

Aus der Poliklinik für Zahnerhaltung und Parodontologie
Ludwig – Maximilians – Universität München
Direktor: Prof. Dr. med. dent. Reinhard Hickel

Evaluation von Bewertungskriterien für praktische Studentenarbeiten im Vergleich zur Bewertung per Augenschein

Dissertation
zum Erwerb des Doktorgrades der Zahnmedizin
an der Medizinischen Fakultät
der Ludwig-Maximilians-Universität zu München

Vorgelegt von
Moritz Peter Baumann
geboren in Bad Homburg
2015

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatterin: Prof. Dr. med. dent. Karin Christine Huth

Mitberichterstatterin: Priv. Doz. Dr. Inga Hege

Dekan: Prof. Dr. med. Dr. h.c. Maximilian Reiser,
FACR, FRCR

Tag der mündlichen Prüfung: 26.03.2015

FÜR MEINE LIEBE FAMILIE

Inhaltsverzeichnis

1	EINLEITUNG	6
1.1	BEGRÜNDUNG DES BEDARFS.....	6
1.2	LITERATURÜBERSICHT ZUM THEMA RELIABILITÄT VERSCHIEDENER BEWERTUNGSMETHODEN IN DER ZAHNMEDIZINISCHEN AUSBILDUNG	8
1.3	VORARBEITEN.....	15
1.3.1	Literaturübersicht zu Kriterien für praktische Arbeiten im Phantomkurs der Zahnerhaltung.....	15
1.3.2	Beschreibung der praktischen Aufgabenstellungen	30
1.3.3	Nominal group technique.....	31
1.3.4	Matrixanalyse.....	36
1.3.5	Definition von Durchfallkriterien	42
1.3.6	Erstellung eines Notenschlüssels	45
2	FRAGESTELLUNG.....	46
3	MATERIAL UND METHODE	47
3.1	ORT UND ZEITRAUM DER STUDIENDURCHFÜHRUNG	47
3.2	ÜBERPRÜFUNG DER VERSTÄNDLICHKEIT DER BEWERTUNGSKRITERIEN	47
3.3	VERGLEICH ZWISCHEN DER BEWERTUNG MIT DEN ERSTELLTEN KRITERIEN UND DER BEWERTUNG PER AUGENSCHHEIN	48
3.4	FRAGEBOGEN FÜR STUDIERENDE UND AUSBILDER ZUM THEMA EVALUATION OBJEKTIVER BEWERTUNGSKRITERIEN FÜR PRAKTISCHE STUDENTENARBEITEN.....	48
3.5	DATENANALYSE	49
4	ERGEBNISSE.....	51
4.1	ÜBERPRÜFUNG DER VERSTÄNDLICHKEIT DER BEWERTUNGSKRITERIEN	51
4.2	RELIABILITÄT DER BEWERTUNGEN PER AUGENSCHHEIN UND PER KRITERIEN DURCH DIE AUSBILDER	53
4.3	RELIABILITÄT DER BEWERTUNGEN PER AUGENSCHHEIN UND PER KRITERIEN DURCH DIE STUDIERENDEN.....	55
4.4	GEGENÜBERSTELLUNG DER ABWEICHUNGEN VON AUSBILDERN UND STUDIERENDEN	57
4.5	FRAGEBOGENAUSWERTUNG	60
5	DISKUSSION	89
5.1	VORARBEITEN.....	89
5.1.1	Nominal group technique.....	89
5.1.2	Matrix-Analyse	91
5.2	DISKUSSION DER METHODE.....	92
5.2.1	Zeitraum der Bewertung	92

5.2.2	Fragebogen.....	92
5.2.3	Datenanalyse.....	93
5.3	DISKUSSION DER ERGEBNISSE	93
6	ZUSAMMENFASSUNG.....	99
7	LITERATURVERZEICHNIS.....	101
8	ABBILDUNGSVERZEICHNIS.....	108
9	TABELLENVERZEICHNIS.....	111
10	ABKÜRZUNGEN.....	113
11	ANHANG.....	114
11.1	AUFGABENSTELLUNG FÜR PRÜFUNG 1	114
11.2	AUFGABENSTELLUNG FÜR PRÜFUNG 2	114
11.3	AUFGABENSTELLUNG FÜR PRÜFUNG 3	114
11.4	AUFGABENSTELLUNG FÜR PRÜFUNG 4	115
11.5	AUFGABENSTELLUNG FÜR PRÜFUNG 5	115
11.6	AUFGABENSTELLUNG FÜR PRÜFUNG 6	115
11.7	FRAGEBOGEN STUDIERENDE	117
11.8	FRAGEBOGEN AUSBILDER.....	121
12	DANKSAGUNG	127
13	LEBENS LAUF	128

1 Einleitung

1.1 Begründung des Bedarfs

In der Ausbildung der Zahnmedizinstudenten spielt die Beurteilung der praktischen Leistungen der Studierenden eine große Rolle. Bevor die Studierenden klinisch am Patienten tätig werden, müssen sie vier praktische Kurse am Phantomkopf absolvieren, um die theoretischen und manuellen Fähigkeiten und Fertigkeiten zu erlernen. Drei Kurse finden im vorklinischen Studienabschnitt statt, ein Kurs im klinischen Studienabschnitt. Dieser ist der Phantomkurs der Zahnerhaltung, Parodontologie und Kinderzahnheilkunde. Von besonderer Wichtigkeit in diesen Kursen ist ein Benotungssystem das objektiv und reproduzierbar ist.

Es gibt unterschiedliche Möglichkeiten die Arbeiten zu bewerten. Zum einen die „glance and grade“ Methode (Vann *et al.* 1983; Jenkins *et al.* 1998) und zum anderen anhand von definierten Bewertungskriterien (Dhuru *et al.* 1978). Es tauchen in der Literatur auch Begriffe auf wie „global rating“ auf der einen Seite und OSCE/OSPE (objective structured clinical/ practical evaluation) auf der anderen Seite (Jünger 2006).

Wird eine Bewertung anhand von „glance and grade“ bzw. per Augenschein vorgenommen, hängt diese von der individuellen Einschätzung des jeweiligen Bewerter ab (Lilley *et al.* 1968; Fuller 1972; Gaines *et al.* 1974; Feil 1982). Dieser Ermessensspielraum kann unterschiedliche Ursachen wie Ausbildung und Erfahrung haben. Des Weiteren sind persönliche Vorlieben und Abneigungen bei der Benotung nicht auszuschließen. Für den Studierenden bedeutet dies oft eine Intransparenz der Benotung und subjektive Verunsicherung hinsichtlich der eigenen Leistung. Auch im Falle einer Verteidigung der Benotung in einer juristischen Auseinandersetzung zwischen Studierendem und Lehrkörper ist eine Benotung per Augenschein schwieriger nachzuvollziehen.

Im Gegensatz zu der Bewertung per Augenschein kann die Bewertung allerdings auch durch Bewertungskriterien, die dem Studierenden schon zu Beginn des Kurses vorliegen, durchgeführt werden. Diese Art der Bewertung soll die Objektivität der Beurteilung durch den ausbildenden Zahnarzt erhöhen und die Selbsteinschätzung des Studenten schulen (Türp 2002).

Die Bewertungskriterien müssen verschiedene Aspekte erfüllen. Zum einen müssen sie eine klinische Relevanz haben, objektiv und reproduzierbar innerhalb eines Bewerbers und zwischen verschiedenen Bewertern sein sowie praktikabel in der täglichen Anwendung (Feil 1982; Buckley *et al.* 2009; Macluskey *et al.* 2011; Usha *et al.* 2012). Vorteile einer Bewertung mit vordefinierten Bewertungskriterien sollen vor allen Dingen Transparenz der Benotung, gesteigerte Objektivität und Vergleichbarkeit der Bewerter (Reliabilität, Verminderung des Prüferinflusses) und nicht zuletzt ein Training der Selbsteinschätzung der praktischen Leistungsfähigkeit durch die Studierenden selbst sein (Jünger 2006). Letzteres erscheint unabdingbar für die Reflexion und Erkennung von Kompetenzdefiziten, um während des Studiums eigene Defizite durch wiederholtes formatives Feedback auszuräumen (Jünger 2006). Dies kann ebenso eine mögliche durch individuelle Veranlagung vorhandene Selbstüber- bzw. unterschätzung der Studierenden mindern. Dies dient im klinischen Alltag dem Wohle des Patienten.

