

## Antworten zu Kapitel 2

### Aufgaben, Standards und Modelle der Evaluation

#### 1 Beschreiben Sie den Unterschied zwischen Evaluation und Evaluationsforschung.

Der Unterschied soll anhand eines Beispiels erläutert werden:

- Ein Schulrat besucht die Realschule in Mannheim-Feudenheim. Nach dem Besuch einiger Klassen, einem Gespräch mit dem Lehrerkollegium, der Schulleitung und der Elternvertretung schreibt er in seinem Bericht: „Die Realschule in Mannheim-Feudenheim ist eine innovative Schule. Das Kollegium arbeitet sehr gut zusammen und es verfolgt auch die aktuellen bildungspolitischen Ziele, die vom Kollegium und den Eltern gut angenommen werden“.

Die vom Schulrat getroffene Aussage lässt sich nicht nachvollziehen, auf der Basis welcher Kriterien, welcher Beobachtungen etc. seine Schlussfolgerungen zustande gekommen sind. Eine Bewertung dieser Art kann man als intuitiv bezeichnen. Solche Vorgehensweisen werden als Evaluation gekennzeichnet.

- Hätte dagegen eine Kommission die Realschule in Mannheim-Feudenheim besucht, die vorher darauf trainiert worden wäre, bestimmte Merkmale zu beobachten, zu registrieren und zu einem Urteil zusammenzufassen, und hätte sich diese Kommission nach der Erhebung der wichtigen Daten dann zusammengesetzt und eine Schlussfolgerung auf der Basis eines gemeinsamen und intersubjektiven Urteils gezogen, so könnte man hierbei von einem rationalen Vorgehen sprechen.

Diese Art von Evaluation könnte dabei zu Recht mit dem Prädikat *Evaluationsforschung* versehen werden.

#### 2 Was versteht man innerhalb der Evaluationsforschung unter dem Begriff *Zielsetzung*?

*Zielsetzungen* in der Evaluation betreffen das Implementieren und Umsetzen von zielgerichteten Vorgehensweisen in einem Evaluationsprozess, sodass eine der Fragestellung entsprechende Aussage aus der Gesamtheit dieses Prozesses abgeleitet werden kann.

Hierzu zählt u.a. die *Zustandsevaluation*, bei der im Vordergrund steht, alle Eigenschaften oder Attribute zu ermitteln, die den Evaluationsgegenstand umfassend beschreiben. Beispiele hierfür sind: Wie ist das Leistungsniveau der 8. Klassen des Bundeslands Saarland einzuschätzen? Wie wird derzeit in den Grundschulen des Landes Rheinland-Pfalz individuelle Förderung von Schülerinnen und Schülern umgesetzt? Wie gut ist das Organisationsklima in Firma O?

Weiterhin wird in der Evaluationsforschung die *Veränderungsevaluation* unter die Zielsetzung subsumiert. Sie zielt darauf ab, die Veränderung des Zustands eines Forschungsobjekts in einem raum-zeitlichen Prozess zu analysieren.

Folgende Beispiele umreißen diese Zielsetzung:

- Welche Veränderungen im Leistungsverhalten haben sich bei den 15jährigen Schülerinnen und Schülern in den Allgemeinbildenden Schulen der Bundesrepublik Deutschland in der Zeitspanne zwischen 2000 und 2013 ergeben?
- Wie hat sich das Training der Pflegenden in der Klinik H in der Zeitspanne von 2004 bis 2007 auf die Auftretenshäufigkeit von Lungenentzündungen bei den Patienten ausgewirkt?

Ein weiteres Beispiel für Zielsetzung in der Evaluationsforschung ist die *Wirksamkeitsevaluation*: Hier geht es um die Frage, ob und welche Wirkung durch eine bestimmte Maßnahme erzielt wurde.

Die nachfolgenden Beispiele beschreiben diese Art von Zielsetzung:

- Haben sich seit der Umstellung der Finanzierung von Universitäten auf ein parame-tergestütztes Modell Veränderungen in der Höhe der Einwerbung von Drittmitteln ergeben, die auf die Umstellung zurückzuführen sind?
- Führt die Möglichkeit des Übertrags von nicht ausgegebenen Mitteln des Vorjahres auf das Nachfolgejahr zu einem veränderten Ausgabeverhalten in den einzelnen Abteilungen einer Behörde im Bundesland Niedersachsen?
- Führt die Umsetzung von Hartz IV zu einer Veränderung des Bewerbungsverhaltens bei den Betroffenen?

