

Beurteilung der Facharztweiterbildung durch Ärzte in Weiterbildung anhand eines validierten Fragebogens

Evaluation of medical residency training programmes by residents using a validated questionnaire

Th. Prien¹ und M. Siebolds²

¹ Klinik und Poliklinik für Anästhesiologie und operative Intensivmedizin, Universitätsklinikum Münster (Direktor: Prof. Dr. H. Van Aken)

² Katholische Fachhochschule Nordrhein-Westfalen, Abteilung Köln, Fachbereich Gesundheitswesen

Zusammenfassung: Es gibt zahlreiche Gründe, die für eine Evaluation der Weiterbildung zum Facharzt durch Ärzte, die sich gegenwärtig in Weiterbildung befinden, sprechen. Hierzu zählen u.a. die Aufdeckung der Stärken und Schwächen eines Programms, die Dokumentation des Ausbildungsklimas, das rechtzeitige Erkennen von problematischen Entwicklungen, die Notwendigkeit, die Attraktivität von Weiterbildungsprogrammen auch für den kommenden ärztlichen Nachwuchs zu gewährleisten, sowie die kontinuierliche Qualitätssicherung. Eine derartige Evaluation wurde bislang jedoch weder im Bereich der deutschen Facharztweiterbildung noch in der Anästhesiologie eingeführt.

Das Ziel dieser Arbeit war, ein validiertes, zuverlässiges und praxistaugliches Instrument zur Evaluation der Anästhesie-Weiterbildung in Deutschland zu entwickeln.

Zu diesem Zweck wurde ein Fragenkatalog erarbeitet, der im Jahre 2002 in Form eines Fragebogens an 97 in Weiterbildung befindliche Ärzte in Deutschland (nicht-zufällige Stichprobe) ausgesandt wurde.

Die Entwicklung des Fragenkatalogs erfolgte nach Datenbankrecherche auf Basis der vorliegenden Literatur. Insgesamt wurden 72 Aussagen (Items) als Beurteilungskriterien formuliert und von den Teilnehmern auf einer 5-Punkte-Rating-Skala gewichtet; zusätzliche Items konnten angegeben werden. Die Items mit der höchsten Bedeutung sowie zusätzliche Items aus der Befragung wurden nach Validierung mittels kognitiver Interviews mit 5 in Weiterbildung befindlichen Ärzten zu einem 11-Item-Fragebogen für die Beurteilung des Weiterbildungsprogramms durch Ärzte, die sich in Weiterbildung befinden, und einem 19-Item-Fragebogen zur Beurteilung der Rotationen bzw. Ausbilder zusammengestellt. Zur Prüfung der Test-Retest-Reliabilität wurden diese beiden Fragebogen zweimal im Abstand von zwei Wochen von 14 Ärzten in Weiterbildung bearbeitet, wobei die Items wiederum auf einer 5-Punkte-Rating-Skala zu beurteilen waren.

Die Ergebnisse der Auswertung zeigen, dass vor allem die folgenden Items durch die Befragten als bedeutsam eingestuft worden sind (Angaben in absteigender

Reihenfolge): klinische Aspekte der Weiterbildung, Kernkompetenz der Ausbilder, Arbeitshygiene, didaktisches Verhalten der Ausbilder, durch die Ausbilder vermittelte Arbeitshygiene, professionelles Verhalten der Ausbilder, Prozessqualität der Weiterbildung, klinische Ausbildung durch weiterbildende Ärzte, Strukturqualität, äußere Faktoren, berufliche Perspektiven, Rahmenbedingungen der Beschäftigung, Ergebnisqualität und Strukturmerkmale der Weiterbildungsstätte. Etwa 3/4 (75%) der in Weiterbildung befindlichen Ärzte war der Meinung, dass man eher die einzelnen Ausbilder als das ganze Programm beurteilen sollte; gut 2/3 (70%) votierten für eine anonyme Auswertung.

Als Ergebnis dieser Studie liegt nun mit den beiden validierten und Test-Retest-stabilen Fragebogen ein praktisches Instrument zur flexiblen, situationsangepassten Evaluation der Weiterbildung zum Facharzt für Anästhesie durch die in Weiterbildung befindlichen Ärzte vor.