Folgerichtig wurde bei der Literaturrecherche über die Reliabilität der Beurteilung klinischer Arbeiten klar, dass das häufig angewandte Benotungssystem per Augenschein eine unbefriedigende Objektivität und Reproduzierbarkeit hervorbringt. In einer Studie von 1998, in der 75 Präparationen von fünf Bewertern unterschiedlicher Ausbildung und Erfahrung anhand des Bewertungssystem per Augenschein benotet wurden, wird die intra- und interpersonelle Variabilität beschrieben (Jenkins *et al.* 1998). Das Ergebnis dieser Studie zeigte, dass im Hinblick auf die interpersonelle Reliabilität im Durchschnitt eine Abweichung von 5-22% in der Benotung vorlag. Das Niveau der Abweichung hing vor allem von der Erfahrung des Prüfers ab. Ebenso wurde deutlich, dass selbst der erfahrenste Bewerter in Bezug auf die intrapersonelle Reliabilität eine Variabilität von 17% in seiner Benotung aufwies. Aus dieser Tatsache wurde abgeleitet, dass ein umfassendes und gut definiertes Bewertungssystem geschaffen werden muss, um die Objektivität zu steigern.

Eine weitere Studie aus dem Jahr 2006 befasste sich mit der Variabilität der Benotung und vergleicht die Bewertung durch „glance and grade“ mit der einer Checkliste (Sharaf *et al.* 2007). Dabei präparierten 30 Zahnmedizinstudenten Kavitäten in Kunststoffzähne, die im Folgenden von drei Zahnärzten der Abteilung mit mehr als zehn Jahren Berufserfahrung bewertet wurden. Die Ergebnisse dieser Studie zeigten, dass eine deutliche Abweichung innerhalb der Bewerter vorlag im

Vergleich der beiden Bewertungsarten „glance and grade“ und Checkliste. Zudem war für die Checkliste noch immer eine unbefriedigende Variabilität bezüglich der inter- und intra-personellen Reliabilität festzustellen. Die Autoren interpretierten dies so, dass weiterhin das Problem der Beständigkeit und Variabilität der Benotung vorliegt und weitere Anstrengungen betrieben werden sollten, ein Bewertungssystem mit klar definierten Bewertungskriterien zu entwickeln.

Es war das Ziel dieser Studie, die für die praktischen Studentearbeiten im Phantomkurs der Zahnerhaltung erarbeiteten Bewertungskriterien hinsichtlich ihrer Reproduzierbarkeit durch Studierende und Ausbilder im Gegensatz zu einer Bewertung per Augenschein zu vergleichen.

1.2 Literaturübersicht zum Thema Reliabilität verschiedener Bewertungsmethoden in der zahnmedizinischen Ausbildung

Das Ergebnis der Literaturrecherche hat ergeben, dass es nur eine kleine Anzahl von Studien über die Exaktheit der Beurteilung von Arbeiten, die durch Studierende angefertigt wurden, gibt. Die Mehrzahl der Veröffentlichungen stammt aus dem englischsprachigen Raum. Aus dem deutschsprachigen Raum gibt es lediglich die Veröffentlichungen von Türp und dessen Kollegen (Türp 2002) und von Kellersmann (Kellersmann 2008), die sich mit diesem Problem befassten. Weitere Veröffentlichungen aus dem angloamerikanischen Raum haben andere Ausbildungssysteme als Grundlage.

Gegenstand der Untersuchung von Türp und seinen Kollegen (Türp 2002) war es die Größenordnung der Abweichung einer Benotung von prothetischen Arbeiten von Studierenden zu untersuchen. Es handelte sich in dieser Studie um Arbeiten eines Phantomkurses. Während dieser Untersuchung wurden 20 Brücken und 20 Interimsprouthesen von drei klinischen Studierenden und drei Zahnärzten untersucht. Die Intention war, die Größenordnung der interindividuellen Streuung bei der Beurteilung der erbrachten Leistungen darzustellen. Folgende Aspekte sollten dabei besonders beachtet werden: die Breite des Notenspektrums, welche der beiden Bewertergruppen benotete strenger und wovon waren die unterschiedlichen Noten der Studienteilnehmer abhängig. (Kellersmann 2008). Die Studie zeigte, dass die an

der Studie teilnehmenden Assistenzärzte bei beiden zu bewertenden Arbeiten strenger als die Studierenden des klinischen Studienabschnitts benoteten.

Kellersmann (Kellersmann 2008) beschäftigte sich 2008 in seiner Veröffentlichung mit der „Reliabilität der Beurteilung vorklinischer Phantomarbeiten bei Einsatz eines strukturierten Bewertungsbogens“ (Kellersmann 2008). Es wurde beschrieben, dass durch die Verwendung des Bewertungsbogens sowohl die Objektivität, als auch die Reproduzierbarkeit der Benotung erhöht werden konnten. Weiter wurde erklärt, dass die Gruppe der Zahnärzte, die die klinischen Behandlungskurse betreuten, zum einen bei der inter-, als auch bei der intrapersonellen Gegenüberstellung die stärkste Übereinstimmung in ihrer Bewertung hatten. Die geringste wurde hingegen bei den Studierenden gefunden. Kellersmann formulierte daraus, dass die erhobenen Ergebnisse beweisen, dass mit der Verwendung des Bogens für die Bewertung die Sicherheit der Benotung zahnmedizinischer Arbeiten deutlich erhöht werden konnte.

Bereits im Jahr 1967 untersuchten Natkin und Guild (Natkin und Guild 1967) die Übereinstimmung der Bewertung studentischer Arbeiten. Es wurden 65 endodontische Studentenbehandlungen extrahierter Zähne bewertet. Sechs Bewerter benoteten die Arbeiten und gaben einen Kommentar ab. Die Bewertung erfolgte nach dem klassischen Verfahren per Augenschein („glance and grade“) (Natkin und Guild 1967). Es ergaben sich klar zu verzeichnende Ungleichheiten in der Benotung. Knapp die Hälfte der befundeten Zähne erzielte für die gleiche Arbeit Noten mit einem maximalen Unterschied von vier Notenstufen. Lediglich bei 8% der Arbeiten lag die Notendifferenz bei weniger als einer Note.

Ein Jahr später überprüften Lilley und Mitarbeiter (Lilley *et al.* 1968) wie die Übereinstimmung der Beurteilung erhöht werden kann. Dabei war von Interesse ob durch die Verwendung von notenwirksamen Kriterien und Fehlerdefinition sowohl die inter- als auch die intraindividuelle Übereinstimmung gefördert werden kann. In dieser Studie wurden 37 anonymisierte Modelle mit Studentearbeiten von drei Bewertern dreimal in einem Abstand von einem Monat benotet. Zu untersuchen waren Amalgamfüllungen, die mit Bewertungen von A bis E (Lilley *et al.* 1968) versehen wurden. Die Bestehensgrenze wurde zwischen C und D festgelegt. Vor dem zweiten Bewerten erörterten die Probanden für jede Stufe Bewertungskriterien, sowie Fehler, die zum sofortigen Durchfallen führten. Anhand der erhobenen Daten erkannte man, dass die interindividuelle Variabilität bezogen auf die Kavität im ersten

Durchgang sehr hoch war (Intraklassenkorrelationskoeffizienten 0,12 – 0,28). Der Intraklassenkorrelationskoeffizient (ICC) ist ein Wert zwischen null und eins. Null ist ein schlechter Wert, eins hingegen ein guter Wert. Diese Variabilität verringerte sich jedoch deutlich nach der Diskussion, so dass ein ICC von 0,51- 0,63 nach dem dritten Durchgang errechnet werden konnte.

