist ein lebendiges System, das seinerseits aus lebendigen, d.h. eigendynamischen Teilen besteht, die komplex miteinander interagieren («nicht-triviales System», vgl. Foerster 1984, 10; Willke

2005, 32). Befolgt die Dozentin diejenigen Kriterien, die sich im letzten Semester als wichtig herausstellten auch im nächsten Semester, kann das genau falsch sein.

Zu den wirklichen Qualitätskriterien gehört also gerade die Situativität und Agilität: Reagiert die Dozentin auf veränderte Verhältnisse kreativ? Wenn dem so ist: Wie allerdings wollten wir Situativität und Agilität messen? Solche Qualität ist nicht leicht und schon gar nicht genau messbar. Dazu kommt: Obwohl offensichtlich Agilität eine wichtige Qualität ist, kann es auch Situationen geben, in denen gerade eine gewisse Sturheit, also die Weigerung der Dozentin, sich an veränderte Verhältnisse anzupassen, ausserordentlich gute Qualität ist. Im Bildungssystem ist also nicht nur das «Produkt» komplex, sondern auch die Qualität. Nicht nur das Modul hat Eigendynamiken, verhält sich also potenziell überraschend eigenwillig zu den Versuchen, es zu beeinflussen. Auch die Kriterien der Güte eines Moduls sind nicht voneinander unabhängig, sondern bilden gemeinsam eine komplexes Netz. Zudem gilt hier wiederum: Wie schon das «Produkt» ist auch dessen Qualität nicht nur komplex, sondern dynamisch: Was heute ein Zeichen guter Qualität ist, kann morgen ein Fehler sein.

Dies trifft sogar für sehr allgemeine Qualitätskriterien wie «gutes Lernklima». So sehr man meinen könnte, dieses Kriterium wäre immer gültig, sind auch hier Einwände möglich. Es kann Situationen geben, in denen es den Lernprozess behindert, wenn ein Dozent das gute Lernklima nahtlos aufrecht zu erhalten versucht, anstatt Störungen desselben (vorübergehend) zuzulassen und den Studierenden zuzutrauen, dass sie das auch aushalten können.

## **Grundproblem: Messung als Störung**

«Qualität entsteht in der pädagogischen Arbeit, wenn die übergeordnete Steuerungsinstanz [...] die Bedingungen schafft, unter denen sich die Selbststeuerung der Teilsysteme gut entfalten kann», schreibt Landwehr (2015, 241). Trifft das zu, ist Fremdsteuerung schon im Versuch ein Qualitätsproblem – und der Versuch der Fremdsteuerung beginnt beispielsweise mit einer «Messung» bzw. einem Versuch, zu messen. Warum ist das bereits ein Qualitätsproblem?

Wählt man ein Qualitätssystem, das wesentlich aus Messungen besteht, so wird man nach möglichst leicht Messbarem suchen, nach sogenannten «Indikatoren».

Damit erreicht man in Bildungsorganisationen ein Qualitätsmanagement – aber keine Qualität. Denn wie gezeigt ist die Qualität hier nicht nur komplex, sondern dynamisch: Fast jeder Indikator ist situationsspezifisch manchmal richtig und manchmal auch falsch. (Zu den immer richtigen Indikatoren gehört etwa der Respekt vor den Studierenden – wenn man daraus aber ein Item in einer Studierendenbefragung macht, kann auch das die Dozierenden in die falsche Richtung lenken.)