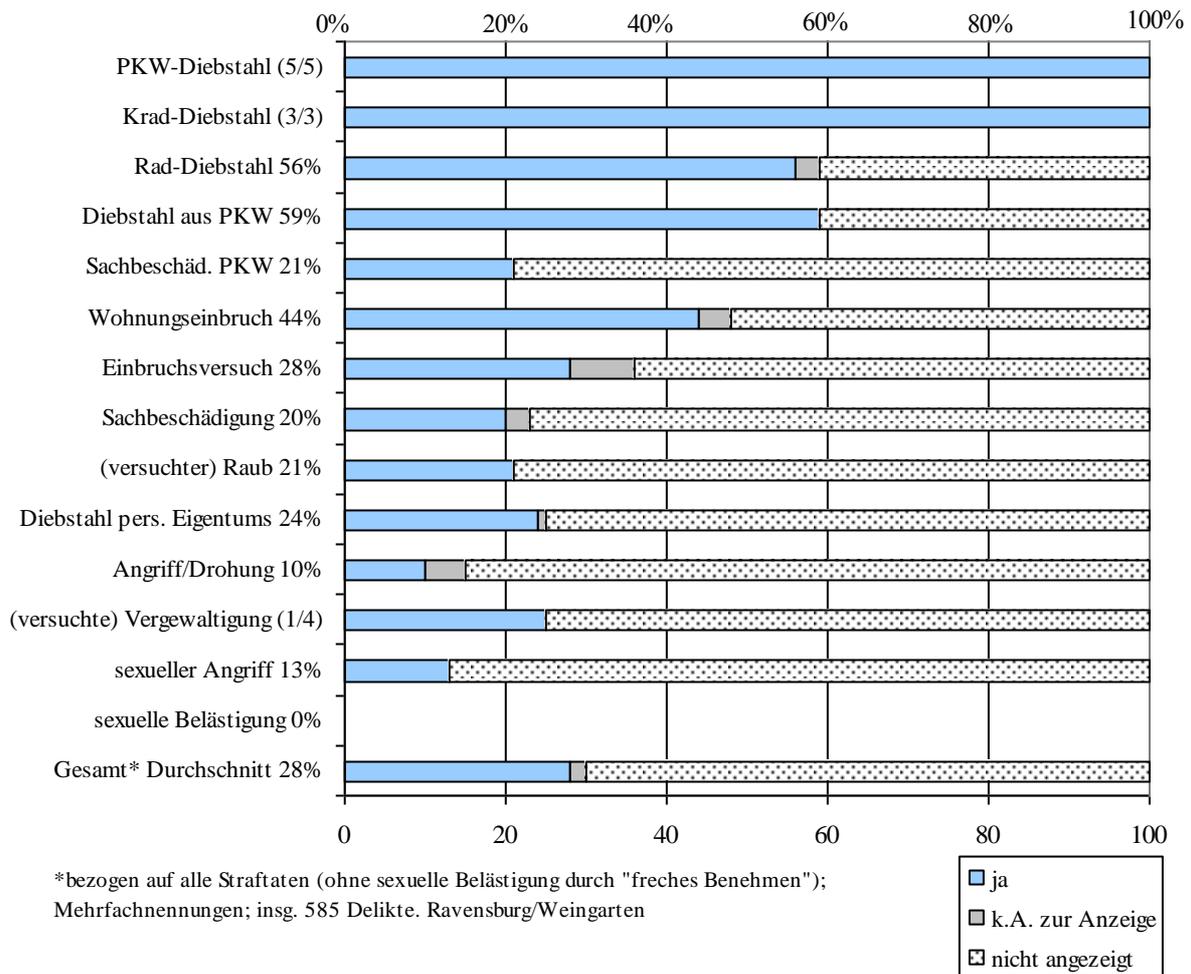
### **3 Beschreiben Sie, was man unter den folgenden drei Begriffen versteht:**

#### **3.1 Beschreiben Sie, was man unter *Zustandsevaluation* versteht.**

Bei der *Zustandsevaluation* sollen alle Eigenschaften oder Attribute ermittelt werden, die den Evaluationsgegenstand umfassend beschreiben.