Im Jahr 1972 beschäftigte sich Fuller (Fuller 1972) mit der Reliabilität der Bewertung und den Einflussmöglichkeiten auf eine Bewertung einzuwirken. Es wurde die zu dieser Zeit weitverbreitete Bewertungsmethode per Augenschein und die durch Checklisten angewandt. Im Rahmen der Studie sollte gezeigt werden, dass durch gezielte Schulung die Übereinstimmung gesteigert werden könnte. Vier Benotungsgruppen bewerteten die praktischen Arbeiten von drei Prüfungen der 67 Erstsemester-Studenten der Universität Iowa. Die Bewerter hatten zur Aufgabe Präparationen zu beurteilen. Die Bewertung erfolgte in drei Runden. Der erste Durchgang wurde per Augenschein durch alle Personen separat durchgeführt. Im Anschluss an diese Runde wurden erneut 25 zufällig ausgewählte Arbeiten zur Feststellung der intraindividuellen Reliabilität begutachtet. Im Anschluss wurden vier Zweier Gruppen gebildet, die mit unterschiedlichen Methoden die zuvor ausgewählten Arbeiten bewerteten. In der ersten Gruppe wurde weiterhin lediglich nach Augenschein begutachtet. Die weiteren Gruppen hatten Checklisten und zum Teil ein Training zuvor zur Verfügung. Die Benotung erfolgte durch Punkte. Die Maximale Punktzahl betrug 50 Punkte. Anhand des Ergebnisses wurden die Arbeiten unterteilt. Das Ergebnis machte deutlich, dass die Schulung und Checklisten nicht zu der erhofften Steigerung der Reliabilität führten. Zu erkennen war, dass die höchste Übereinstimmung zwischen den Bewertungsdurchgängen bei den erfahrensten Assistenten (ICC von 0,831) erreicht wurde. Erwartungsgemäß konnte die niedrigste Reliabilität bei den Assistenten mit der geringsten Erfahrung gemessen werden (ICC von 0,663 und von 0,472).

Im folgenden Jahr setzten sich Hinkelman und Long (Hinkelman und Long 1973) mit einer vergleichbaren Problematik, der Beurteilung zahnärztlicher Arbeiten, auseinander. Um die Bewertung durchzuführen wurde ein Bogen erstellt, der die Präparations- und Restorationsphase beurteilte. Der Bogen basierte sowohl auf Lehrbuchinhalten als auch auf Ideen der Assistenten. Die Einzelkomponenten jeder Phase wurden mit einem Drei-Punkte-System (Hinkelman und Long 1973) benotet. Demnach bekam die Arbeit, an der nichts zu beanstanden war, drei Punkte, eine zu

akzeptierende Arbeit zwei Punkte und eine nicht zu akzeptierende Arbeit keine Punkte. Sollte es für einen Teilaspekt keine Punkte geben, war es nicht mehr möglich die Prüfung zu bestehen. Aus allen Einzelbewertungen wurde ein Durchschnittswert gebildet. Diese Art der Bewertung wurde für das komplette Semester verwendet. Im Anschluss wurde die Verlässlichkeit untersucht. Hierzu wurden 60 Inlaypräparationen ausgewählt. Diese Arbeiten wurden von vier Assistenten ein weiteres Mal beurteilt. Von diesen vier, praktizierten zwei erst ein Jahre, die übrigen beiden sieben Jahre. Die 60 zu beurteilenden Arbeiten wurden in zwei 30er Gruppen unterteilt, die im Abstand von vier Wochen und einer dem Training dienenden Diskussion bewertet wurden. Im Anschluss an die Bewertungen wurde die prozentuale Übereinstimmung der Assistenzärzte aus der Drei-Punkte-Bewertung (Hinkelman und Long 1973) miteinander verglichen. Das Ergebnis zeigte, dass die Übereinstimmung vor dem Halbzeitgespräch bei 56,3%, danach bei 58% lag.

Eine weitere Untersuchung der interindividuellen Reliabilität wurde von Gaines und dessen Kollegen (Gaines *et al.* 1974) durchgeführt. Das Ziel dieser Studie bestand darin die Kontinuität der Beurteilung von Bewertern, die unterschiedliche Systeme verwendeten, zu untersuchen. Zu diesem Zweck wurden in zwei Bewertungsdurchgängen zwei Bewertungssysteme angewandt. Folgende Systeme wurden zur Beurteilung der Arbeiten genutzt: Zum einen die Methode, die bereits Verwendung fand, indem für Teilbereiche Punkte von 1-5 vergeben wurden. Zum anderen eine modifizierte Methode (Gaines *et al.* 1974). Diese bestand darin, dass jedem Punktwert eine bildliche Darstellung angefügt wurde. An der Studie nahmen sieben Mitarbeiter der prothetischen Abteilung teil, die acht beliebig auserwählte Arbeiten in den zuvor erwähnten zwei Durchgängen beurteilten. Die Autoren beschreiben in Ihrer Veröffentlichung, dass es signifikante Unterschiede im Vergleich der Teilnehmer bei der Benotung der Prüfungsarbeiten gab. Der größte Unterschied bezifferte 22,6 Punkte, der Durchschnittswert lag zwischen 31,63 und 54,25 Punkten. Der Korrelationskoeffizient wurde auch ermittelt. Für die erste Bewertung betrug dieser 0,26, für den zweiten Durchgang mit der modifizierten Bewertungsmethode bei 0,56. Gaines und seine Mitarbeiter gingen davon aus, dass die hinzugefügte Beschreibung eine größere Kontinuität mit geringeren Varianzen mit sich bringt.

Im Jahr 1978 führten Dhuru mit seinen Mitarbeitern (Dhuru *et al.* 1978) eine Studie über ein Bewertungssystem, welches auf Kriterien basierte, durch. An dieser Untersuchung nahmen zwei Zahnarztgruppen mit unterschiedlicher klinischer

Erfahrung teil. Die erfahrene Gruppe bestand aus 8 Teilnehmern, die zwischen einem und neun Jahren Berufserfahrung hatten. Die zweite Gruppe bestand aus vier Teilnehmern, die nur zwei Monate Erfahrungen in der Lehre sammeln konnten. Es wurden zwei Bewertungsdurchgänge durchgeführt. Die erste Beurteilung führten alle Zahnärzte nach ihren eigenen klinischen Maßstäben anhand eines 10-Punkte-Bogens durch (Dhuru *et al.* 1978). Für den zweiten Bewertungsdurchgang wurden 10 Punkte festgelegt, die für eine positive Bewertung erfüllt werden mussten. Wurden Aspekte nicht erfüllt, führte dies zum Punktabzug. Die Bestimmung der Reliabilität wurde mit Hilfe einer ICC berechnet. Es konnte bei der Auswertung festgestellt werden, dass der Korrelationskoeffizient in beiden Gruppen anstieg. Der Wert erhöhte sich für die erfahrenen Zahnärzte um 0,131, für die unerfahrenen um durchschnittlich 0,072. Der deutlichste Zuwachs konnte für beide Bewertergruppen bei der Beurteilung der Präparation für ein Inlay verzeichnet werden.

Goepferd und Kerber (Goepferd und Kerber 1980) verglichen 1980 in Ihrer Veröffentlichung zwei unterschiedliche Bewertungssysteme um zu beweisen, dass die Reliabilität durch Verwendung einer Checkliste erhöht werden kann. In diesem Fall wurde zum einen die Bewertung per Augenschein der kompletten Arbeit und zum Anderen der Teilschritte durch Kriterien einer Checkliste durchgeführt. Goepferd und Kerber resümierten, dass sowohl die inter- als auch die intraindividuelle Reliabilität, jedoch nicht statistisch signifikant, gesteigert werden konnte.

Das Problem der studentischen Beurteilung veranlasste Feil (Feil 1982) 1982 dazu sich über die Erhöhung der Objektivität Gedanken zu machen. Für ihn und seine Kollegen ging es zu Beginn darum, die Ursachen für Fehler in der Beurteilung der studentischen Arbeiten herauszufinden. Sie kamen zu dem Ergebnis, dass die Ursachen dafür sowohl die Bewerter selbst, die Bewertungskriterien, die nicht klar genug definiert sind, als auch eine unbeständige Bewertungsbreite (Feil 1982) sind. Eine Erhöhung der Reliabilität sollte erzielt werden, indem das Bewertungssystem per Augenschein verbessert wurde. Ansatzpunkte waren eine Erhöhung der Spezifizierung durch Reduzierung der subjektiven Beurteilung und die Einführung eines Feedbacks. Zu diesem Zweck wurde ein Bewertungsbogen eingeführt, der Kriterien beinhaltet, die erfüllt werden müssen. Um die Rückmeldung zu erhöhen, beurteilten die Studierenden ihre Arbeiten bevor dies die Assistenten taten. Man erhoffte sich dadurch, dass die Transparenz für den Studierenden erhöht wurde. Folgende Arbeiten wurden bewertet: VMK-Kronenpräparation, mod-

Kavitätenpräparation und mod-Restauration. Der ICC für den Bewertungsbogen lag für die VMK-Kronenpräparation bei 0,68, für die mod-Kavitätenpräparation bei 0,80 und für die mod-Restauration bei 0,74.