Summary: There are many reasons why an evaluation of medical residency programmes by residents who are currently receiving specialized medical training may be beneficial. These include, for example, the identification of the strengths and weaknesses of a programme, the documentation of the educational climate, an early recognition of problems, the need to ensure the attractiveness of the programme for the future generation of residents, and a continuous quality improvement. However, no formal evaluation of residency programmes has yet been introduced in Germany, neither for post-graduate medical training in general, nor for training in the field of anaesthesiology in particular. The objective of the present study was to develop a validated, reliable and practical evaluation instrument for the assessment of German training programmes in anaesthesiology.

Based on the information obtained from database searches for relevant literature, a set of questions was devised and presented to a non-random sample of 97 residents participating in medical residency programmes by means of a questionnaire mail survey conducted in Germany in 2002.

Altogether, 72 statements (items) were defined to serve as evaluation criteria for the assessment of residency programmes. These items were rated according to their relative importance on a 5-point Likert scale by the survey participants, who were also allowed to add new items. Those items that received the highest importance scores as well as the items that were included by the participating residents were validated by a cognitive interview with 5 residents and finally taken up into an 11-item list compiled for programme evaluation and into a 19-item list compiled for rotation or faculty evaluation, resp. In order to assess the test-retest reliability of these instruments, the two questionnaires were administered to 14 residents, who rated the individual items twice within a period of two weeks, again using a 5-point Likert scale.

The results obtained from these ratings show that the residents attach a particular importance to the following items (in descending order): clinical aspects of the residency programme, core competencies of the clinical teachers, departmental standards of work hygiene, didactic competence of the teachers, implementation and control of work hygiene standards by the teachers,

professional conduct of the teachers, procedural quality of the programme, clinical training provided by the faculty, structural quality of the programme, factors extrinsic to the training programme, career perspectives, basic conditions of employment, assessment of learning outcomes and student achievement, and aspects related to the location of the hospital providing the training. Approximately three-quarters (75%) of the residents believe that single rotations should be evaluated rather than the whole programme while two-thirds (70%) of the residents voted for an anonymous evaluation.

By the development of the two validated questionnaires that have been shown to be reliable in test-retest situations, the study has produced a flexible and stable instrument for the evaluation of anaesthesia residency programmes by residents.

Schlüsselwörter: Weiterbildung zum Facharzt – Programmevaluation – Fragebogen

Keywords: Medical Residency – Programme Evaluation – Questionnaire.

Einleitung

In den USA sehen die allgemeinen Anforderungen an die Anästhesiologie-Weiterbildungsprogramme unter anderem vor, dass diese regelmäßig durch Ärztinnen und Ärzte in Weiterbildung zur Fachärztin / zum Facharzt für Anästhesie (ÄiW) evaluiert werden. Ferner muss eine vertrauliche Evaluation der Ausbilder möglich sein (www.acgme.org).

Aus der Perspektive der Weiterbildungsberechtigten bzw. Abteilungsleiter sprechen die folgenden Gründe für eine Evaluation der Weiterbildung durch ÄiW: Aufdeckung der Stärken und Schwächen eines Programms, Dokumentation des Ausbildungsklimas, rechtzeitige Warnung vor problematischen Entwicklungen, Gewährleistung der Attraktivität von Weiterbildungsprogrammen für den ärztlichen Nachwuchs, kontinuierliche Qualitätssicherung. Insbesondere im Kontext einer Etablierung und Zertifizierung von Qualitätssicherungssystemen (z.B. EN ISO 9001, KTQ) ist eine regelmäßige Evaluation der Kundenzufriedenheit (hier der ÄiW) eine wesentliche Komponente. Aus Sicht der ÄiW sprechen vor allem die Chance zur Einflussnahme auf die Weiterbildung und die Möglichkeit, ihre Bedürfnisse mitzuteilen bzw. Kritik zu äußern für eine regelmäßige Weiterbildungs-evaluation.

In Deutschland wurde eine derartige Evaluation der Weiterbildung durch ÄiW bisher nicht durchgeführt. Es war daher das Ziel dieser Arbeit, ein Instrument zu entwickeln, mit dem die Zufriedenheit der ÄiW mit ihrer Weiterbildung regelmäßig gemessen werden kann. Das Evaluationsinstrument sollte validiert, zuverlässig und für möglichst viele Anästhesie-Weiterbildungsstätten in Deutschland geeignet sein.