Ein Beispiel: Kriminalisten gehen unter anderem der Frage nach, wie sich das Anzeigeverhalten von Betroffenen darstellt. Zur Erläuterung wird auf das Anzeigeverhalten im Kreis Ravensburg/Weingarten im Jahre 1994 (Heinz & Spiess, 1995) zurückgegriffen. Im Rahmen eines Projektes zur Viktimisierung, Anzeigeerstattung und Einschätzung der Arbeit der Polizei durch die Bürger wurde in mehreren Projektstädten Baden-Württembergs die Bevölkerung befragt, mit welchen Straftaten sie konfrontiert worden waren und in welchem Umfang sie diese bei der Polizei angezeigt hatten.

**Anzeigenerstattung bejaht in Prozent**



**Abbildung 2.9** Zustandsevaluation. Bei weitem nicht alle Straftaten, denen Bürger zum Opfer fallen, werden auch bei der Polizei angezeigt. Wie ein Projekt in Baden-Württemberg 1994 ergab, werden z.B. nur 44 % aller Wohnungseinbrüche und 20 % aller Sachbeschädigung zur Anzeige gebracht (Heinz & Spiess, 1995)

**3.2 Beschreiben Sie, was man unter Veränderungsevaluation versteht.**

Im Rahmen der *Veränderungsevaluation* wird die Veränderung des Zustands eines Forschungsobjekts in einem raum-zeitlichen Prozess analysiert.

Folgendes Beispiel steht für die Veränderungsevaluation: In einer Pressemitteilung der Universität Koblenz-Landau ist Folgendes zu lesen: „Die Einführung von Teamstrukturen in einem Unternehmen gelingt umso besser, je besser das Organisationsklima in Bezug auf Information, Transparenz und Klarheit über betriebliche Vorgänge von der Belegschaft empfunden wird. Zu diesem Ergebnis kam Prof. Dr. Günter F. Müller vom Fachbereich Psychologie der Universität in Landau bei einer Untersuchung in einem Unternehmen der Informatikbranche. Er befragte die 134 Mitarbeiterinnen und Mitarbeiter der Firma mit Hilfe des von ihm entwickelten „Landauer Inventar zur Diagnose des Organisationsklimas“ (LIDO), nach-

dem die Belegschaft erste Erfahrungen mit neu eingeführten Teamstrukturen gesammelt hatte.“ (Quelle: <http://www.uni-protokolle.de/nachrichten/id/52263/>; Zugriff 12.1.2013)  
Dieses Beispiel demonstriert Veränderungen auf der Basis neuer Teamstrukturen.

### **3.3 Beschreiben Sie, was man unter *Wirksamkeitsevaluation* versteht.**

Mit Hilfe der *Wirksamkeitsevaluation* will man herausfinden, ob und welche Wirkung durch eine bestimmte Maßnahme erzielt wurde.

Beispiel: Bei der Untersuchung der Wirkung eines Medikaments stellt sich z.B. heraus, dass unerwartete Nebenwirkungen auftreten, welche die Fahrtüchtigkeit erheblich beeinträchtigen. Dabei wird deutlich, dass es eine Hauptwirkung gibt – Eindämmung des Heuschnupfens –, aber auch Ermüdungserscheinungen, die als Nebenwirkungen zu beschreiben sind. Die Evaluationsforschung mit dem Ziel der Wirksamkeitsangabe hat die Aufgabe, nicht nur die intendierten, sondern auch die nicht intendierten Wirkungen zu erfassen.

### **4 Warum benötigt man *Gütekriterien* zur Bestimmung, wie gut eine Evaluation ist?**

*Gütekriterien* stellen Festlegungen dafür dar, eine geplante Evaluationsforschung am State of the Art zu orientieren und bereits vollzogene Forschungen dieser Art daraufhin zu beurteilen, ob die Regeln der Kunst eingehalten wurden. Die Gütekriterien beschreiben, inwieweit dies gelungen ist und sie geben auch ein Maß für dieses Gelingen an. Hätte man solche Informationen nicht zur Verfügung, so könnte ein Evaluator nur eine Eigenbewertung durchführen, die vielleicht von anderen Evaluatoren nicht geteilt wird.