Vann und seine Kollegen (Vann *et al.* 1983) erforschten die Wirkung, die die Verwendung von Kriterien auf die Reproduzierbarkeit und Objektivität der Beurteilung hatten. Im Rahmen dieser Studie wurden 3 verschiedene Bewertungssysteme angewandt. Zum einen wurde ein „glance-and-grade-System“ verwendet. Bei dieser Art der Bewertung wurde die Gesamtleistung ohne Kriterien begutachtet und mit Noten von A bis E versehen. Zum anderen gab es die Methode durch eine Checkliste. Diese Liste beinhaltete 10 Punkte für die Beurteilung der Kavität. Jeder dieser Punkte bekam einen Wert von 1, bedeutend schlecht, bis 5, bedeutend sehr gut. Der letzte Bewertungsansatz beinhaltete eine analytische Methode. Jeder der zuvor erwähnten Punkte wurde mit einer Kriterienliste ausgestattet. Jedes dieser Kriterien wurde erneut mit 1-5 Punkten bewertet. Diese Methode ist identisch mit der analytischen Methode von Goepferd und Kerber (Goepferd und Kerber 1980). Für jede Bewertungsmethode wurden Informationszettel mit den Arbeitsanweisungen ausgegeben. Zu bewerten waren 30 zufällig ausgewählte Klasse-II-Präparationen von 3 Assistenzärzten und 3 Studierenden. Nach einem Tag wurden die Präparation in geänderter Reihenfolge erneut nach dem gleichen Bewertungssystem begutachtet. Vann und seine Mitarbeiter formulierten, dass keine Bewertungsmethode der anderen überlegen sei in Bezug auf die inter- und intrapersonelle Reliabilität. Alle 3 untersuchten Methoden können problemlos angewandt und auch für andere klinische Felder verwendet werden.

Im Jahr 1987 arbeiteten Bedi und seine Mitarbeiter (Bedi *et al.* 1987) an einer Studie zur Beurteilung des Einflusses von Kriterien, die mit Bildern versehen wurden auf die Reliabilität der Benotungen. Zu beurteilen waren fünf Modelle mit fünf Kavitätenpräparationen bei Milchzähnen. An der Studie nahmen drei Gruppen teil. Die erste Gruppe bestand aus fünf Zahnärzten, die in der Lehre tätig waren. Die zweite Gruppe bestand aus 14 Examenskandidaten. Die dritte Gruppe setzte sich aus 16 Erstsemesterstudierenden zusammen. Folgendermaßen wurde bewertet: Zu Beginn wurden die Prüfungsarbeiten dreimal mit dem herkömmlichen Systems per Augenschein mit Schulnoten bewertet. Die Durchgänge erfolgten in einem Abstand von einer Woche. Im Anschluss an diese drei Durchgänge wurden den Schulnoten bildliche Bewertungskriterien angefügt (Bedi *et al.* 1987). Anhand dieser Kriterien

wurden die Prüfungsarbeiten erneut dreimal begutachtet. Der Zeitabstand zwischen den jeweiligen Bewertungsdurchgängen betrug erneut eine Woche. Die statistische Auswertung zeigte, dass alle Teilnehmer der Studie durch die Nutzung der Kriterien eine höhere interindividuelle Reliabilität erzielten. Der durchschnittliche Wert des Kappakoeffizienten betrug 0,71. Die deutlichste Steigerung der Reliabilität durch die Verwendung der Bewertungskriterien konnte in der Gruppe der Examenskandidaten verzeichnet werden. Die Durchschnittsnote nach Verwendung der Kriterien mit Bildern lag bei den Examenskandidaten bei 3,4, bei den Zahnärzten bei 3,76. Bedi *et al.* (Bedi *et al.* 1987) gingen davon aus, dass die Bewertungskriterien, die mit Bildern versehen wurden, eine positive Wirkung auf das Lernverhalten der Studenten hätte, da durch diese Methode die Selbsteinschätzung vereinfacht würde.

Jenkins und Mitarbeiter (Jenkins *et al.* 1998) führten 1998 eine Untersuchung mit Teilnehmern mit unterschiedlicher Lehrerfahrung durch. In zwei Bewertungsdurchgängen wurden 75 Klasse-II-Präparationen für Amalgamfüllungen beurteilt. Die Teilnehmer der Studie hatten schriftliche Informationen ausgehändigt bekommen über die Idealform der Präparation. Eine Checkliste für die Bewertung wurde nicht ausgehändigt. Alles in allem konnten 13 Noten abgegeben werden für jede Präparation. Die Noten wurden im Anschluss in Punkte umgewandelt. Sowohl die Erfahrensten als auch die unerfahrensten Probanden hatten eine Differenz von fünf Punkten und weniger. Die noch verbleibenden Teilnehmer der Studie zeigten Abweichungen, die größer als sieben Punkten waren hinsichtlich der intraindividuellen Übereinstimmung. Um die Übereinstimmung der Beurteilungen der Teilnehmer darzustellen hat man alle Resultate mit denen von Bewerter eins gegenübergestellt (Kellersmann 2008). Bei diesem Untersuchungsansatz konnte man erkennen, dass auch in diesem Fall Differenzen größer als fünf Punkte zu erkennen waren. Außerdem wurde deutlich, dass die erfahrensten Zahnärzte die besten Noten verteilten.

Aly Sharaf (Sharaf *et al.* 2007) setzte sich 2006 mit der intra- und interpersonellen Variabilität in der Bewertung zahnmedizinischer Arbeiten in der Kinderzahnheilkunde auseinander. In dieser Studie wurden zwei verschiedene Bewertungssysteme miteinander verglichen. Zum einen per glance and grade, zum anderen mittels einer Checkliste und Kriterien (Sharaf *et al.* 2007). Teilgenommen haben an dieser Untersuchung drei Mitarbeiter der Abteilung mit mehr als zehn Jahren Erfahrung. Zu beurteilen waren sieben verschiedene Präparationen an Kunststoffmilchzähnen

von 30 Studierenden. Es standen den Probanden keine Hilfsmittel wie Lupenbrille oder Sonden zur Verfügung. Jede der Arbeiten wurde dreimal von jedem der drei Probanden allein beurteilt. Im ersten Durchgang wurden die Arbeiten per glance and grade (Sharaf *et al.* 2007) mit einem Punktwert von eins bis zehn bewertet. Der zweite Bewertungsdurchgang erfolgte nach drei Tagen. Zur Bestimmung der intrapersonellen Variabilität wurde erneut das Bewertungssystem per Augenschein verwendet. Im Anschluss an diese beiden Bewertungsdurchgänge wurden in Zusammenarbeit mit den Studierenden für jede Art der Präparation Kriterien entwickelt. Die dritte Bewertung der Präparationen erfolgte durch Zuhilfenahme der Liste der Kriterien und des Zehn-Punkte-Skala (Sharaf *et al.* 2007). Die Ergebnisse dieser Studie zeigten, dass die intrapersonelle Variabilität im Vergleich der drei Bewerter nicht signifikant war. Bei nahezu allen Präparationsarten war die interpersonelle Variabilität der Untersucher jedoch signifikant. Die Autoren formulierten, dass es weder für die Bewertung per Augenschein (glance and grade) noch für die Bewertung durch Kriterien eine Verbesserung der Beständigkeit gab. Somit bestünde noch immer das Problem einer unzureichenden interpersonellen Reliabilität und Variabilität.

1.3 Vorarbeiten

1.3.1 Literaturübersicht zu Kriterien für praktische Arbeiten im Phantomkurs der Zahnerhaltung

Mit Hilfe einer Literaturrecherche wurden zu Beginn der Studie Kriterien zusammengestellt.

Der **erste Themenkomplex** beschäftigte sich mit den Kriterien, die eingehalten werden sollten bei der **Präparation einer Klasse II Kavität im Seitenzahnbereich mit einer Karieslokalisation mesial, distal und okklusal**.

An erster und wichtigster Stelle war als Kriterium zu nennen, dass der zu präparierende und, oder der Nachbarzahn nicht verletzt werden dürften. Durch unbeabsichtigtes Anpräparieren von Nachbarzähnen, besonders in approximalen Bereichen steige die Gefahr für Karies und parodontale Erkrankungen (Qvist *et al.* 1992). Es müsse besonders auf den Schutz des zu präparierenden und der

Nachbarzähne geachtet werden, um weitere unbeabsichtigte Therapieschritte vermeiden zu können (Lussi und Gygax 1998).