Die Fokussierung auf Messungen ist aus einem zweiten Grund qualitätsschädlich. Das Stichwort heisst «bürokratischen Zielverschiebung», ein Effekt, den schon von Merton (1940) beschrieben hat. Das Phänomen ist offenbar aktuell, denn es wird von zahlreichen Forscherinnen und Forschern aufgegriffen, so z.B. von Rösli (2015). Er bringt die Problematik in einen Zusammenhang mit Hirnforschung und Organisationsentwicklung. Gemeint ist Folgendes: Jede Form von Messung bzw. Evaluation setzt ein quasi bürokratisches Anreizsystem, das jenseits des eigenen Willens der Mitarbeiterin liegt, selbst gute Arbeit zu leisten. Insbesondere dann, wenn die Ergebnisse solcher Messungen bzw. Evaluationen Auswirkungen auf die Dozierenden haben («Sanktionen», «lohnwirksam», «öffentlich»), werden die Lehrenden versuchen, sich so zu verhalten, dass die Evaluationen «gut herauskommen». Damit findet eine «Zielverschiebung» statt: Ihr Ziel ist nicht mehr gute Lehre, sondern gute Evaluationen. Wenn in den Evaluationen die Studierenden gefragt werden, ob die Dozierenden die Studierenden über die Modulziele informiert haben, werden die Dozierenden das tun – auch dann, wenn das nicht sinnvoll ist, und vermutlich auch in einer Art und Weise, in der es nicht sinnvoll ist: nämlich in einer Beamerfolie zu Beginn, mehr oder weniger kommentarlos und ohne Diskussion. Für gute Qualität wichtiger wäre die Art und Weise des Umgangs mit diesen Modulzielen in Vorbereitung, Beschreibung, dem Verlauf eines Moduls und in der Art und Weise des Leistungsnachweises – doch dies alles wird nicht «gemessen», weil eben das schwierig zu messen wäre. Auf Messung fokussierte Qualitätssysteme brauchen leicht messbare Items.

Strathern sagt es so: «When a measure becomes a target, it ceases to be a good measure» (1996, 4, zitiert nach Osterloh 2013, 104). Die Verhaltensänderung, die entsteht, wenn sich Dozierende an den Items der Evaluation zu orientieren

beginnnen, würde Osterloh (a.a.O., 105) als «perverses Lernen» bezeichnen, welches auf drei Effekten beruht: «Zielverschiebung, [...] Verdrängung der intrinsischen Motivation und Versuche das System zu schlagen».

Zehnder et. al (2017, 22 und passim) zeigen mit Bezug empirische Forschung, dass komplexe Tätigkeiten ganz generell von der Führung nicht sinnvoll via Kontrolle und Vereinbarung übernommen werden kann.

Es gäbe noch weitere starke Argumente, warum die Fokussierung auf Messung im Bildungssystem schädlich ist, doch reichen die genannten zweifellos aus, um deutlich zu sehen: Der beschriebene Zirkel von «Messungen» und «Massnahmen» mit seiner Fokussierung auf das möglichst exakt Messbare gehört definitiv zur Industrieproduktion und nur dorthin. Anders sieht es aus, wenn man den «Demingkreis» für komplexe Qualität uminterpretiert. Aus der Phase «check» als quantifizierender Messphase wird dann eine Phase «discovery» als gemeinsame Betrachtungs- und Interaktionsphase mit dem Ziel, nicht nur Stärken und Schwächen im Dialog zu entdecken, sondern gleichzeitig Ideen für Veränderungen miteinander zu entwickeln. Damit wird auch deren Umsetzung wahrscheinlicher.

Als zusätzliches, wichtiges Argument für eine Fokusverschiebung weg von quantifizierender Messung sei noch eines wenigstens erwähnt: Immer dort, wo gelingendes Beziehungsgeschehen für die Qualität wichtig ist, stören Kontrolle und extrinsische Anreize stark, weil die Echtheit der Kommunikation elementare Qualitätsbedingung wird. Sobald Studierende, bewusst oder unbewusst, aber zu Recht, wahrnehmen, dass Dozierende etwas «wegen der Qualität» und nicht «wegen ihnen und ihrem Lernprozess» tun, ist das eine Qualitätsverminderung. Letztlich funktioniert nur echte intrinsische, «dozierendeneigenwillige» Qualitätsorientierung wirklich: nur der je eigene Wille zur guten Lehre.