Methodik

Mittels "Medline"- und "XMED"-Datenbankrecherche und Sichtung der Literaturlisten relevanter Arbeiten wurden insgesamt 62 relevante Publikationen identifiziert. Darauf basierend wurden zwei Listen mit Aussagen (Items) zur Beurteilung der Weiterbildung erstellt: eine Liste mit 41 Items, die Gesamtzufriedenheit der ÄiW mit der Weiterbildung betreffend, und eine Liste mit 31 Items, die Zufriedenheit mit dem Ausbilder einer Rotation betreffend. Die Items wurden in sinnverwandten Kategorien zusammengefasst und zu einem Fragebogen zusammengestellt, wobei jedes Item auf einer Rating-Skala nach Likert von 1 - 5 bewertbar war, mit den Eckwerten 5 (sehr wichtig) und 1 (weniger wichtig). Die Teilnehmer hatten auch die Möglichkeit, eigene Items hinzuzufügen. Ebenfalls erfragt wurden Strukturmerkmale der Weiterbildungsstätte und folgende soziodemographische Daten: Alter, Geschlecht, Weiterbildungsjahr, verheiratet / feste Bindung, Kinder, akademische Karriere geplant? Der Fragebogen wurde im 2. Quartal 2002 an 25 anästhesiologische Einrichtungen gesandt, deren Leiter dem Erstautor bekannt waren. Die erhaltenen Daten wurden mittels SPSS-Version 10.0.7 elektronisch ausgewertet. Zur deskriptiven Statistik wurden neben dem arithmetischen Mittel für jedes Item auch Schiefe und Exzess ermittelt. Schiefe- und Exzess-Daten zeigten keine Normalverteilungen. Daher wurde zur Prüfung auf statistisch signifikante Unterschiede in der Beurteilung einzelner Items durch Untergruppen von ÄiW der verteilungsunabhängige U-Test von Wilcoxon, Mann und Whitney verwendet.

Nach Auswertung dieses ersten "Gewichtungs"-Fragebogens wurde eine Itemreduktion vorgenommen, um

eine höhere Akzeptanz für eine serielle Anwendung des zu schaffenden Evaluationsinstruments zu erzielen. Dabei wurden folgende Prinzipien berücksichtigt:

- I) Es wurden die Items ausgewählt, die in Phase 2 von 97 ÄiW durchschnittlich am höchsten bewertet worden waren.
- II) Es wurden nur durch 5 ÄiW validierte Items ausgewählt (s.u.)
- III) Zusatzitems und Kommentare auf den Fragebögen wurden berücksichtigt.
- IV) Bei der Rotationsevaluation mussten alle 7 von *Irby* [3] angegebenen Dimensionen der klinischen Lehre durch Items berücksichtigt werden.

Die Validität der Items wurde mittels kognitiven Interviews mit 5 ÄiW nach dem von *Kurz et al.* [5] angegebenen Verfahren gesichert. Auf Grund dieser Validitätsprüfung wurden einzelne Items umformuliert oder gestrichen.

Nach Itemreduktion und Validitätsprüfung umfasste die Liste zur Beurteilung der Zufriedenheit mit der Weiterbildung 11 Items, die Liste zur Zufriedenheit mit einem Ausbilder 19 Items. Die Items wurden zufällig angeordnet, um Halo-Effekte infolge Kategorisierung ähnlicher Items zu verhindern.

Ein Pilotexemplar dieser Fragebogen wurde den ca. 60 ÄiW der Klinik und Poliklinik für Anästhesiologie und operative Intensivmedizin des Universitätsklinikums Münster zweimal im Abstand von zwei Wochen gegeben. Um die zu den beiden Testzeitpunkten bearbeiteten Fragebogen einem bestimmten ÄiW zuordnen zu können und gleichzeitig die Anonymität zu wahren, waren die Fragebogen mit einem nur für den Ausfüllenden erkennbaren Alias zu kennzeichnen. Die erhaltenen Daten wurden mittels SPSS-Version 10.0.7 elektronisch ausgewertet. Dazu wurde der Übereinstimmungsgrad zu den beiden Testzeitpunkten mit dem Spearman-Korrelationskoeffizienten und mittels *Cronbachs* Alpha-Statistik ermittelt.

Ergebnisse

Aus 13 Anästhesieabteilungen kamen zusammen 104 ausgefüllte Fragebogen zurück. Davon konnten 97 ausgewertet werden; 7 wurden nicht berücksichtigt, da sie erkennbar von Fachärzten bearbeitet worden waren. 2 Bogen konnten keinem Krankenhaus zugeordnet werden. Alle Rückläufe kamen entweder aus Universitätskliniken oder Lehrkrankenhäusern (Tab. 1). Kleine Weiterbildungsstätten sind damit in der Stichprobe unterrepräsentiert (Tab. 2).