Es musste bei der Präparation des approximalen Kastens auf die zervikale Höhe geachtet werden. Sollte die Präparationsgrenze zu weit zervikal liegen, habe dies weitreichende Folgen. Subgingivale Restaurationsränder führten demnach dazu, dass bei direkten Restaurationen die Möglichkeit der absoluten Trockenlegung stark eingeschränkt sei. Außerdem träten vermehrt gingivale Entzündungen, Attachmentverlust und Knochenresorptionen auf, da die Biologische Breite verletzt werde (Lanning *et al.* 2003).

In Folge der Literaturrecherche hatte sich herausgestellt, dass die Substanzschonung des zu behandelnden Zahnes besonders beachtet werden müsse. Ist die Karies wie bei einer Klasse II Kavität okklusal und im approximalen Bereich lokalisiert, dürfe sich die Größe der Kavität lediglich nach der Ausdehnung der Karies orientieren. Es solle demnach bei der Präparation die Tiefe der okklusalen Kavität beachtet werden, da eine Verletzung der Pulpa vermieden werden könne. Die Gestaltung der Präparation solle kleiner und defektbezogener sein. Die Präparationsform müsse nicht mehr zur Retention dienen, da für adhäsive Füllungen keine Unterschnitte notwendig seien. Dementsprechend werde es ausreichen lediglich von Karies betroffenes Dentin zu entfernen (Summitt 2006).

Die Gestaltung der okklusalen Kavität richte sich ausschließlich nach der Ausdehnung der Karies und soll nicht nur in Richtung Pulpa, sondern auch in oro-vestibulärer Dimension defektbezogen gestaltet werden. In der Folge der Weiterentwicklung der Restaurationsmaterialien hatten sich auch die Anforderungen an die Präparation geändert. Dies bedeutete, dass keine Unterschnitte, wie für Amalgam notwendig, präpariert werden müssen. Die Ausdehnung der Kavität könne sich an der Defektgröße orientieren. Als strikte Regel sollte eingehalten werden, dass die Kavität so klein wie möglich gehalten werden sollte (Frankenberger *et al.* 2003).

Nicht nur die okklusale Tiefe der Präparation musste beachtet werden, auch die Ausdehnung der approximalen Kästen durfte in mesio-distaler Relation nicht zu groß gestaltet werden, da sonst eine Gefährdung der Pulpa bestünde. Die Präparation der Karies sollte durch minimalinvasive Methoden (Mount und Ngo 2000) durchgeführt

werden und lediglich durch die Kariesausdehnung begrenzt werden, um den Zahn zu schonen.

Die Präparation musste so gestaltet werden, dass die anschließende Füllung suffizient gelegt werden konnte. Dazu zählte, dass der Rand der Präparation definiert war und keine losen Schmelzprismen verblieben. Lose Schmelzprismen sollten entfernt werden, da sie brüchig seien und im Folgenden zu postoperativen Hypersensibilitäten und Sekundärkaries führen können, welches den Austausch der Füllung notwendig machen könne (Summitt 2006). Zu den Gründen für das Scheitern von Seitenzahnrestaurationen zählten insuffiziente Füllungsänder, postoperative Hypersensibilitäten und Sekundärkaries (Hickel und Manhart 2001). Anhand dieser Ausführungen erkannte man, dass bei der Randgestaltung der Präparation besonders darauf geachtet werden sollte, dass eine definierte Präparationsgrenze vorliege, die es ermöglichte eine suffiziente Füllung zu applizieren.

Ein weiteres schwerwiegendes Kriterium war die Traumatisierung des Zahnes in Folge nicht patientenäquivalenten Arbeitens. Die Präparation eines Zahnes ohne Wasserkühlung könne schwerwiegende Folgen für den Zahn haben. Die Pulpa des Zahnes akzeptiere einen Temperaturanstieg von mehr als 5,5°C nicht, daher sei es unumgänglich, dass der Schleifkörper durch Wasser gekühlt werde. Die Folgen einer Präparation ohne Wasserkühlung könnten eine temporäre oder dauerhafte Schädigung der Pulpa sein (Cavalcanti *et al.* 2002).

Eine der Leitlinien der Behandlung, unabhängig von der Gestaltung der Präparation, war die vollständige Entfernung der Karies. Das wichtigste Ziel der Behandlung ist die komplette Entfernung der dentalen Karies um die größte Sicherheit für das gesunde Restzahngewebe und damit den Vitalerhalt der Zahnpulpa zu erreichen (Javaheri *et al.* 2010).

War die Karies im Bereich des approximalen Kontaktpunktes lokalisiert und wurde in Folge dessen ein approximaler Kasten während der Präparation angelegt, sollten die Ränder dessen sondierbar und der approximale Kontakt folgerichtig aufgelöst sein, um später kontrollieren zu können, ob der Füllungsrand dicht war und somit keine Randspalten vorlägen. Randspalten könnten zu Sekundärkaries führen. Sekundärkaries habe einen negativen Effekt auf den klinischen Erfolg (Kramer *et al.* 2009). Anhand dieser Aussage wurde klar, dass bei der Präparation schon darauf

geachtet werden sollte, dass alle Voraussetzungen geschaffen wurden für eine suffiziente Versorgung mit guter klinischer Prognose.

Ein weiter Aspekt der Substanzschonung befasste sich mit der Gestaltung und Ausdehnung der Randanschrägung. Bei Klasse II Präparationen im Seitenzahnggebiet sollte eine Anschrägung deutlich, definiert und nicht zu breit an der vestibulären und oralen Begrenzung des approximalen Kastens angelegt werden. Falle diese Anschrägung zu breit aus, sei dies sowohl zahnschädigend als auch hinderlich für eine spätere suffiziente Füllungsrandgestaltung (Wilson *et al.* 1990).

Der **zweite Themenkomplex** beinhaltete die Kriterien, die bei einer anatomisch korrekten und ästhetisch hochwertigen **direkten Kompositfüllung mit Approximalkontakten im Seitenzahnbereich** zu erfüllen waren.

Nicht nur bei der Präparation sondern auch bei der späteren Versorgung durch eine Füllung hatte der Schutz des Nachbarzahnes eine hohe Priorität. Durch die iatrogene Verletzung des Nachbarzahnes während der Präparation durch den Zahnarzt werde die Gefahr für eine Karies steigen und damit verbunden die Notwendigkeit für eine restaurative Therapie des Nachbarzahnes (Lussi 1995). Aus diesem Grund wurde empfohlen den Nachbarzahn so gut es geht zu schützen, um iatrogene Verletzungen des Nachbarzahnes zu vermeiden.

Ein weiteres wichtiges Kriterium bei der Anfertigung einer Füllung sowohl im Front- als auch im Seitenzahnggebiet war, dass sie nicht locker sein dürfte. Eine lockere Füllung hatte eine Neuanfertigung zur Folge. Resultat einer lockeren Füllung, die belassen und nicht therapiert wurde, könne Sekundärkaries sein. Sekundärkaries sei der Hauptgrund für den Austausch und die Neuanfertigung von Amalgam-, Komposit- und Glasionomerzementfüllungen. Aus diesem Grund müsse darauf geachtet werden, dass bei der Anfertigung einer Füllung alle Ansprüche an eine suffiziente Füllung, wie Randdichtigkeit, keine Lockerung, erfüllt werden, damit keine kariesbildenden Prozesse zwischen Zahn und Füllung stattfinden könnten (Burke *et al.* 1999).

Ein weiteres wichtiges Kriterium bei der Anfertigung einer Füllung war die korrekte Herstellung eines Approximalkontaktes. Die Anwesenheit eines Approximalkontaktes sei essentiell um die Position des Zahnes zu erhalten und ihm Stabilität zu verleihen.

Außerdem diene er als Barriere für Essen und erleichtert die Reinigung der Zähne. Die Form des Approximalkontaktes sei abhängig von der Kontur des Zahnes und eine umschriebene Fläche, weniger ein kleiner einzelner Punkt. (Chuang *et al.* 2011). Ein fehlender Approximalkontakt sei einer der Hauptgründe für einen frühzeitigen Austausch oder eine Reparatur der Restauration (Opdam *et al.* 2004). Der Approximalkontakt sollte jedoch nicht einfach nur vorhanden sein, sondern auch an der richtigen Position und nicht zu schwach gestaltet werden. Wenn der Approximalkontakt zu schwach sei, könnten durch die Einlagerung von Speiseresten daraus parodontale Probleme und Karies entstehen (El-Badrawy *et al.* 2003).