Daraus folgt: Gute Qualität im Bildungswesen wird durch Messungen gestört. Hier (und ebenso im Gesundheitswesen und in verschiedenen anderen verwandten Feldern) braucht es ein Qualitätssystem, das komplexer und schlauer denkt.

Der Verzicht auf Messungsfokussierung bedeutet allerdings nicht den Verzicht auf Studierendenbefragung: Man kann auch Fragen stellen, ohne messen zu wollen.

Das ist eine der entscheidenden Pointen. Man kann nämlich Fragen stellen in der erkennbaren Absicht, aus den Antworten Ideen für Verbesserungen zu entwickeln («discovery»). Dies ist bereits ein Beispiel für das erste von zwei Grundprinzipien für eine Theorie des Qualitätsmanagements im Bildungssystem.

## **Theoriebezogenes Fazit:**

### **Zwei Grundprinzipien für Arbeit an der Qualität**

1. Man kann Qualität auch verbessern, ohne sie zu messen.
2. Man kann Qualität auch verbessern, ohne sich auf fixe Kriterien festzulegen.

Ad 1: Es lassen sich in konkreten Institutionen bzw. Situationen zweifellos Aktivitäten benennen, die plausiblerweise die Qualität verbessern. Ein Beispiel: Gegenseitige Hospitationen von Dozierenden untereinander, optional gestützt durch unterschiedliche, anregende Kriterienblätter und Anleitungen, unter denen man wählen kann, werden zwar nicht immer, aber mit relativ hoher Wahrscheinlichkeit sich insgesamt auf die Qualität der Lehre positiv auswirken – vorausgesetzt, sie werden nicht primär «absolviert», weil man dazu verpflichtet wurde, sondern die Dozierenden wollen von sich aus oder können inhaltlich dafür gewonnen werden, zumindest zu einem gewissen Teil. Dann aber wird man jenseits aller Messungen der Annahme zustimmen: Es ist plausibel davon auszugehen, dass so Qualität der Lehre verbessert wird. In diesem Sinn lassen sich – Kreativität vorausgesetzt – möglicherweise sogar endlos qualitätsförderliche Aktivitäten (er-)finden. (Dazu könnte auch gehören: ein Jahr lang auf solche Aktivitäten zu verzichten.)

Ad 2: Hat man das Konzept der «dynamischen Qualität» verstanden, so wird deutlich, dass man darauf mit einem dynamischen Umgang mit Kriterien reagieren kann. Es geht nicht mehr darum, absolut gültige Kriterien zu finden, sondern Hinsichten, die *jetzt* hilfreich sind. Morgen werden es – ganz dynamisch – andere sein, aber heute wollen wir uns diesen widmen.

## Pragmatisches Fazit

Das Qualitätssystem im Bildungssystem wird also nicht ein Zirkel sein, sondern ein dynamisches System, ein «agiles Qualitätsmanagement» (Arn 2024). Was könnte dazu gehören?

### Laufender Qualitätsdiskurs

Wird in einer Bildungsinstitution insbesondere unter den Lehrenden und den Studierenden, durchaus auch mit Mitarbeitenden in Administration und Führung, ein gerne gelegentlich streitbarer Diskurs über die Frage geführt, was Qualität aktuell und konkret ausmache, so wird dieser Diskurs alleine schon plausiblerweise die Qualität befördern. Er verstärkt das Bewusstsein über und den Fokus auf die Güte der Lehre.

Dieser primär mündliche Diskurs soll regelmässig in kompakte schriftliche Formulierungen münden, also in Kurzpapiere, vor allem jedoch in Beiträge in entsprechenden elektronischen Foren, vielleicht Tweets usw. In Text geronnener Diskurs wird dabei stets als Zwischenstand für die Fortführung dieser Fachgespräche verstanden – immer als vorläufiger und zugleich engagierter Stand.

