

Stake, Robert E. (2004): Standards-Based & Responsive Evaluation. Thousand Oaks: Sage Publications, 329 Seiten, ISBN 0-7619-2665-8

Bereits der Titel enthält die zentrale Botschaft des Buches: die Verbindung einer messorientierten (*standard-based*) mit einer erfahrungsorientierten (*responsive*) Perspektive. Robert Stake erweitert und vertieft mit diesem Werk seinen responsiven Evaluationsansatz. Er versteht sich als Alternative zu den bis in die 70er Jahre vorherrschenden „geschlossenen“ Evaluationsmodellen und fordert, die Wertperspektiven der Programmteilnehmer in die Evaluation einzubeziehen und empfänglich (responsiv) auf die Anliegen der Beteiligten zu reagieren. Wegweisend für die Weiterentwicklung der Evaluationsmethodologie sind Stakes Werke „The Countenance of Educational Evaluation“¹ und „Evaluating the Arts in Education: A Responsive Approach“².

Mit dieser Publikation ergänzt Stake seine Fürsprache für qualitative Evaluationszugänge um eine kritische Würdigung des quantitativen Paradigmas.

Zu Beginn werden grundlegende Begriffe wie formativ/ summativ anschaulich erläutert (Kap. 1). Die Ausführungen zu Schlüsselbegriffen wie „Kriterien“ und „Standards“ illustrieren Stakes offene und wertesensible Haltung. So definiert er Kriterien als Aspekte bzw. Attribute des Evaluationsgegenstands (*evaluand*), entlang derer eine Beschreibung und Bewertung vorgenommen werden kann. Beispielsweise kann ein Lehrbuch anhand seiner Lesbarkeit, seines Informationsgehalts oder seiner Aktualität beurteilt werden. Standards legen darüber hinaus fest, in welchem Ausmaß diese Kriterien erfüllt sein müssen, damit das Lehrbuch als gut oder schlecht eingestuft werden kann.

Stakes Erläuterungen zeigen auf wie schwierig es ist, Standards für die Bestimmung der Qualität eines Gegenstands festzulegen. Was als gut oder schlecht gilt, wird seiner Meinung nach weniger durch den Gegenstand selbst bestimmt. Die Frage „Does the quality belong to the melon or to the experience?“ beantwortet er folgendermaßen: Die Qualität wird in der Interaktion zwischen der Melone und dem sie Essenden bestimmt. Sie wurzelt in der menschlichen Erfahrung. Hieraus ergibt sich die Problematik die meist impliziten Standards der den Gegenstand Nutzenden zu identifizieren: Wie saftig, reif und fruchtig muss eine Melone sein, damit sie als gut klassifiziert werden kann? Häufig, vor allem bei komplexen und noch in der Entwicklung befindlichen Evaluationsgegenständen, sind Beteiligte nicht in der Lage, verbindliche Standards für die Beurteilung zu liefern. In der Praxis hat es sich als zweckdienlich erwiesen, zunächst den Evaluationsgegenstand zu studieren und anschließend Standards zu bestimmen. In einem ersten Schritt können Standards willkürlich gesetzt werden, deren Funktionalität später geprüft wird.

Die Beschreibung und Bewertung eines Evaluationsgegenstandes kann auf zwei Wegen erfolgen: durch kriterien-gestützte und/oder durch interpretative Evaluation. Erstere bildet die Qualität durch Zahlen, Skalen und Standards ab. Sie repräsentiert eine quantitative, kriterienbasierte Betrachtungsweise. Bei der interpretativen Evaluation wird Qualität durch subjektive Erfahrung erfasst, durch verbale Beschreibung und multiperspektivische Einschätzung. Stake bezeichnet sie als qualitative Evaluation - episodisch und ganzheitlich. Standard-based Evaluation (Kap. 3) nutzt vorrangig Evaluationsmethoden, die kriterien-gestützt Erkenntnissen erzeugen. Zentrale Steuerungselemente (*factors*) dieser Vorgehensweise sind Bedarfe der Zielgruppe, Programmziele, -leistung und -kosten sowie Evaluationskriterien, Evaluationsstandards und Synthese der Ergebnisse.