Die Lage des Approximalkontaktes sei von entscheidender Bedeutung. Der Kontaktpunkt dürfe weder zu weit inzisal oder zu weit koronal liegen. Der Approximalkontakt solle im koronalen Drittel zu finden sein, da ansonsten die interdental Gingiva den interdentalen Raum nicht ausfüllen könne (Chuang *et al.* 2011). Wenn der Kontaktpunkt 5 mm Abstand zum Knochen habe, sei die Papille zu 100% erhalten. Wenn die Distanz 6 mm Abstand betragen würde, wäre die Papille nur noch zu 56% erhalten, bei 7 mm oder mehr Abstand sogar nur noch zu 27% vorhanden (Tarnow *et al.* 1992). Dies machte deutlich wie wichtig die korrekte Lage des approximalen Kontaktes in okklusaler-zervikaler Richtung war.

Es musste jedoch bei der Herstellung des Approximalkontaktes darauf geachtet werden, dass dieser nicht zu stark ist oder verblockt wird. Die Technik des „prewedging“ verhinderte das Problem von Verblockungen nach Anfertigung einer direkten Restauration (Loomans *et al.* 2006). Zu starke Approximalkontakte oder Verblockungen führten zu Unannehmlichkeiten des Patienten, unerwünschten Zahnbewegungen, einer Barriere am Kontaktpunkt für Zahnseide und resultierten schließlich in einer Traumatisierung des Parodontiums (Chuang *et al.* 2011).

Ein Randspalt, auch von kleiner Größe, könne in der Folge dazu führen, dass weitere Therapieschritte notwendig werden. Durch einen Randspalt zwischen Restauration und Zahn steige die Sensibilität des Zahnes und könne bis zu irreversiblen Schäden an der Pulpa führen (Rodrigues Junior *et al.* 2010). Die Hauptfolge eines Randspaltes sei eine Microleakage, das in der Folge zu Sekundärkaries führen könne (Mjor und Toffenetti 2000).

Eine Vorstufe des Randspaltes war der Unterschuss am Füllungsrand. Ein Unterschuss könne schon zum Misserfolg der Restauration, mit den bekannten

weiteren Folgen wie erhöhte Sensitivität und Sekundärkaries führen (Opdam *et al.* 2004).

Es solle jedoch nicht nur ein Unterschuss des Füllungsmaterials am Restaurationsrand vermieden werden, sondern auch ein Überschuss. Bei der Anfertigung einer direkten Restauration müssen die Anatomie und die Form des jeweiligen Zahnes und des umliegenden Gewebes beachtet und respektiert werden. Werden dies nicht getan, könne es in der Folge zu Entzündungen der Gingiva und des Zahnhalteapparates kommen, ausgelöst durch die vermehrte Anlagerung von Plaque in vom Behandler geschaffenen Bereichen und Nischen. Um dies zu vermeiden, solle auch nach dem Legen einer Füllung und der Ausarbeitung, der Rand der Füllung kontrolliert werden (Parkinson 1976).

Bei der Versorgung eines Zahnes mit einer direkten Restauration solle eine suffiziente Kauflächenmodellation das Ziel sein. Dies bezog sich sowohl auf den Frontzahnbereich als auch auf den Seitenzahnbereich. Es reiche nicht aus die fehlenden Zahnbereiche aufzufüllen. Es solle die ideale Form neu aufgebaut werden, um eine optimale Funktion zu erhalten (Raghu und Srinivasan 2011). Im Seitenzahnbereich zählten zu den anatomischen Parametern die Anzahl, Lage und Form der Höcker, die Zentralfissur, die Gestaltung der Randleiste mit Grube in Richtung Zahnmitte und Einziehung koronal des Approximalkontaktes (Benninghoff 2008).

Mit den aktuellen Füllungstechniken könnten die ästhetischen Ansprüche zufrieden gestellt werden. So sei es möglich die farbliche Integrität, eine ideale anatomische Form und Funktion, optimale proximale Kontakte und eine optimale Randedichtigkeit zu erreichen. Die materialtechnischen Voraussetzungen müssen lediglich vom Behandler umgesetzt werden (Terry 2005). Es sei auch möglich, ein sehr ästhetisch ansprechendes Ergebnis zu erzielen, selbst wenn man den Farbton nicht ganz treffe, solange die Form ideal sei. Allerdings sei die korrekte Farbauswahl sehr wichtig. In Kombination mit einer unzureichenden Formgebung sähe das Ergebnis nicht natürlich aus und werde den Patienten nicht zufrieden stellen (Margeas 2010). Gerade im Frontzahnbereich sei die große Herausforderung, die Frakturlinie eines Klasse IV Defektes für das menschliche Auge weitestgehend unsichtbar zu machen. Dafür müsse der behandelnde Zahnarzt sowohl das richtige Material, in der korrekten

Farbe, als auch die korrekte Anwendung des Restaurationsmaterials beherrschen (Fahl 1997).

Bei der Modelation der Restauration müsse darauf geachtet werden, dass das Material homogen sei. Eine suffiziente Inkrementadaptation reduziere die Gefahr für Inhomogenitäten und sei zwingend notwendig, um die mit der Polymerisation zusammenhängende Schrumpfung des Komposites so gering wie möglich zu halten. Mit der richtigen Technik solle erreicht werden, dass die Schrumpfung in Richtung der Kavitätenwände und nicht entgegen gesetzt gerichtet werden solle. Erfolge die Schrumpfung weg von der Kavitätenwand bestünde die Gefahr der Randspaltbildung (El-Badrawy *et al.* 2003).

Aus den beschriebenen Inhomogenitäten und Spannungen können in der Nutzungsphase durch den Patienten Frakturen der Restauration entstehen. Der wichtigste Faktor für die Langlebigkeit einer Zahnrestauration sei der Behandler selbst. Schaffe der Behandler eine suffiziente Kauflächengestaltung mit richtiger Inkrementtechnik und verhindere somit Inhomogenitäten, sinke sowohl das Risiko für Frakturen durch normale Belastung sowie Bruxismus als auch für Sekundärkaries (Demarco *et al.* 2012).

Unabhängig von Farbe und Form der Gestaltung der Front- oder Seitenzahnfüllung werde die Restauration immer mangelhaft erscheinen, wenn die Oberflächenbearbeitung insuffizient erfolge. Die Restauration sehe nicht nur besser aus, sie habe auch eine bessere Prognose für eine längere Zeit. Denn nicht nur die Ästhetik auch die Funktion stünde im Vordergrund (Goldstein 1989). Eine glatte Oberfläche und die Entfernung von Irregularitäten auf der Oberfläche reduzierten die Zahnsteinbildung, Plaqueanhaftung, Gingivale Irritationen, Sekundärkaries, Abrasion des Füllungsmaterials und steigerten das Wohlbefinden des Patienten. Daher müsse größte Sorgfalt betrieben werden nicht nur bei der Gestaltung der Füllung sondern auch bei der Ausarbeitung und Politur (Lesage 2011). Der Zahnarzt solle nachdem er die Restauration ausgearbeitet habe diese zudem ausreichend polieren. Für die Politur von Kompositen sollen adäquate Polierinstrumente verwendet werden, die an die Größe der Füllkörper angepasst sei. Denn nur so könne eine zufriedenstellende Politur der Restauration erreicht werden (Berger *et al.* 2011).

Abschließend sollte erneut erwähnt werden, dass eine patientenadäquate Behandlung mit Wasserkühlung auch bei der Ausarbeitung einer direkten

Restauration verlangt wurde. Sowohl kurze Pausen während des Beschleifens als auch eine Wasserkühlung des Zahnes auf 30-32°C erzielten den erwünschten positiven Effekt (Ottl und Lauer 1998).

Das **dritte Themengebiet** beinhaltet die Bewertungskriterien für die **Präparation einer klassischen Stufenteilkrone für eine indirekte Goldrestauration.**

Die Gefahr der Verletzung der Zahnpulpa bestehe bei der Präparation für eine Goldteilkrone durch zu großen Substanzabtrag sowohl direkt durch Eröffnung der Pulpa als auch durch Traumatisierung dieser oder inadäquate Präparationstechniken. Die Kavitätenpräparation solle mit ausreichender Wasserkühlung durchgeführt werden, um einen Temperaturanstieg zu vermeiden und somit das Pulpagewebe zu schützen (Ozturk *et al.* 2004).

Erfolge ein zu starker Substanzabtrag steige die Gefahr einer Pulpaverletzung an. Unter einer Restdentindicke von 0,25 mm gehe die Anzahl der Odontoblasten um 23% zurück und finde eine minimale Tertiärdentinbildung statt (About *et al.* 2001). Eine Restdentindicke von mindestens 0,5 mm sei notwendig, um eine Pulpaschädigung zu vermeiden (Murray *et al.* 2003).