### Laufende Qualitätskreativität

Dazu kommt passenderweise eine laufende Entwicklung von Ideen zur Verbesserung der Qualität, je alleine und/oder im Austausch. Allerdings geht es um mehr als um Ideen und vor allem geht es hier um mehr als um einen Diskurs: Es geht um eine Kultur fortlaufenden Experimentierens. Damit wird die vielbeschworene «lernende Organisation» tatsächlich Realität – und dies impliziert auch, dass Scheitern erlaubt ist. Wäre das nicht der Fall, wäre die Rede vom «Experimentieren» blosser Schein.

Das nun ist der Punkt, an dem das Wort «Massnahme» wieder seinen Raum finden kann: Welche «Massnahmen» schaffen den Raum für «Lust am lehrreichen Scheitern», für Risikobereitschaft in der Lehre kombiniert mit engagierter, hartnäckiger Suche nach gelingendem Lernen?



Hier passt die «Qualitätsparty»: In wenig strukturierter, eher informell gestalteter Form wird Gelegenheit für Austausch geboten, der durch kurze Inputs und kreative Aufträge stimuliert wird. Individuelle ebenso wie gemeinsame Aktivitäten können entwickelt werden. Qualitätspartys können unter Dozierenden, unter Studierenden, gemeinsam oder auch kombiniert mit der Teilnahme von Mitarbeitenden aus Administration und Führung veranstaltet werden.

Hier passen auch alle Formen von Kommunikation über gelingende Qualität wie über misslingende (Fehlerkultur!): Hospitationen, schriftliche Kommunikationen erfolgreicher Lehraktivitäten, Intervisionsgruppe, Pleiten-, Pech- und Pannen-Contest usw.

Randbemerkung: Wenn überhaupt Preise für gute Lehre, Ratings und Ähnliches, dann immer für Prozesse, nicht für Ergebnisse. Risikobereitschaft, fortlaufendes Experimentieren, Suchen jenseits der Trampelpfade des Mainstreams ist das, was sich zu honorieren lohnt, weil es andere Dozierende mit motiviert, eigene Versuche zu unternehmen. Blosser Studierendenzufriedenheit (die vielleicht auch mal davon lebt, dass bescheidene Erwartungen an die Studierenden mit originellen Dozierendenauftritten bzw. Medieneinsätzen kombiniert werden) als Anlass für Preise oder Ratingstufen zu nehmen, würde wenig förderliche Hierarchien unter Dozierenden zementieren.

## **Schluss: Organisationseinheit «Qualität» und Wissenschaftlichkeit**

Es dürfte Sinn machen, eine Qualitätsgruppe zu pflegen, mit (teilweise) rotierenden Mitgliedschaften und mit den Aufgaben

- fortlaufend Brennpunkte auszumachen,
- die Dynamik des fortlaufenden Qualitätsdiskurses wie der fortlaufenden Qualitätskreativität entsprechend zu stimulieren,
- bei grösseren Problemen gezielt die gelbe Karte heilsam auf den wunden Punkt zu legen,
- dafür zu sorgen, dass die Erfolge gefeiert werden.

Apropos Wissenschaftlichkeit: Der aktuelle Stand der Diskussion über wissenschaftliche Methodik in der Evaluation scheint der zu sein, dass dieselben

methodischen Standards einzuhalten seien wie bei der qualitativen und der quantitativen Forschung. Das wäre immerhin das – trifft für die meisten real existierenden Evaluationen offensichtlich nicht zu: Denn strikte Grundregel empirischer Forschung (wie theoretischer Forschung auch), ist die konzise Formulierung einer Forschungsfrage. Mit gutem Grund mühen sich Studierende bei den ersten Gehversuchen wissenschaftsorientierten Abfassens schriftlicher Arbeiten mit der Formulierung, genauer gesagt: mit dem Denken einer klaren und sinnvollen Forschungsfrage oft ab. Allerdings nicht nur sie – genauso ergeht es den meisten Doktorandinnen und Forschenden insbesondere im sozial- und geisteswissenschaftlichen Bereich. Denn nicht selten besteht darin die grössere Denkleistung als in allem anderen, was zu einer Forschungsarbeit gehört. Die meisten real existierenden Evaluationen etwa in Bildungsinstitutionen hingegen lassen diesen Arbeitsschritt – schwupps! – einfach aus: In aller Regel werden Fragebögen mit Items generiert, ohne je ein Erkenntnisinteresse sorgfältig formuliert zu haben. Mit Pauschalfragen von der Art: „Wie gut ist dieser Studiengang?“ würde sich eine Vollblutforscherin nie zufrieden geben. Sie würde herausarbeiten, welche präzisere Frage konkret in diesem Fall, bei diesem Studiengang, interessant wäre. Wo allerdings schon von Anfang an die spannende Frage fehlt, will es auch nicht recht erstaunen, wenn sich nachher für die Resultate niemand interessieren mag.