¹ Aufsatz in deutscher Sprache veröffentlicht in Wulf, Christian (Hg.) (1972): Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen. München

² Stake, Robert E. (1975): Evaluating the Arts in Education: A Responsive Approach. Columbus/Ohio

Responsiv Evaluierende (Kap. 4) zeichnen sich dadurch aus, dass sie sich durch persönliche Erfahrung leiten lassen sowie die Beteiligten und ihre Werte einbeziehen. Um die Qualität des Programms zu erschließen, sucht die evaluierende Person nach persönlichen und szenischen Eindrücken, die im weiteren Verlauf korrigiert und angepasst werden. Responsive Evaluation schließt kriterien-gestützte Vorgehensweisen nicht aus. Sie wird sich aufgrund der häufig festgestellten mangelnden Aussagekraft von standardisierten Verfahren für die Bestimmung der Qualität nicht auf sie allein verlassen.

Kapitel 5 skizziert Fallstricke im Datenerhebungsprozess und gibt Hinweise zu Anwendungsbereichen und –formen von Evaluationsmethoden. Interessant ist die Diskussion der Besonderheiten evaluationsspezifischer Datenquellen wie Antworten von Betroffenen, Mitarbeitenden, Verantwortlichen und der interessierten Öffentlichkeit. Zentrale Bedeutung haben für ihn die Beschreibungen, Bewertungen und Ideen der Mitarbeitenden und Programmverantwortlichen. Sie wissen sehr viel über das Programm und können durch ihre Hinweise die detektivische Arbeit der Evaluierenden unterstützen. In der Regel dienen informelle Gespräche dazu, eigene Vermutungen und Annahmen zu relativieren oder und zu erhärten. Bewusst sein sollten Evaluierende sich über die Gefahr des „going native“ wie Stake – in Anlehnung an Ethnographen - die kritiklose Übernahme der Perspektive der Beteiligten nennt. Neben den typischen Datenerhebungsinstrumenten wie Interviews und Surveys skizziert Stake die Bedeutung der Rekonstruktion von (Spannungs-)Themen (*issues*) im Evaluationsprozess. In der responsiven Evaluation ist ein wichtiges Erkenntnisinstrument, die Geschichte des Programms, seine Ausgangssituation und die Art und Weise der Implementation nachzuerzählen. Die geschichtlichen Informationen werden in einer Art Evaluationstagebuch in Form von Gesprächsnotizen, E-Mail-Korrespondenz, Aufnahmen und Photographien gesammelt. Nach Stake agieren die Evaluierenden ähnlich Kulturanthropologen oder Theaterregisseuren: wichtige programmspezifische Ereignisse und Aktivitäten werden gesammelt und dem Publikum entlang zentraler Themen verständlich präsentiert.

Anschließend an Ausführungen zu Analyse, Synthese und Meta-Evaluation (Kap. 6) widmet sich Stake der besonderen Bedeutung von Stakeholdern im Evaluationsprozess (Kap. 7). Hier werden u. a. weitere beteiligtenorientierte Ansätze wie die partizipative, nutzungsfokussierte und demokratische Evaluation dargestellt. Im vorletzten Kapitel werden interpretative Zugänge zu zentralen Evaluations-Issues wie Programmnormierung und Know-How-Generierung bei Stakeholdern erläutert. Kapitel 9 setzt sich kritische mit Evaluation als evidenzbasierter Forschung auseinander. Abschließend werden ethische Implikationen und Verantwortlichkeiten diskutiert.

Das Buch ist verständlich und anschaulich geschrieben, teils kommt das Gefühl eines persönlichen Gesprächs zwischen Autor und Leser/Leserin auf. Bemerkenswert ist die dichte und abwechslungsreiche Kapitelgestaltung: Inspiriert durch fiktive Diskurse Galileo Galileis endet jedes Kapitel mit einem humoristischen Dialog zwischen der Evaluatorin Phyllis und ihrem Vorgesetzten Mr. Sagredo. Ein digitaler Appendix (<ftp://webftp.sagepub.com/StakeAppendix>) liefert zusätzliche Materialien wie Projektskizzen und Evaluationsberichte.

Fazit: Rezeptartige Handlungsanweisungen suchen Evaluierende hier vergebens. Vielmehr stellt Stake Orientierungs- und Reflexionshilfen in Form von Phasenmodellen, Schaubildern und Erfahrungsberichten bereit. Für konzeptionelle Entscheidungen steht Interessierten mit diesem Werk ein reichhaltiger Fundus aus 45 Jahren Evaluationspraxis zur Verfügung, belegt durch eine Vielzahl an Arbeitsproben, Cartoons und Geschichten. Praktikerinnen und

Praktiker erhalten mit der Zusammenführung der beiden Perspektiven eine umfangreiche Hilfestellung für die Anlage situationsangemessener Evaluationen, die der Vielfalt der Evaluationskontexte Rechnung tragen. In Stakes Worten: „What works, here and now, usually is a combination of responsive and standards-based, with one taking the lead, the other filling in and rounding out.“