Wie die ÄiW dieser Stichprobe die Items beurteilten ist übersichtsweise in Tab. 2 und 3 dargestellt. Die detaillierten Daten aller Items können via Internet (Dokument "Beurteilung der Weiterbildung"; <http://anaest.uni-muenster.de/homepages/prien>) eingesehen werden. Die deskriptive Statistik der 5 jeweils am höchsten und am niedrigsten bewerteten Items

Tabelle 1: Beschreibung der Stichprobe.

		N
Krankenhausart	Lehrkrankenhaus	39
	Universitätsklinik	56
	k.A.	2
Zahl der ÄiW	3 bis 5	15
	6 bis 10	18
	mehr als 20	64
Weiterbildungsjahr	erstes	14
	zweites	12
	drittes	17
	viertes	22
	fünftes	19
	> 5	9
	k.A.	4
Alter*	32,1 ± 3,41 (25 - 43)	
Geschlecht	weiblich	40
	männlich	55
	k.A.	2
verheiratet/ feste Bindung?	ja	29
	Nein	65
	k.A.	3
Kinder?	ja	32
	nein	61
	k.A.	4
akademische Karriere geplant?	ja	22
	nein	67
	k.A.	8

(k.A. = keine Angaben)
* Mittelwert ± Standardabweichung (Spannweite).

zeigt Tabelle 3. Schiefe- und Exzessdaten lassen erkennen, dass keine Normalverteilungen vorliegen.

Tabelle 3 zeigt die Bedeutung der Item-Kategorien in absteigender Reihenfolge. Diese wurde bestimmt, indem die Mittelwerte der Items einer jeden Kategorie ihrerseits gemittelt wurden. Die durchschnittlich höchste Bedeutung wurde danach den Items beigemessen, die das "klinische Angebot" beschreiben, also klinische Aspekte der Weiterbildung wie Patientenspektrum, Spektrum der durchgeführten Anästhesieverfahren, klinische Kompetenz der Ausbilder. Relativ große Bedeutung wurde auch den "arbeitshygienischen" Faktoren beigemessen. Relativ gering bewertet wurden dagegen die Merkmale der Strukturqualität sowie die Ergebnisqualität.

Etwa drei Viertel der Teilnehmer war der Meinung, dass man eher die einzelnen Ausbilder/Oberärzte als das ganze Programm beurteilen sollte; gut zwei Drittel votierten für eine anonyme Auswertung.

Eine Analyse der Itembewertungen in Abhängigkeit von den erfragten soziodemographischen Daten findet sich im Internet-Dokument (s.o.). Dort können auch die Ergebnisse der Validitätsprüfung und des Itemreduktionsprozesses eingesehen werden.

Tabelle 2: Bedeutungsanalyse. Statistische Kenngrößen der fünf am höchsten und am niedrigsten bewerteten Items.

	N	MW	SD	Schiefe	Exzess
Gutes Arbeitsklima	97	4,77	,44	-1,682	1,778
Ausbilder ist klinisch kompetent	97	4,74	,44	-1,125	-,750
Ausbilder lässt mich neue Verfahren und Techniken lernen	97	4,73	,45	-1,064	-,887
Klinische Kompetenz des Chefs und der Oberärzte	97	4,67	,55	-1,840	4,550
Anwendung vieler Verfahren und Techniken	97	4,67	,57	-1,895	4,372
...					
Promotionsmöglichkeit	95	2,89	1,31	,199	-1,054
Möglichkeit zur Teilzeitarbeit	97	2,86	1,27	-,001	-1,053
Hauseigene Kindertagesstätte	96	2,71	1,41	,261	-1,243
Möglichkeit zur wissenschaftlichen Tätigkeit (akademische Laufbahn)	93	2,63	1,29	,561	-,716
Patientenklientel (sozioökonomischer Status)	96	2,25	,98	,626	,294

MW = Mittelwert, SD = Standardabweichung

Tabelle 3: Bedeutung der Item-Kategorien.

