

# Kriteriendimensionen zur Auswahl und Bezeichnung hoch relevanter Programmelemente

## Checkliste für Monitoring und Evaluation

Goldie MacDonald

2013

Original:

[http://www.wmich.edu/sites/default/files/attachments/u350/2014/Indicator\\_checklist.pdf](http://www.wmich.edu/sites/default/files/attachments/u350/2014/Indicator_checklist.pdf)

von Katharina Klockgether 2014 in Absprache mit Wolfgang Beywl übersetzt

Diese Checkliste beinhaltet praxisbezogene [Kriteriendimensionen](#), die bei der Auswahl und Bezeichnung von relevanten [Programmelementen](#) für Monitoring und Evaluation berücksichtigt werden sollten. Die Auswahl relevanter Programmelemente kann eine komplexe und zeitintensive Aufgabe sein. Gelegentlich wird der Prozess ihrer Auswahl und Bestimmung gegenüber den [Stakeholdern](#) nicht transparent gemacht.

Meistens jedoch kommen hierfür Menschen zusammen, die zu [Evaluation](#) und Monitoring über sehr unterschiedliche Kenntnisse verfügen. Deshalb ist zu klären: Wie kann die Qualität möglicher Programmelemente eingeschätzt werden und wie können Stakeholder dazu angeregt werden, sich daran zu beteiligen? Mit der Checkliste sind drei Absichten verbunden: Sie (1) unterstützt den Bearbeitungsprozess und stellt ein einheitliches Vokabular für den Dialog mit den Stakeholdern bereit, um die Programmelemente auszuwählen, sie (2) verstärkt die enge Verbindung der Programmelemente mit den [Evaluationsfragestellungen](#) und sie (3) hilft, eine [Datenerhebung](#) anzulegen, die möglichst klar auf die intendierten [Nutzungen](#) der [Ergebnisse](#) ausgerichtet ist.

Zuallererst muss geklärt werden, was unter einem relevanten Programmelement verstanden wird<sup>1</sup>. In dieser Checkliste wird ein Programmelement als dokumentierbare oder messbare Informationseinheit, bezogen auf das jeweilige [Programm](#), verstanden (z. B. Eigenschaften des Programms, Aspekte der Programmumsetzung, Outcomes). Häufig stellen Programmelemente einen wesentlichen Marker oder eine Näherung bezüglich des Implementierungsstands oder der Outcomes eines Programms dar. Im Falle von Evaluation und Monitoring benötigt ein Programmelement eine [Operationalisierung](#) und eine methodologisch saubere, strenge Datenerhebung. Ein Programmelement kann sowohl auf qualitativen wie auf quantitative [Daten](#) basieren.

Diese Checkliste soll die Monitoring- oder Evaluationsverantwortlichen unterstützen, optimal geeignete und ressourceneffiziente Programmelemente zu identifizieren. Dies kann in Zusammenarbeit mit wichtigen Stakeholdern geschehen, insbesondere mit denen, die bezüglich des Programms Entscheidungen treffen und diese potenziell auf die gewonnenen Ergebnisse abstützen. Die Checkliste sollte frühestmöglich für die [Evaluationsplanung](#) eingesetzt werden, um den Dialog über verschiedene Optionen von relevanten Programmelementen, inklusive praktischer Überlegungen zur Datenerhebung anzureichern und anzuregen.

Die folgenden Kriteriendimensionen sind nach Alphabet ihres englischsprachigen Originals und nicht nach Priorität angeordnet. Die Relevanz und Nützlichkeit eines jeden Kriteriums variiert je nach [Kontext](#) und Nutzerin bzw. Nutzer. Das vorliegende Instrument enthält nicht alle Kriteriendimensionen, die zur Auswahl von Programmelementen berücksichtigt werden könnten.

---

1 Im englischsprachigen Original wird der Terminus „Indikator“ genutzt, den wir durch „Programmelement“ ersetzen, auch da in der Evaluationstheorie ein dringender Bedarf für eine Bezeichnung dieses Sachverhaltes besteht. Unter „[Indikator](#)“ wird in der sozialwissenschaftlichen Literatur in der Regel etwas anderes verstanden, was auch Goldie McDonald betont.