### **5 Erklären Sie den Begriff *Gütekriterium*.**

Ein *Gütekriterium* ist ein Maßstab oder Gütesiegel zur Beurteilung, wie gut oder wie schlecht bspw. eine Evaluationsforschung vollzogen worden ist. Das Kriterium wird zur quantitativen oder qualitativen Beschreibung der Güte der Evaluationsforschung herangezogen.

### **6 Beschreiben Sie Kriterien, die zur Klasse der *allgemeinen Gütekriterien* gehören.**

Ein *allgemeines Gütekriterium* sagt unter anderem aus, inwieweit es gelungen ist, eine allgemeine von einem Auftraggeber übermittelte Fragestellung so zu formulieren, dass daraus eine Frage entsteht, die Ausgangspunkt der Evaluationsforschung ist.

Ein anderes Gütekriterium allgemeiner Art betrifft die Objektivierbarkeit: Hierbei geht um die Frage, ob eine Person, die nicht in die Evaluationsforschung einbezogen war, auf Basis eines vorgelegten Berichts alle Schritte des vollzogenen Evaluationsprozesses transparent nachvollziehen kann.

### **7 Beschreiben Sie Kriterien, die zur Klasse der *psychometrischen Gütekriterien* gehören.**

*Psychometrische Gütekriterien* beziehen sich auf die Güte von Messungen. Sie werden quantitativ angegangen.

Nehmen wir ein konkretes Beispiel für eine Evaluationsfragestellung:

Führt die Umsetzung von Hartz IV zu einer Veränderung des Bewerbungsverhaltens der Betroffenenengruppe?

Nunmehr sind mit Blick auf die Veränderung des Bewerbungsverhaltens folgende Fragen zu stellen: Ist die Messung objektiv? Ist sie zuverlässig (Reliabilität)? Wird mit der Messung das Merkmal erfasst, das intendiert war (Validität)? Werden Personen mit einer anderen Muttersprache als Deutsch benachteiligt usw.? (Fairness).

Diese Kriterien werden mit statistischen Kennziffern (z.B. Korrelationen) beschrieben.

### **8 Warum lassen sich die psychometrischen Gütekriterien auch als Kriterien der Produktqualität beschreiben?**

Jedes Instrument zur Erfassung von Merkmalen stellt eine Operationalisierung für diese dar. Die mit Hilfe der gefundenen Operationalisierung erfolgte Messung ist als ein Resultat oder Produkt zu verstehen. Durch die Quantifizierung bzw. Qualifizierung mit Gütekriterien wird dieses Produkt hinsichtlich seiner Quantität bzw. Qualität näher beschrieben.

### **9 Was versteht man unter**

**(9.1) interner Validität und**

**(9.2) externer Validität?**

#### **9.1 Was versteht man unter *interner Validität*?**

Ist man in der Lage, eine eindeutige Aussage über die Annahme oder die Ablehnung der entsprechenden Hypothese zu liefern, so kann hierfür auch der Grund dargelegt werden. In diesem Falle ist eine Erklärung für das untersuchte Phänomen bekannt. Man spricht deshalb davon, dass die *interne Validität* gegeben sei.

#### **9.2 Was versteht man unter *externer Validität*?**

Gelingt, es z.B. den Erfolg einer Therapiemaßnahme zu wiederholen, der zunächst in einer Klinik unter Laborbedingungen nachgewiesen wurde, ihn auf andere Kliniken und andere Therapeuten zu übertragen, so kann man sagen, dass eine Generalisierung gelungen ist. In diesem Fall ist *externe Validität* gegeben.

### **10 Welche Funktion kommt den so genannten *Evaluationsstandards* zu?**

*Evaluationsstandards* kann man als Konventionen (resultierend aus einem Verständigungsprozess von Experten) dafür bezeichnen, wie Evaluationsforschung durchgeführt werden soll. Sie gelten als Minimalstandards.

Sie dienen zur Orientierung bei der praktischen Umsetzung in der Evaluationsforschung. Der Evaluator weiß dann, ob er bestimmte – von Experten gesetzte – Maßstäbe eingehalten hat und ist dann in seinem Vorgehen weniger angreifbar, weil er sich auf diese Experten und demnach die Einhaltung von Standards berufen kann.