Item-Kategorie	Mittelwert
Klinische Aspekte der Weiterbildung	4,67
Kernkompetenz der Ausbilder	4,63
Arbeitshygiene	4,49
Didaktisches Verhalten der Ausbilder	4,45
Arbeitshygiene vermittelt durch Ausbilder	4,43
Professionelles Verhalten der Ausbilder	4,33
Prozessqualität der Weiterbildung	4,24
Klinische Ausbildung durch weiterbildende Ärzte	4,11
Strukturqualität der Weiterbildung	3,90
Äußere Faktoren	3,73
Berufliche Perspektiven	3,33
Rahmenbedingungen der Beschäftigung	3,11
Ergebnisqualität der Weiterbildung	3,04
Aspekte der Weiterbildungsstätte	2,71

Für die Test-Retest-Reliabilitätsprüfung dieser beiden reduzierten Itemlisten (siehe Anhang) konnten insgesamt nur 14 bzw. 12 Fragebogenpaare berücksichtigt werden, die eindeutig einer Person zuzuordnen und vollständig ausgefüllt waren. Bei den Items zur Beurteilung der Weiterbildungsstätte variierten die Korrelationskoeffizienten nach *Spearman* zwischen 0,53 und 0,96 (beidseitiges p jeweils $< 0,05$), die Alpha-Statistiken nach *Cronbach* zwischen 0,70 und 0,94. Unter den eingeschränkten Voraussetzungen weisen diese Kennziffern auf eine hohe Test-Retest-Reliabilität des Befragungsinstruments hin. Bei den Items zur Beurteilung der Rotationen/Ausbilder variierten die Korrelationskoeffizienten nach *Spearman* zwi-

schen 0,59 und 0,97 (beidseitiges p jeweils $< 0,05$), die Alpha-Statistiken nach *Cronbach* zwischen 0,77 und 0,98. Das Item "...arbeitet effektiv mit Pflegenden u. Operateuren zusammen" erwies sich nicht als Retest-stabil und wurde daher für den endgültigen Fragebogen gestrichen.

Diskussion

Nach Lage der Literatur haben sich vor allem die US-amerikanischen Psychiater in der AADPRT (American Association of Directors of Psychiatry Residency Training) mit der Evaluation der Weiterbildung beschäftigt [7]. Eine ausführliche Darstellung der theoretischen Konzepte, des Standes der Untersuchungen, der bisher entwickelten Instrumente und der praktischen Erfahrungen findet sich im Internet-Dokument (s.o.). Im deutschen Sprachraum wurde eine systematische Evaluation der Weiterbildung durch ÄiW bisher nicht durchgeführt. Auch findet sich in der internationalen Literatur kein entsprechendes Instrument für das Fachgebiet Anästhesiologie. Als Ergebnis dieser Arbeit liegen nun zwei validierte und zuverlässige Fragelisten vor, die für die Evaluation der Weiterbildung durch Ärztinnen und Ärzte in Weiterbildung zur Fachärztin/zum Facharzt für Anästhesie (ÄiW) verwendet werden können (siehe Anhang), eine zur Evaluation des gesamten Weiterbildungsprogramms und eine zur Beurteilung einzelner Ausbilder. Die beiden Instrumente können flexibel genutzt werden, an die Gegebenheiten der einzelnen Weiterbildungsstätte angepasst.

Bei der Entscheidung, ob man diese Instrumente einsetzen möchte, muss man möglichen Nutzen (siehe Einleitung) und Aufwand gegeneinander abwägen, darf dabei die methodenimmanenten Probleme und

Anlage 1: Fragebogen zur Evaluation des Weiterbildungs-Programms.

Wenn Sie an die Weiterbildung (WB) an Ihrer anästhesiologischen Einrichtung denken, wie bewerten Sie die folgenden Aussagen?

¹Ausbilder = Leitender Arzt, Oberärzte, ggf. Fachärzte

	trifft voll zu	trifft eher zu	unentschieden	trifft eher nicht zu	trifft gar nicht zu
Das Arbeitsklima ist gut	⑤	④	③	②	①
Die Ausbilder ¹ sind klinisch kompetent.....	⑤	④	③	②	①
Die angewandten Verfahren und Techniken sind zeitgemäß..	⑤	④	③	②	①
Die Ausbilder ¹ behandeln und beurteilen mich fair.....	⑤	④	③	②	①
Es gibt ausreichende klinische Ausbildungsmöglichkeiten für die einzelnen Inhalte der Weiterbildung.....	⑤	④	③	②	①
Die Ausbilder ¹ behandeln mich respektvoll.....	⑤	④	③	②	①
Es gibt ein Teamgefühl in der Anästhesieabteilung.....	⑤	④	③	②	①
Ich werde meiner klinischen Kompetenz entsprechend eingesetzt (keine Unter- bzw. Überforderung)	⑤	④	③	②	①
Die Supervision ist ausreichend (in qualitativer und quantitativer Hinsicht).....	⑤	④	③	②	①
Ich werde ausreichend für Weiterbildung freigestellt.....	⑤	④	③	②	①
Das WB-Programm reagiert ausreichend auf Feedback/ Kritik von den WBsteilnehmern.....	⑤	④	③	②	①