**Akzeptanz in der Praxis und Verwendungsgeschichte**

**Das Ausmaß, in welchem ein Programmelement mit der Verwendungspraxis der Gegenwart und der Vergangenheit übereinstimmt**

Gelegentlich wählen wir ein Programmelement nur deswegen aus, weil es schon in der Vergangenheit genutzt wurde. Diese Kriteriendimension verlangt danach, die Vorteile und Grenzen von Programmelementen zu bestimmen, die schon in der Vergangenheit in Gebrauch waren (auch in anderen [Kontexten](#) und Situationen). Dieser Hinweis soll es vermeiden, ein Programmelement nur deshalb festzulegen, weil er vorher schon einmal genutzt wurde. Wie bei einigen der folgenden Kriteriendimensionen mag die Anwendung dazu führen, ein Programmelement NICHT zu nutzen. Die Diskussion der Akzeptanz in der Praxis und Verwendungsgeschichte eines Programmelements kann auch dazu führen, die [Datenerhebung](#) neu oder anders ausrichten zu müssen (z. B. sollte die Prägnanz oder die Sensitivität des Programmelements gesteigert werden, um vorhandene Wissenslücken in Bezug auf das Programm zu schließen).

**Anwendbarkeit in verschiedenen Settings**

**Das Ausmaß, in welchem ein Programmelement in unterschiedlichen Settings passend ist**

Diese Kriteriendimension ist besonders bei Datenerhebungen in und an unterschiedlichen Standorten von Belang (z. B. Cluster- oder Mehrebenen-Evaluation). Häufig ist die Möglichkeit, vergleichbare und konsistente [Daten](#) an unterschiedlichen Standorten zu erheben, ausschlaggebend für den Grad der Nützlichkeit eines Programmelements.

**Datenverfügbarkeit**

**Das Ausmaß, in welchem Daten für die (Evaluations-)Studie zugänglich sind**

Bei dieser Kriteriendimension müssen je Programmelement potenzielle [Datenquellen](#) (primäre und sekundäre) identifiziert werden. Die Überlegungen sollten, je nach Bedarf, die Verfügbarkeit von Daten zu einem bestimmten Zeitpunkt (z. B. Daten zur Ausgangslage) und über die Zeit hinweg beinhalten. Zu beachten ist, dass ein Programmelement nicht ausschließlich deswegen gewählt werden sollte, weil Daten dazu vorhanden sind.

**Belastung der Beteiligten und Betroffenen**

**Das Ausmaß, in welchem die Datenerhebung die [Beteiligten](#) und [Betroffenen](#) belastet**

Vor der Entscheidung für ein Programmelement sollte das Ausmaß der (personellen und finanziellen) Belastung geprüft werden. Falls sich das zu evaluierende Programm inhaltlich mit anderen Programmen aus demselben Kontext überschneidet, sollten mögliche Kooperationen (in Bezug auf Datenerhebungen) in Betracht gezogen werden. Dies führt sowohl zu geringerer Belastung der Beteiligten und Betroffenen durch die Datenerhebung als auch zu einer maximalen Ausschöpfung der begrenzten Ressourcen der (Evaluations-)Studie.

**Begriffsschärfe und Bedeutungsklarheit**

**Das Ausmaß, in welchem ein Programmelement eindeutig ist und den [Evaluationsgegenstand](#) korrekt wiedergibt oder repräsentiert**

Ein Programmelement sollte NICHT verwendet werden, wenn es den Evaluationsgegenstand (d. h. das Objekt der Evaluation), einschließlich seiner Rahmenbedingungen wie dem Kontext des Programms, verzerrt oder nicht korrekt repräsentiert. Von den Beteiligten und Betroffenen sollte das Programmelement als aussagekräftige Beschreibung oder als Marker für den Zustand