Eine optimale Präparationstechnik solle den Anspruch erheben, den Nachbarzahn vor iatrogener Verletzung zu schützen (Hugo 1999).

Die ideale okklusale Tiefe liege bei 0,8 - 1 mm, um die Mindestschichtstärke für Gold zu gewährleisten. Dazu müssen die zentrischen Höcker um 1,2- 1,5 mm, die Scherhöcker um etwa 1 mm gekürzt werden (Klaiber und Hugo 1998). Die Tiefe des approximalen Kastens solle 2,5 mm betragen (Summitt 2006). Die Präparationsgrenze müsse sondierbar und definiert sein. Zur Erzielung einer ausreichenden Metallstärke im Randbereich müsse sich die Schliifffläche deutlich vom Verlauf der Zahnoberfläche abheben (Klaiber und Hugo 1998; Summitt 2006).

Ein weiteres zu beachtendes Kriterium bei der Präparation für eine Goldteilkrone war der Öffnungswinkel. Dieser durfte weder zu groß noch zu gering sein. War der Öffnungswinkel zu groß können in der Folge Retentionsprobleme auftreten. War der Öffnungswinkel zu klein, werde die Positionierung der Goldrestauration schwierig. Der Öffnungswinkel für die vestibuläre Wand solle zwischen 3 und 5 Grad, in der Summe 6 bis 10 Grad, betragen (Summitt 2006). Der Hauptgrund für ein Scheitern einer indirekten Goldrestauration sei mit 33,7% Sekundärkaries, gefolgt mit 32,7%

mangelnde Retention, gefolgt von endodontischen Behandlungen (29,6%), insuffiziente Randgestaltung (3,1%) und Extraktionen (1%) (Stoll *et al.* 1999). Es dürften keine Unterschnitte vorhanden sein, da es sonst nicht möglich sei die Restauration zu Positionieren und wieder Abzuheben. Des Weiteren sei es nicht möglich ein suffizientes Provisorium herzustellen (Summitt 2006). Eine Ansträgung der Präparationsgrenze der approximalen Kästen wurde empfohlen (Summitt 2006). Durch den Einsatz des SonicSys Aufsatzes für Goldrestorationen könnten nahezu perfekte Ränder präpariert werden und damit das Ziel eines Federrandes erreicht werden (Hugo 1999). Die Gestaltung des Federrandes am Scherhöcker wurde mit 45 Grad zur Zahnachse definiert (Summitt 2006). Die Scherhöcker sollten im Sinne eines Außenschliffs zum Kaukantenschutz angeschrägt werden (Federlin *et al.* 2006). Stützhöcker jedoch sollten durch eine Stufe mit Ansträgung überkuppelt werden. Die Schichtstärke müsse dabei immer bedacht werden. Bei der Überkuppelung mit angeschrägter Stufe läge bei Oberkiefermolaren die Stufe etwa auf Höhe der größten Zirkumferenz. Sie wäre nach Ansträgung noch ca. 1 mm breit (Klaiber und Hugo 1998).

Am Übergang vom approximalen Federrand zum Höckerschutz bzw. Kaukantenschutz sollten keine scharfen Ecken vorliegen. Die approximalen Ansträgungen sollten fließend in die angrenzenden Ansträgungen übergehen. Es sollten jedoch nicht nur weiche Übergänge an den verschiedenen Ansträgungen vorliegen sondern auch an den Übergängen von Präparationsboden zu den benachbarten Wänden (Summitt 2006).

Bei der Gestaltung des approximalen Kastens war darauf zu achten, dass die Präparationsgrenze nicht im Kontaktpunktbereich liege. Der Approximalkontakt sollte deutlich aufgelöst sein. Der Rand sollte in gesunder Zahnstruktur liegen, außerdem wurde empfohlen, dass der Nachbarzahn deutlich separiert wurde (Osborne und Summitt 1998).

Der folgende **vierte Abschnitt** befasste sich mit den zu beachtenden Kriterien bei der **Präparation einer Keramikteilkrone**.

Auch bei dieser Präparation musste selbstverständlich beachtet werden, dass der Nachbarzahn geschützt wurde vor iatrogenen Verletzungen. Zu den Folgen einer Verletzung des Nachbarzahnes gehörte die erhöhte Gefahr der Kariesbildung und

Plaqueeinlagerung. Diese iatrogene Verletzung des Nachbarzahnes könne für Patienten mit einem hohen Kariesrisiko eine negative Prognose für diesen Zahn bedeuten. Außerdem könnten diese Läsionen im Rahmen späterer Untersuchungen fehlinterpretiert werden, dass diese einen kariogenen Ursprung haben (Lussi und Gygax 1998). Weiter verschlechtert werde die Prognose für einen Zahn, wenn nicht patientenäquivalent gearbeitet werde. Zu den geforderten Maßnahmen gehörte die schon in anderen Themenkomplexen beschriebene Wasserkühlung während der Präparation (Ottl und Lauer 1998; Cavalcanti *et al.* 2002; Cavalcanti *et al.* 2005).

Es zeigte sich, wenn die Präparationsgrenzen unterhalb der Schmelz-Zementgrenze liege, dass der Halt der Restauration deutlich sinke. Maßnahmen zur absoluten Trockenlegung, die im Fall einer Eingliederung einer keramischen Restauration notwendig seien, werden zudem erschwert. (Federlin *et al.* 2004).

Eine Reduktion von 1,5 mm an Zahnschicht im Kaulastbereich sei ausreichend für Lithiumdisilikatpresskeramiken. Diese Materialstärke sei erforderlich, um die notwendige Stabilität des Materials zu erreichen und mögliche Brüche des Materials zu vermeiden (Banks 1990).

Die Mindestschichtstärke solle jedoch nicht nur okklusal oder approximal eingehalten werden, sondern auch im Bereich zwischen den Höckern, also im Isthmus. Die Isthmusbreite solle an der engsten Stelle 2 mm nicht unterschreiten (André Hutsky 2011), um keine Sollbruchstelle zu kreieren.

Die Wände des Isthmus und der approximalen Kästen sollten nicht zu parallel oder gar unter sich gehend gestaltet werden. Der Öffnungswinkel dürfe weder negativ noch zu gering sein. Der Öffnungswinkel solle 6 bis 10 Grad betragen. Generell gälte, dass eher zu divergierend als zu parallel präpariert werden solle. Seien die Öffnungswinkel zu klein werde die Positionierung in der Endposition der keramischen Arbeit nur sehr schwierig oder nicht möglich sein (Frankenberger *et al.* 2008). Sollte der Öffnungswinkel negativ und damit die Präparation unter sich gehend sein, sei die Eingliederung eines Provisoriums und später der Restauration nicht möglich. Unterschnitte dürften nicht vorhanden sein. Sollte es dennoch dazu kommen, dass Unterschnitte vorhanden seien, müssten diese ausgeblockt werden, damit eine spätere suffiziente Versorgung möglich werde (Banks 1990).

Die Gestaltung der Präparation richtete sich in erster Linie nach dem Substanzverlust durch frühere Restaurationen, Karies oder ein Trauma. Es sollte nur so viel präpariert

werden, dass beim Eingliedern eine eindeutige Lagefixierung der Keramikrestauration in der definierten Endposition sichergestellt wäre. Somit werden die Voraussetzungen geschaffen, dass Klarheit bei der Eingliederung über die Endposition herrsche und die Randgestaltung der Restauration beurteilt werden könne. Außerdem werde vermieden, dass ein zu großer Substanzabtrag durch die Präparation erfolge (Frankenberger et al. 2008).

Ein zu großer Substanzabtrag habe einen negativen Einfluss auf die langfristige Prognose. An verschiedenen Bereichen des präparierten Zahnes bestünde durch den zu großen Substanzabtrag die Gefahr für eine Pulpaverletzung. Zu diesen Bereichen gehörten die Höckerspitzen, ganz besonders bei jungen Zähnen, der Isthmus und die approximalen Kästen in mesio-distaler Richtung (André Hutsky 2011). Das Ziel müsse es sein, so wenig wie möglich Zahnschicht zu entfernen und gleichzeitig die Ansprüche an die Schichtstärke zu erfüllen (Frankenberger et al. 2008). Mit den aktuellen Instrumenten für die Kavitätengestaltung und den neuen Materialien habe man die Möglichkeit den Zahn bestmöglich zu schützen und so bestmöglich minimalinvasiv zu präparieren (St-Georges *et al.* 2003).