Bemerkungen:

Grenzen nicht außer Acht lassen. Ein "messbarer" Nutzen einer derartigen Evaluation wurde bislang nur für die Ausbilderbeurteilung nachgewiesen, ist aber nicht unumstritten.

Dominiert wird die Evaluationsthematik von der Anonymitätsfrage: Sollen die beurteilenden ÄiW grundsätzlich anonym bleiben oder nicht? In den USA sind anonyme Beurteilungen die Regel. Anonymität wird dabei als Voraussetzung für eine ehrliche Beurteilung angesehen. Auch in dieser Befragung votierten 70% der befragten ÄiW für die Anonymität. Auf Seiten der beurteilten Ausbilder ist allerdings zu befürchten, dass unangemessene Schlechtbeurteilungen erfolgen können, ohne dass dem Ausbilder die Chance gegeben wird, diese zu korrigieren. Um dem vorzubeugen, wird bei anonymisierten Systemen immer eine aggregierte Auswertung vorgenommen, in die die Beurteilungen einer Mindestzahl von ÄiW eingehen.

Die Anonymitätsfrage ist bei kleinen Anästhesieabteilungen mit wenigen ÄiW von besonderer Brisanz. Letztlich ist die Wahrung der Anonymität nur bei

Weiterbildungsstätten mit mehr als etwa fünf ÄiW möglich, und das auch nur bei Beurteilung des gesamten Programms. Eine anonyme, zeitnahe Beurteilung einzelner Ausbilder bzw. Rotationen ist auch an großen Weiterbildungseinrichtungen unmöglich, da der Ausbilder weiß, wer ihn beurteilt hat, wenn nur ein bis zwei ÄiW gleichzeitig eine Weiterbildungsrotation beenden. Eine anonyme Ausbilderbeurteilung ist daher nur zeitfern möglich, indem Einzelbeurteilungen gesammelt und aggregiert ausgewertet werden. Von Nachteil ist dabei natürlich die verlängerte Reaktionszeit.

Trotzdem sollte auf die Evaluation einzelner Ausbilder bzw. Rotationen nach Möglichkeit nicht verzichtet werden, denn die Zufriedenheit der ÄiW wird entscheidend durch das persönliche Engagement der Ausbilder, deren klinische, didaktische und menschlichen Kompetenzen und Qualitäten determiniert [4]. Auch in dieser Befragung meinten 75% der befragten ÄiW, man solle eher die einzelnen Ausbilder als das gesamte Programm beurteilen. Dabei ist die Evaluation der Rotationen bzw. Ausbilder für die Führung einer klinischen Abteilung, in der fachärztliche Weiter-

Anlage 2: Fragebogen zur Evaluation einer Weiterbildungsrotation bzw. eines Ausbilders.

Wenn Sie an die zurückliegende Rotation denken, wie bewerten Sie die folgenden Aussagen?

Rotation = Einsatz und Ausbildung in einem bestimmten Bereich Ihrer anästhesiologischen Einrichtung

Ausbilder = die/der während der Rotation für ihre Weiterbildung zuständige Ärztin/Arzt (z.B. Bereichsleiterin/-arzt)