Doch ist der Irrtum viel umfassender: Anerkennt man das vorher erläuterte Messproblem, die Komplexitätsthematik und die weiteren oben beschriebenen Phänomene, wird deutlich, dass klassische quantitative Forschung und übliche qualitative (Interview-Interpretations-)Forschung als Herangehensweise schon an sich eher ungeeignet ist. Wissenschaftliche Zugänge, die hier greifen, können in der Organisationsethnologie, in der Aktionsforschung und verwandten Zugängen zu finden sein. Diese Forschungsformen vermeiden es, Menschen und menschliches Handeln zu einem blossen Objekt wissenschaftlicher Beobachtung und Messung zu degradieren. Sie sind sinnigerweise mit den Prinzip verbunden, den Beforschten «eine Stimme zu geben» und berücksichtigen den Einfluss der Forschung auf die interessierenden Phänomene. Damit endlich werden Dozierende und Studierende von Schrauben unterschieden und kommen wieder in den Subjektstatus.

## Quellen und weiterführende Literatur

- Arn, Christof (2024): Agile Hochschuldidaktik. Beltz-Juventa, Weinheim.
- Besser, Ralf (2004): Transfer. Damit Seminare Früchte tragen. Beltz, Weinheim.
- Deming, William E. (1986): Out of the Crisis, MIT, Mass, Cambridge.
- Foerster, Heinz von (1984): Prinziples of self-organisation – In a socio-managerial context. In: Ulrich, Hans und Probst, Gilbert (Hrsg.): Self-organization and management of social systems, Springer, Berlin. Seiten 2–24.
- Landwehr, Norbert (2015): Prozesse im pädagogischen Qualitätsmanagement. In: Bergelehner, Florian; Wilbers, Karl (Hrsg.): Schulisches Prozessmanagement. Einführung, Praxisreflexion, Perspektiven. epubli, Berlin. Seiten 226–248.
- Merton, Robert K. (1940): Bureaucratic structure and personality. In: Social Forces 18. University Press, Oxford. Seiten 560–568
- Osterloh, Margit (2013): Das Paradox der Leistungsmessung und die Nachhaltigkeit der Forschung. In: Nova Acta Leopoldina NF 117, Nr. 398. Wissenschaftliche Verlagsgesellschaft, Stuttgart. Seiten 103–113.
- Röösli, Franz (2015): Initialisierung musterbrechender Managementinnovation. Eine interdisziplinäre Betrachtung. Eul-Verlag, Lohmar.
- Rosenstiel, Lutz von (2007): Organisationspsychologie. 6. Auflage. Kohlhammer, Stuttgart.
- Strathern, Marilyn (1966): From improvement to enhancement: An anthropological comment on the audit culture. In: Cambridge Anthropology 19. Berghahn, New York und Oxford. Seiten 1–21.
- Willke, Hellmut (2005): Systemtheorie II: Interventionstheorie. Lucius & Lucius, Stuttgart.
- Zehnder, Christian; Herz, Holger; Bonardi, Jean-Philippe (2017): A productive clash of cultures: Injecting economics into leadership research. In: The Leadership Quarterly, Volume 28, Issue 1, February 2017, Pages 65–85. <http://dx.doi.org/10.1016/j.leaqua.2016.10.004> .