### 11 Erklären Sie den Unterschied zwischen Standards und psychometrischen Gütekriterien.

- *Standards* sind Minimalstandards, nach „oben“ gewissermaßen „offen“. Sie sind jederzeit auf der Basis veränderter Experteneinschätzungen veränderbar. Sie lassen sich nicht quantifizieren.
- *Gütekriterien* stellen statistisch untermauerte und quantifizierte Größen dar: Sie basieren – insbesondere bei den Hauptgütekriterien (Objektivität, Reliabilität und Validität) – auf Maßen, welche ihrerseits standardisiert sind (hier bewegen sich die numerischen Werte zwischen 0 und 1).

### 12 Erläutern Sie einige *Hauptkriterien der Evaluationsstandards!*

Zu den *Hauptkriterien* der Evaluationsstandards gehören Nützlichkeits-, Durchführungs- und Genauigkeitsstandards.

- Unter die *Nützlichkeitsstandards* werden unter anderem die Ermittlung der Beteiligengruppen (Beteiligte, Betroffene, Verantwortliche usw.) oder die Glaubwürdigkeit des Evaluators subsumiert: Hierzu rechnet man auch die Qualifizierung des Evaluators durch eine fachliche Kompetenz.
- Im Kontext der *Durchführungsstandards* wird erwartet, dass der Evaluator praxisverträgliche Verfahren einsetzt, bei denen sich Aufwendungen und Ertrag in einer angemessenen Relation halten.
- Bei den *Genauigkeitsstandards* wird eine genaue Beschreibung des Evaluationsgegenstands ebenso erwartet wie Transparenz hinsichtlich der Angemessenheit und Genauigkeit der Informationsquellen.

### 13 Was versteht man unter Evaluationsmodellen?

*Evaluationsmodelle* stellen Instrumente zur Steuerung von vollständigen Evaluationsprozessen oder Ausschnitten solcher Prozesse dar. Sie kommen zustande aufgrund der systematischen Durchdringung von allgemeinen oder spezifischen Evaluationsfragestellungen. So ging Tyler (1949) der Frage nach, welche Kriterien existieren, um das Erreichen eines Curriculums zu evaluieren. Balzer (2005) entwickelte ein normatives Modell, das aus der Sicht eines Evaluators als eine Vorschrift verstanden werden kann, wie der Evaluationsprozess vollzogen werden muss, um ihn erfolgreich zu Ende zu bringen.

Je nach Fragestellung ergeben sich jeweils andere theoretische, methodische und praktische Vorgehensweisen sowie Interpretationen von Evaluationsvorgängen und Evaluationsprozessen.

### 14 Erläutern Sie anhand eines Beispiels das *Modell von Tyler (1949)* und erklären Sie, weshalb es sich hierbei um eine methodische/methodologische Orientierung handelt?

Tyler hat einen Evaluationsprozess auf wesentliche Abfolgeprozesse reduziert. Damit hat er zunächst einen Beitrag zur Methodik des Evaluationsprozesses geleistet. Die einzelnen Teilprozesse werden sodann detailliert angegangen, so unter anderem die „Beschreibung der Ziele

der Evaluations durch geeignete Verhaltenskategorien“. Diese Kategorien klassifiziert er. Und diese Klassifikation (Tyler-Matrix) ist Ausgangspunkt für eine rationale Konstruktion von verhaltensbezogenen Items.

**15 Wie funktioniert der *bewertungsorientierte Ansatz* von Owens (1973) und Wolf (1975)?**

**Erläutern Sie das Modell anhand eines von Ihnen konstruierten Beispiels.**

**15.1 Wie funktioniert der *bewertungsorientierte Ansatz* von Owens (1973) und Wolf (1975)?**

Durch ein Vorgehen, das sich am bewertungsorientierten Ansatz orientiert, wird dem einzelnen Evaluator ein sehr subjektbezogenes Vorgehen im Evaluationsprozess entzogen. Es wird an Stelle eines subjektiven Vorgehens ein transparentes und objektiviertes Prozedere bei der Evaluation etabliert. Dieses Ziel wird dadurch gewährleistet, dass anstelle einer Einzelfallentscheidung eine Jury die Entscheidung über das Vorgehen bei der Evaluationsforschung fällt. Dabei stützt sich die Entscheidung auf einen Diskurs zwischen einer Art von Anwalt und Gegenanwalt, welche durch Fragen und „Verhöre“ seitens der Jury ergänzt werden. Man nennt diesen Ansatz auch *gegnerschaftsorientierte Evaluation* (adversary evaluation). Damit wird ausgedrückt, dass durch den Diskurs aller Beteiligten der Evaluationsgegenstand, das adäquate methodische Vorgehen, Überlegungen über das Design und andere wichtige Sachverhalte des Evaluationsprozesses systematisch durchdrungen werden und durch einen Diskurs eine optimierte Lösung (gegenüber einer Einzelentscheidung) gefunden wird.

