

2 Evaluationsprojekte: Strukturen, Abläufe und Qualitätsstandards

Übungsaufgaben

- (1) Erläutern Sie, welche Aspekte man in einem Evaluationsvertrag festhalten sollte und warum die Erstellung eines Evaluationsvertrags wichtig ist.

Musterantwort

Wichtigster Bestandteil eines Evaluationsvertrages ist die genaue Definition des Evaluationsauftrags. Im Vertrag sollte festgehalten werden, welche Leistung von wem bis zu welchem Zeitpunkt erbracht sein muss. Zusätzlich sollten alle Rechte und Pflichten aller Parteien im Vertrag definiert werden; hierzu gehören auch Geheimhaltungspflichten, Nutzungs- und Verwertungsrechte (bspw. der erhobenen Daten), Zahlungspflichten des Auftraggebers und Verfahrensschritte im Falle von Konflikten. Ein solcher Vertrag ist wichtig, um Missverständnisse möglichst zu verhindern und Probleme schnell zu erkennen und zu lösen.

- (2) Erläutern Sie in Grundzügen die vier Oberkategorien von Evaluationsstandards der Deutschen Gesellschaft für Evaluation und geben Sie jeweils ein Beispiel für einen Standard pro Kategorie.

Musterantwort

Die vier Oberkategorien lauten (1) Nützlichkeitsstandards (Die Evaluation soll den Informationsbedürfnissen aller Beteiligten Rechnung tragen; Beispiel: Ziel und Zweck der Evaluation müssen im Vorhinein transparent expliziert werden); (2) Durchführbarkeitsstandards (Die Evaluation soll hinsichtlich ihres Aufwands realistisch und angemessen sein und auf größtmögliche Akzeptanz sowie Effizienz abzielen; Beispiel: Kosten und Nutzen der Evaluation sollten in einem angemessenen Verhältnis zueinander stehen); (3) Fairnessstandards (Die Evaluation soll ethischen Richtlinien und moralisch relevanten Bedürfnissen der Betroffenen entsprechen; Beispiel: Die Teilnahme an einer Evaluationsstudie muss freiwillig sein); (4) Genauigkeitsstandards (Die Evaluation soll adäquate Informationen über den Evaluationsgegenstand erbringen und »state of the art« sein; Beispiel: Es sollten quantitative und qualitative Daten erhoben werden.).

- (3) Benennen Sie die vier Arten der Validität und erläutern Sie, was diese jeweils bedeuten. Geben Sie für jede Validitätsart ein Beispiel an, bei dem diese verletzt ist.

Musterantwort

Die vier Arten lauten (1) Interne Validität; (2) Konstruktvalidität; (3) Externe Validität; (4) Validität des statistischen Schlusses. Ein Beispiel für eine Verletzung der internen Validität wäre, wenn aufgrund eines nicht-experimentellen Designs nicht ausgeschlossen werden könnte, dass eine Erhöhung der Merkmalsausprägung auf einem Wirksamkeitskriterium (bspw. Kenntnisse aufgrund einer Schulung zum Thema X) nicht auf die Schulung, sondern auf andere Ursachen zurückgeht. Ein Beispiel für eine Verletzung der Konstruktvalidität wäre, wenn das Instrument zur Erfassung der »Kenntnisse« nicht nur Wissen, sondern auch Einstellungen, erfassen würde. Ein Beispiel für eine Verletzung der externen Validität wäre, wenn die Ergebnisse, die im Rahmen der Evaluation einer Maßnahme im Kontext A ermittelt wurden, nur auf Kontext A, aber nicht auf Kontext B, C oder D angewendet werden können. Ein Beispiel für eine Verletzung der Validität des statistischen Schlusses wäre, wenn zur Auswertung der Daten ein statistisches Verfahren verwendet würde, bei dem die Voraussetzungen im vorliegenden Fall nicht erfüllt sind.

- (4) Beschreiben Sie die drei Hauptgütekriterien für die Messung von Konstrukten.

Musterantwort

Die drei Hauptgütekriterien lauten (1) Objektivität – die erhobenen Daten sollten unabhängig davon sein, wer die Erhebung durchführt (Durchführungsobjektivität), wie die Daten ausgewertet werden (Auswertungsobjektivität) und wer die ausgewerteten Daten interpretiert (Interpretationsobjektivität); (2) Reliabilität – die Messung sollte möglichst frei von Messfehlern sein; (3) Konstruktvalidität – die erhobenen Daten sollten eindeutig interpretiert werden, d.h. es muss sichergestellt werden, dass auch wirklich das gemessen wurde, was gemessen werden sollte (und nichts anderes).

- (5) Benennen und erläutern Sie drei Nebengütekriterien für die Messung von Konstrukten.

Musterantwort

Drei (ausgewählte) Nebengütekriterien sind (1) die Ökonomie – die Messung sollte für alle Beteiligten möglichst ressourcenschonend sein; (2) die Unverfälschbarkeit – die Messung sollte von den Betroffenen nicht gezielt verfälscht werden können; (3) die Fairness – die Daten und deren Interpretation sollten nicht davon abhängen, von welcher ethnischen, religiösen, soziokulturellen oder sonstigen sozialen Kategorie eine Person gehört.