Die Ausbilderin/der Ausbilder...	trifft voll zu	trifft eher zu	unentschieden	trifft eher nicht zu	trifft gar nicht zu
... ist klinisch kompetent.....	5	4	3	2	1
... beurteilt mich fair.....	5	4	3	2	1
... ist fähig, klinische Entscheidungsprozesse zu vermitteln.....	5	4	3	2	1
... behandelt mich mit Respekt.....	5	4	3	2	1
... schafft ein gutes Arbeitsklima.....	5	4	3	2	1
... verfügt über breites Hintergrundwissen.....	5	4	3	2	1
... erläutert verständlich und ausreichend.....	5	4	3	2	1
... geht gut mit den Patienten um.....	5	4	3	2	1
... gibt konstruktive Rückkopplung über meine Leistungen.....	5	4	3	2	1
... hat Spaß an der Ausbildung (Enthusiasmus).....	5	4	3	2	1
... ist selbstkritisch, erkennt eigene Grenzen.....	5	4	3	2	1
... zeigt aktives Lehrverhalten (reagiert nicht nur auf Fragen)....	5	4	3	2	1
... setzt mich meiner klinischen Kompetenz entsprechend ein (keine Unter- bzw. Überforderung).....	5	4	3	2	1
... schafft eine Atmosphäre, die das Lernen erleichtert.....	5	4	3	2	1
... reagiert positiv auf Feedback/ Kritik von mir.....	5	4	3	2	1
... definiert die Anforderungen an mich klar.....	5	4	3	2	1
... verhält sich unbedrohlich.....	5	4	3	2	1
... überwacht mich ausreichend, hat genügend Zeit für mich..	5	4	3	2	1

Bemerkungen:

bildung stattfindet, von besonderer Brisanz. Da in der augenblicklichen deutschen Weiterbildungsrealität die Ausbildung der ÄiW ja nur eine von mehreren Aufgaben der damit beauftragten Ärzte (in der Regel Oberärzte) ist, muss sich der Weiterbildungsbefugte mit der Frage auseinandersetzen, wie er mit Schlechtleistungen umgehen will, wenn der betreffende Ausbilder ansonsten gute Leistungen erbringt.

Die beiden entwickelten Fragebogen sind ein Kompromiss zwischen Vollständigkeit und Praktikabilität: Der Anspruch, möglichst viele Faktoren und Spezialinteressen zu erfassen, würde die Berücksichtigung möglichst vieler Items erfordern, die Anforderungen der Praxis zwingen zur Beschränkung. Daher wurde im ersten Entwicklungsschritt ermittelt, welche Faktoren den ÄiW besonders wichtig sind. Darauf basierend wurden dann Items eliminiert, die für die Ge-

samtheit im Durchschnitt weniger wichtig waren. Bei diesem Vorgehen werden Items eliminiert, die für Subgruppen von größerer Bedeutung sind als für das Gesamtkollektiv. Diese banale Überlegung wird durch die Ergebnisse der Subgruppenanalyse, auf deren Darstellung hier aus Platzgründen verzichtet wurde, belegt.

Einige Beispiele:

- Lokale Spezifika können dazu führen, dass einzelne Items an dieser Weiterbildungsstätte eine andere Bedeutung haben als im Durchschnitt.
- Unterschiedliche Karriereplanungen beeinflussen die Gewichtung bestimmter Items.
- Mit zunehmender Weiterbildungsdauer nimmt die Bedeutung der Weiterbeschäftigung am Ende der Weiterbildung zu.
- Frauen haben einen höheren Bedarf an geringeren

- Arbeitszeiten, wobei spekuliert werden darf, dass dies wohl durch Kinderversorgung bedingt ist.
- Verheiratete bzw. fest gebundene ÄiW sind weniger mobil als ledige ÄiW. Daher sind neben der Zufriedenheit des Lebenspartners am Ort der Weiterbildung auch Faktoren von besonderer Bedeutung, die eine komplette Weiterbildung an einem Ort ermöglichen, ferner die Weiterbeschäftigung nach Abschluss der Weiterbildung.

Um eine Erfassung von Subgruppenbelangen nicht völlig auszuschließen, wird den Befragten im endgültigen Fragebogen die Möglichkeit gegeben, in einem Freitextteil nicht abgefragte Aspekte anzusprechen. Dabei gilt es zu bedenken, dass Freitextpassagen immer die Gefahr einer Anonymitätsaufhebung und damit potentiell eine datenschutzrechtliche Problematik bergen.

Es liegt in der Natur einer Befragung, dass damit nur subjektive Beurteilungen erfasst werden, im Gegensatz zu objektiven Markern. Überspitzt formulieren Kritiker, dass es sich um einen Popularitätswettbewerb [6] handelt, bei dem die angenehmsten Ausbilder in Folge eines Halo-Effektes die höchsten Wertungen bekommen. Auch ist zu berücksichtigen, dass eine einzelne unobjektive Schlechtbewertung die Durchschnittsbewertung stark beeinflussen kann, wenn insgesamt nur wenige Beurteilungen vorliegen [1]. Diesem Einwand muss entgegeng gehalten werden, dass die "Kundenzufriedenheit" ja ein Qualitätskriterium ist, in diesem Fall die Zufriedenheit der ÄiW mit dem Weiterbildungsprogramm bzw. den Ausbildern. Kundenbefragungen werden daher auch von der EN ISO 9001 gefordert. Zudem haben Befragungen den Vorteil, dass sie ein Urteil von vielen Individuen zu einer Vielzahl von Aspekten und niedrigen Kosten bieten.