**15.2 Erläutern Sie das Modell anhand eines von Ihnen konstruierten Beispiels.**

Diesen Ansatz kann man sich folgendermaßen vorstellen: Bevor ein Evaluationsprozess implementiert wird, will sich ein Evaluator über sein weiteres Vorgehen sicher sein. Hierzu lädt er verschiedene Experten ein, die er drei Gruppen zuordnet: einer Jury, einem Anwalt und einem Gegenanwalt. Anwalt und Gegenanwalt haben die Aufgabe, anhand einer vorgegebenen inhaltlichen Fragestellung den Evaluationsprozess zu erörtern. Der Anwalt beginnt mit Fragen und Aussagen; es folgt die Gegenrede des Gegenanwalts mit Gegenfragen, Infragestellungen von Aussagen des Anwalts etc. Die Jury kann ebenso zu dem erörterten Gegenstand und den bisherigen Aussagen Fragen stellen und beide – Anwalt und Gegenanwalt – „verhören“.

Dieser Erörterungsprozess kommt dann zu einem Ende, wenn der Evaluationsprozess mit allen notwendigen Einzelteilen vollständig diskutiert ist und die Jury festlegen kann, welche einzelnen Teilschritte wie zu vollziehen sind, damit im zu vollziehenden Evaluationsprozess vom Evaluator kein Fehler begangen wird.

**16 Was ist ein *nutzungsorientiertes Modell*? Erklären Sie das Beispiel von Kirkpatrick und Kirkpatrick (2006).**

*Nutzungsorientierte Ansätze* gehen der Frage nach, welche Zwecke mit welchen Arten von Ergebnissen einhergehen. Hierbei hat Kirkpatrick (2006) vier Ebenen formuliert:

- (1) Reaktionsebene (reaction),
- (2) Ebene des Lernerfolges (learning),

- (3) Verhaltensebene (behavior) und
- (4) Ebene objektiver Endergebnisse (results)