<p>des Evaluationsgegenstandes angesehen werden. Außerdem sollten die Beteiligten und Betroffenen überlegen, ob zusätzliche Informationen benötigt werden, um die jeweiligen Daten für ein spezifisches Programmelement adäquat oder vollständig interpretieren zu können.</p>
<p><input type="checkbox"/> <b>Kulturelle Passung und Relevanz</b></p> <p><b>Das Ausmaß, in welchem ein Programmelement kulturell adäquat ist, in Bezug auf seinen Inhalt, Fokus und die zugehörigen Datenerhebungen</b></p>
<p>Die Stakeholder sollten bestimmen, ob das Programmelement (d. h. das Konstrukt und die <a href="#">Datenerhebungsmethode</a>) für den jeweiligen Kontext adäquat und relevant ist. Meist ist bei dieser Entscheidung ein vertiefender, aussagekräftiger Dialog mit den Stakeholdern notwendig.</p>
<p><input type="checkbox"/> <b><a href="#">Datenqualität</a></b></p> <p><b>Das Ausmaß, in welchem gesammelte Daten vollständig, reliabel und <a href="#">valide</a> sind</b></p>
<p>Die <a href="#">Kriterienpunkte</a> für die Datenqualität variieren je nach Kontext, abhängig von den Erwartungen und <a href="#">sozialen Werten</a> der Stakeholder. Der jeweilige Stellenwert verschiedener Dimensionen und Arten der Datenqualität ist ebenfalls variabel. Die Evaluationsverantwortlichen sollten daher die Stakeholder in die Entwicklung von Standards für Datenqualität einbeziehen.</p>
<p><input type="checkbox"/> <b>Einsatz von Ressourcen</b></p> <p><b>Das Ausmaß der benötigten Ressourcen (z. B. finanzielle Mittel, Personal, Zeit), das für Datenerhebung, <a href="#">Auswertung</a> und <a href="#">Nutzung</a> der Ergebnisse benötigt wird</b></p>
<p>Die Auswahl eines relevanten Programmelements setzt voraus, dass eine klare Vorstellung über die benötigten Ressourcen zur Erhebung und Auswertung der Daten besteht. Diese Kriteriendimension erfordert ein Abwägen von Wissen oder Kompetenzen, die für die Verwendung des Programmelements benötigt werden. Häufig führt die Diskussion über den Einsatz von Ressourcen zu harten Entscheidungen bezüglich <a href="#">Durchführbarkeit</a>, Qualität und Aktualität der Daten.</p>
<p><input type="checkbox"/> <b>Neutrale Sprache</b></p> <p><b>Das Programmelement ist ungerichtet formuliert, nicht schon vor der Datenerhebung positiv oder negativ ausgerichtet</b></p>
<p>Programmelemente sollten NICHT mit einer Ausrichtung oder Präferenz in eine bestimmte Richtung formuliert werden. Häufig werden <a href="#">Outcomes</a> von Programmen als Aussagen über Erreichtes oder Fortschritt formuliert (z. B. Zuwachs oder Abnahme). Programmelemente sollten neutral und ungerichtet formuliert werden. Die <a href="#">Interpretation</a> oder <a href="#">Bewertung</a> der Daten sollte erst erfolgen, wenn die Datenerhebung abgeschlossen ist. Beispielsweise sollte ein Programmelement besser mit „Wissen“ als mit „Wissenszuwachs“ oder besser mit „Häufigkeit des Tabakkonsums bei Jugendlichen“ als mit „abnehmender Konsum von Tabak bei Jugendlichen“ bezeichnet werden.</p>
<p><input type="checkbox"/> <b>Offenheit für unerwartete und <a href="#">nicht-intendierte Resultate</a></b></p> <p><b>Das Ausmaß, in welchem ein Programmelement (oder ein Set) die Möglichkeit bietet, unerwartete oder nicht-intendierte Aspekte des Programms zu registrieren</b></p>
<p>Meist konstruieren wir Programmelemente, mit welchen bestimmt werden kann, ob erwartete Resultate eintreten, sei es in Bezug auf Programmimplementierung oder auch auf Outcomes. Einige der Programmelemente sollten jedoch ausreichend flexibel sein, um auch die Registrierung unerwarteter oder nicht-intendierter Resultate zu ermöglichen.</p>



**Anbahnung der Ergebnisnutzung**

**Das Ausmaß, in welchem die Nutzung und die Nutzenden der Ergebnisse bekannt und akzeptiert sind**

Diese Kriteriendimension betrifft folgende zwei Themen: wer die gesammelten Daten nutzen soll und wie die erzeugten Informationen genutzt werden sollen. Die Kriteriendimension soll gewährleisten, dass bei der Erstellung von Programmelementen eine explizite und gezielte Verbindung zur vorgesehenen Nutzung der Ergebnisse hergestellt wird. Demzufolge ist es ausschlaggebend, dass die gesammelten Daten zum intendierten Nutzen passen (d. h. sie müssen die Art und Qualität von Informationen liefern, um die intendierte Nutzung erreichen zu können).

**Relevanz für die Evaluationsfragestellungen**

**Das Ausmaß, in welchem ein Programmelement zur Beantwortung der Evaluationsfragestellungen beiträgt**

Mit dieser Kriteriendimension soll geprüft werden, inwiefern das Programmelement mit den Evaluationsfragestellungen verknüpft ist. Wenn ein Programmelement nicht eindeutig zur Beantwortung der Fragestellungen beiträgt, sollten die Stakeholder genau prüfen, ob das Programmelement tatsächlich beigezogen werden soll. Außerdem sollten die Stakeholder darauf achten, ob die gewählten Programmelemente ausreichend Diversität mit sich bringen, um angemessen zur Beantwortung der Evaluationsfragestellungen beizutragen. Das Set der Programmelemente könnte sich zu stark auf einen Ausschnitt des Evaluationsgegenstandes konzentrieren (z. B. Verwendung finanzieller Ressourcen, Ausführung geplanter Aktivitäten, Durchführung des Programms oder Dienstleistungen). Ein Beispiel: Wenn die Stakeholder Fortschritt in Richtung intendierter Outcomes dokumentieren wollen, dürfen die Gegenstandselemente keinesfalls auf Implementation oder Outcomes beschränkt sein.