Von nicht unerheblicher Bedeutung ist auch, ob die ÄiW zur Beurteilung innerhalb eines Weiterbildungsprogramms verpflichtet werden oder ob sie dies freiwillig tun. Innerhalb eines Weiterbildungsprogramms könnte ja die regelmäßige Evaluation sowohl des gesamten Programms als auch der einzelnen Rotationen etabliert werden. Dies ist eine Führungsentscheidung, denn die Erfassung der ÄiW-Zufriedenheit mit der Weiterbildung und vor allem der Umgang der in Weiterbildung befindlichen Ärzte mit den erhobenen Daten kann die Kultur an der Weiterbildungsstätte erheblich beeinflussen. Dies verdeutlichen die Ängste der ÄiW, die diese gegenüber einer offenen, nicht-anonymen Beurteilung haben. Kann man die ÄiW zur Evaluation der Weiterbildung verpflichten? Wenn dieses Vorgehen unter allen Beteiligten konsentiert ist, spricht grundsätzlich nichts dagegen, obwohl in der Praxis die Umsetzung erheblich erschwert sein dürfte, wenn man die Evaluation anonym vornimmt. Soll man die ÄiW zur Evaluation der Weiterbildung verpflichten? Wird der Kunde zur Evaluation eines Produktes verpflichtet, dann besteht die Gefahr, dass die Evaluation unter Zwang nicht nach bestem Wissen und

Gewissen erfolgt. Eine innere Bereitschaft zur Evaluation ist Voraussetzung für verwertbare Aussagen. Die Evaluation sollte daher grundsätzlich freiwillig sein.

Eine weitere wichtige Grenze für solche Instrumente wie die hier entwickelten Fragebogen ist der Mangel an Benchmarking-Daten. Es kann also keinerlei Aussage darüber getroffen werden, ob ein durchschnittlicher Zufriedenheits-Score von beispielsweise "4,1" nun gut oder schlecht ist. Die subjektive Zufriedenheit ist nämlich erheblich von den Erwartungen determiniert [2], und diese können unterschiedlich sein bzw. sich ändern. In diesem Sinne hat der Fragebogen - "nur" - eine Indikatorfunktion. Liegt eine Schlechtbeurteilung vor oder tritt eine graduelle Verschlechterung im Zeitverlauf ein, sollte dies zum Anlass genommen werden, den Ursachen nachzugehen.

Dank

Unser Dank gilt den Kolleginnen und Kollegen in Weiterbildung, die mit ihren Angaben zur Entwicklung des Fragebogeninstrumentes mit beigetragen haben.

Literatur

1. Cohan RH, Dunnick NR, Blane CE, Fitzgerald JT. Improvement of faculty teaching performance: efficacy of resident evaluations. *Acad Radiol* 1996;3:63-67.
2. Herzberg F. *Work and the nature of man*. New York: Cleveland; 1966.
3. Irby DM. Clinical teacher effectiveness in medicine. *J Med Educ* 1978;53:808-815.
4. Irby DM, Gillmore GM, Ramsey PG. Factors affecting ratings of clinical teachers by medical students and residents. *J Med Educ* 1987;62:1-7.
5. Kurz K, Prüfer P, Rexroth M. Zur Validität von Fragen in standardisierten Erhebungen. Ergebnisse des Einsatzes eines kognitiven Pretestinterviews. *ZUMA-Nachrichten* 1999;44:83-107.
6. McLeod PJ, James CA, Abrahamowicz M. Clinical tutor evaluation: a 5-year study by students on an in-patient service and residents in an ambulatory care clinic. *Med Educ* 1993;27:48-54.
7. Yudkowsky R, Elliot R, Schwartz A. Two perspectives on the indicators of quality in psychiatric residencies: Program directors' and residents'. *Acad Med* 2002;77:57-64.

Korrespondenzadresse:

Prof. Dr. med. *Thomas Prien*, M.Sc.
Klinik und Poliklinik für Anästhesiologie und operative Intensivmedizin
Universitätsklinikum Münster
Albert-Schweitzer-Straße 33
D-48129 Münster.