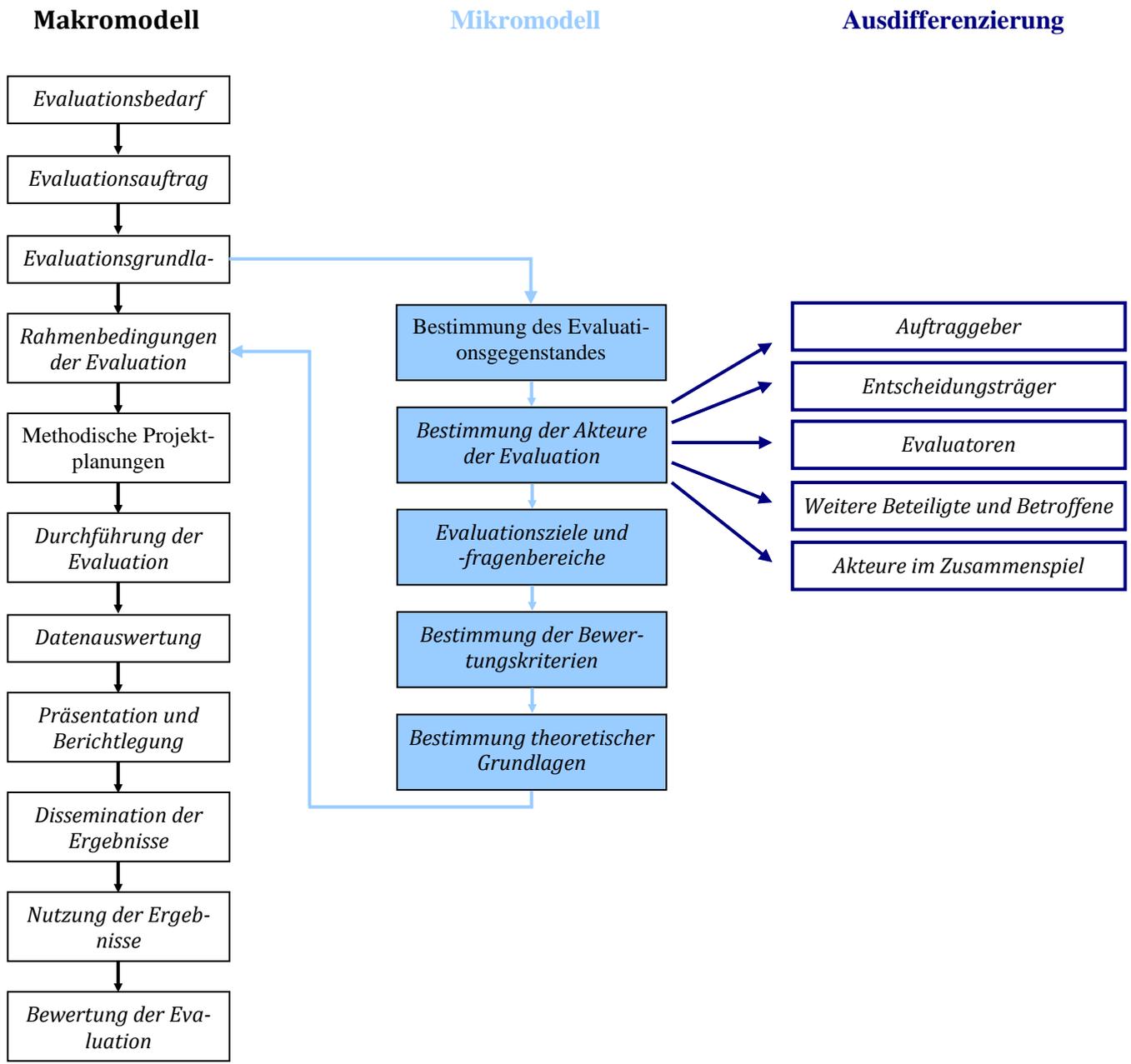
Kirkpatrick hat sein Modell insbesondere auf den Bereich Lernen und Bildung bezogen. Daher ist es sinnvoll, die dargestellten Ebenen durch Fragen zu spezifizieren. Sie können mit dem skizzierten inhaltlichen Hintergrund wie folgt formuliert werden:

- (1) Wie reagieren die Lernenden auf die Bildungsmaßnahme? (reaction)
- (2) Haben sich die Kenntnisse und Fähigkeiten der Lernenden verbessert? (learning)
- (3) Hat sich das Verhalten der Lernenden am Lernort verändert? (behavior)
- (4) Welche Ergebnisse erzielte die Bildungsmaßnahme insgesamt und im Detail? (results)

Durch den Bezug auf die genannten Ebenen wird unmittelbar einsichtig, dass mit dem Modell von Kirkpatrick der jeweilige Nutzen hinterfragt und erfasst wird.

**17 Was ist unter einem *normativen Ansatz* zu verstehen? Erklären Sie die Hauptebenen des Modells von Balzer (2005).**

*Normative Modelle* schreiben vor, welches Vorgehen eingeschlagen werden muss, damit ein Evaluator zu einem erfolgreichen und dem State of the art genügenden Evaluationsprozess kommt. Das Modell von Balzer (2005) ist zunächst als ein Makromodell dargestellt, das die Gesamtheit des Evaluationsprozesses in folgende Hauptbestandteile des Evaluationsprozesses (Makromodell) untergliedert (s. Abb. 2.8, linker Teil; S. 49). Demnach beginnt der Evaluationsprozess mit der Feststellung des Evaluationsbedarfs und endet mit der Bewertung der Evaluation selbst. Anschließend werden Mikromodelle formuliert – hier nicht vollständig dargestellt –, welche sodann ausdifferenziert werden (beispielhaft s. Abb. 2.8, rechter Teil, S. 49)



**Abbildung 2.8** Normativer Ansatz. In der Abbildung wird die Integration des **Makromodells** (linker Teil) und **Mikromodells** (Mitte) des normativen Ansatzes von Balzer (2005) dargestellt. Im rechten Teil der Abbildung wird der Bereich „Bestimmung der Akteure der Evaluation“ in seine Komponenten (Auftraggeber, Entscheidungsträger usw.) ausdifferenziert