**Evidenzstärke oder substantielle Güte**

**Das Ausmaß und die Qualität der Informationen, die das Programmelement als adäquate Beschreibung oder Messgröße des Evaluationsgegenstandes unterstützen**

Das Ausmaß, in welchem ein Programmelement als fachlich gut fundiert in einem bestimmten Fachbereich angesehen wird. Der Nachweis der Nützlichkeit eines Programmelements kann über veröffentlichte und unveröffentlichte Literatur, Erfahrungswissen oder soziale Werte der Stakeholder, Beratung durch programmnahe Personen oder über Erkenntnisse aus Monitoring oder Evaluation von ähnlichen Programmen erfolgen. Die Stakeholder sollten dabei festlegen, ob die gefundene Evidenz ausreichend ist, um den Einsatz des Programmelements zu unterstützen.

**Mehrwert innerhalb eines Sets von Programmelementen**

**Das Ausmaß, in dem ein einzelnes Programmelement zu einem Set zusätzliche Bedeutung beisteuert**

Häufig wird ein Programmelement von Stakeholdern eher als alleinstehende Informationseinheit anstatt als Bestandteil eines Sets wahrgenommen, welches im gemeinsamen Zusammenspiel auf die Evaluationsfragestellungen eingehen soll. Neben der Einzel-Bewertung des Wertes eines Programmelements sollten Stakeholder eine Bewertung des Programmelements bezüglich seines Wertes im Rahmen eines Sets vornehmen, um seine Nützlichkeit vollständig einschätzen zu können. Diese Kriteriendimension trägt zur Identifizierung von Doppelungen und Redundanzen bei der Datenerhebung bei und führt möglicherweise zum Einsatz von besser handhabbaren und einer geringeren Zahl von guten Programmelementen.



## Disclaimer

Die Checkliste ist Bestandteil der „Evaluation Checklists Website“ des Evaluation Center an der Western Michigan University. Die Nutzenden erhalten sie als kostenlosen Service. Die Verfassenden haben die Checklisten nicht auf einen spezifischen Bedarf zugeschnitten; die Nutzenden wenden die Checklisten nach eigenem Ermessen und Urteil an. Es werden keine Zusicherungen oder Garantien gegeben, dass diese Checkliste für den spezifischen Zweck geeignet ist, für den die Nutzenden sie einsetzen wollen. Die Verfassenden stellen solche Garantien oder Zusicherungen ausdrücklich in Abrede.

Diese deutschsprachige Fassung erscheint auf der Website von Univation – Institut für Evaluation, Köln. Herausgebende der deutschen Versionen der auf dieser Webseite veröffentlichten Checklisten sind Wolfgang Beywl und Katharina Klockgether. Die Übersetzungen erfolgen mit Erlaubnis der Verfassenden; mehrdeutige Textstellen sind teils in Rücksprache mit ihnen geklärt. Teils finden sich Erläuterungen der Übersetzenden in Fußnoten. Bei der Übersetzung von Akteurs-Begriffen wie „*participant*“ werden geschlechtsneutrale Bezeichnungen bevorzugt (z. B. Teilnehmende) oder beide grammatikalischen Geschlechter genannt (z. B. Kunden und Kundinnen). Gelegentlich wird im Plural formuliert auch wenn im Original der Singular steht. In Ausnahmefällen wird zur Vereinfachung des Satzaufbaus/bei Platzmangel (in Fragebogen, Tabellen ...) das generische Maskulinum genutzt (z. B. Einwohnerzahlen). Teils werden den übersetzten Begriffen die englischsprachigen Originalbegriffe in doppelten spitzen Klammern und kursiv ausgezeichnet nachgestellt. Beispiel: Auftraggebende «*clients*».

Mehr Informationen zur Zwecksetzung und Konzeption der Checklisten finden sich auf: <http://www.wmich.edu/evaluation/checklists>

Besonders hingewiesen sei auf die Leitlinien für eine akzeptable Verwendung der Checklisten:

<http://www.wmich.edu/evaluation/checklists/acceptable-use>

Copyright beim Verfasser; für die Übersetzung ins Deutsche zusätzlich beim Herausgeber der deutschen Checklisten.

**Zitierweise:** MacDonald, Goldie (2013): Kriteriendimensionen zur Auswahl und Bezeichnung hoch relevanter Programmelemente. 2014 übersetzt von Katharina Klockgether und Wolfgang Beywl.

Downloadbar von <http://univation.org/checklisten>.