Die Präparation für Keramikrestorationen solle so gestaltet werden, dass keine Gefahr für Insuffizienzen oder Frakturen für Zahn oder Keramik entstehen. Aus diesem Grund solle keine Anschrägung des Randes präpariert werden. Es gebe keinen Vorteil, wenn eine Randanschrägung erfolge, weder im okklusalen, noch im approximalen Bereich. Aus diesem Grund sollten Anschrägungen vermieden werden (Summitt 2006). Die Begrenzung der beschliffenen Oberfläche solle einen Winkel von 90° haben. Würde diese Forderung konsequent umgesetzt, erreiche man zwangsläufig eine definierte Präparationsgrenze, welche zwingend notwendig sei, um im späteren Verlauf der Herstellung der Keramikrestauration Klarheit zu haben wo die Präparation ende und wo die Restauration begänne. Zu den Folgen eines ungenau gestalteten Randes gehörten Sekundärkaries, Frakturen und postoperative Hypersensivitäten (Hickel und Manhart 2001).

Die approximalen Ränder am Kastenboden und den Flanken in oro-vestibulärer Richtung sollten generell soweit aufgezogen werden, dass die Kavitätenränder den Nachbarzahn nicht berührten. Somit sei sichergestellt, dass der Approximalkontakt aufgelöst sei und die Präparationsgrenze nicht im Bereich des Kontaktes läge (Banks 1990). Zugleich sei die Präparationsgrenze sondierbar. Dies sei notwendig für die

Abformung und erleichterte die Kontrolle des Sitzes während der Einprobe und Überschussentfernung des Einsetzwerkstoffes erheblich (Frankenberger *et al.* 2008).

Ein weiteres zu beachtendes Kriterium während der Präparation für eine Keramikrestauration war die Gestaltung der Innenkanten und Übergänge der verschiedenen Bereiche der Präparation. Alle Zähne, die für Keramikinlays oder Teilkronen präpariert wurden, sollten eine geglättete Oberfläche mit abgerundeten Linien und Winkel haben (Banks 1990). Die Präparation solle mit einer extrem geglätteten Oberfläche, abgerundeten Innenwinkeln und scharf begrenzten Außenwinkeln durchgeführt werden (Mizrahi 2011).

Die Präparationsflächen sollten finiert werden, um die Retention zu steigern. Um ein Höchstmaß an Oberflächenglätte zu erreichen, müssten die geeigneten Schleifkörper ausgewählt werden. Laut Ayad steigere durch das Finieren der Präparationsflächen die Benetzbarkeit des Dentins und somit der Haftverbund des Einsetzwerkstoffes (Ayad *et al.* 2009).

Der **fünfte Themenkomplex** befasst sich mit den zu erfüllenden Kriterien bei **Trepanation und der Schaffung der Zugangskavität an einem Molaren**.

An Hand der aktuellen Literatur erkannte man, dass der Einsatz von Kofferdam während einer Wurzelkanalbehandlung eine Standardbehandlungsmethode sei. Die Infektionskontrolle, der Schutz des Patienten sowie die Erfolgsaussichten der Wurzelkanalbehandlung könnten durch die Verwendung von Kofferdam gesteigert werden. Aus diesen Gründen könne die Verwendung eines Kofferdams nur von Vorteil sein und wird daher gefordert (Lin 2011).

Die Gestaltung einer zu kleinen Zugangskavität reduziere die Übersicht und vergrößere die Gefahr, dass eine unvollständige Entfernung des Pulpagewebes vorgenommen werde. Ein in der Folge nicht gerader Feilenzugang erschwere eine ausreichende Reinigung des Wurzelkanalsystems und im nächsten Schritt eine suffiziente Wurzelkanalfüllung. Diese negativen Faktoren verschlechterten die Prognose des wurzelkanalbehandelten Zahnes (Patel und Rhodes 2007). Das Pulpdach solle komplett entfernt werden, damit eine gute Übersicht bestünde und somit die Möglichkeit die Eingänge in die Wurzelkanäle zu finden (European Society of Endodontology 2006).

Die Größe der Zugangskavität richte sich nach den anatomischen Gegebenheiten des jeweiligen Zahnes. Um eine gute Voraussetzung für eine erfolgreiche Behandlung des Wurzelkanals zu schaffen, solle die Größe der Zugangskavität so gewählt werden, dass die gesamte Ausdehnung des Wurzelkanalsystems dargestellt werde. Werde die Zugangskavität größer angelegt, bedeutete dies eine unnötige Schwächung der Restzahnschubstanz (Krasner und Rankow 2004).

Es sei zu beachten, dass die Pulpa immer in der Mitte des Zahnes anzufinden ist und der Boden der Pulpakammer auf Höhe der Schmelz-Zement-Grenze. Wenn man sich an diese anatomischen Begebenheiten halte und während der Trepanation beachte, sei die Gefahr einer Perforation verringert. Zudem stimme dann sowohl die Lage als auch die Form der Zugangskavität (Johnson 2009). Es könne sowohl während der Suche nach den Eingängen der Wurzelkanäle am Boden der Kronenpulpa als auch im späteren Verlauf während der Aufbereitung der Wurzelkanäle zur Perforation kommen. Käme es während der Behandlung zu einer Perforation verschlechtere sich die Prognose des Zahnes. Dabei seien die Lage der Perforation, die Möglichkeit der Deckung und der Zeitpunkt der Deckung entscheidend für den weiteren klinischen Erfolg der Behandlung und die Erhaltungswürdigkeit des Zahnes. Bereits das Anpräparieren des Pulpakammerbodens verschlechterte langfristig die Prognose des Zahnes, da die Stabilität des Zahnes sinke und damit die Gefahr einer Fraktur steige (Leo 2012).

Ausschlaggebend für den Erfolg einer Wurzelkanalbehandlung sei es, dass alle Wurzelkanäle gefunden, anschließend aufbereitet und mit einer Füllung versorgt werden (Patel und Rhodes 2007). Denn genauso wichtig wie die Diagnosestellung und Behandlungsplanung sei die gute Kenntnis des Wurzelkanalsystems, denn dies sei notwendig für den Erfolg der Wurzelkanalbehandlung (Li *et al.* 2011).

In der Fachliteratur wurde beschrieben, dass ein geradliniger Feilenzugang zwingend notwendig sei, um den Wurzelkanal aufbereiten zu können. Es müssten sowohl die Zugangskavität ausreichend groß, als auch die Kanaleingänge klar dargestellt werden (Krasner und Rankow 2004).

Bei der Präparation des Wurzelkanals solle immer der Verlauf und die ursprüngliche Breite des Kanals beachtet werden. Werde jedoch der Eingang in den Wurzelkanal zu gering und nicht konisch aufbereitet, sei ein geradliniger Feilenzugang und eine spätere suffiziente Wurzelfüllung nicht möglich (Young *et al.* 2007). Wenn der

koronale Bereich nicht ausreichend konisch aufbereitet werde, erhöhe sich die Gefahr für den Verlust von Arbeitslänge während des Aufbereitens und später bei der Wurzelfüllung (Davis *et al.* 2002).

Eine erfolgreiche Wurzelkanalbehandlung sei nur möglich, wenn Karies aus der Zugangskavität komplett entfernt und eine ausreichende Spülung betrieben wurde. Zu der Reinigung zählten sowohl die mechanische Reinigung mittels Instrumenten als auch die chemische durch Spüllösungen (Basrani 2011).

Der **sechste und letzte Themenkomplex** listet die Kriterien auf, die zu beachten sind während einer **Wurzelkanalfüllung eines bleibenden Frontzahnes mittels manueller step-back Aufbereitungstechnik und folgender lateraler Kompaktion**.

Die Literaturrecherche hatte ergeben, dass es während der Aufbereitung und Füllung der Wurzelkanäle erforderlich sei, dass der Zahn durch einen Kofferdam isoliert werde. Der Kofferdam verhindere, dass der Patient Instrumente verschlucke oder aspiriere. Außerdem werden die Erfolgsaussichten sinken, wenn die Wurzelkanalbehandlung ohne Kofferdam durchgeführt werde (Ahmad 2009).

Voraussetzung für eine suffiziente Wurzelfüllung sei eine der Wurzelgröße entsprechende Aufbereitung des Wurzelkanals. Der Wurzelkanal müsse ausreichend groß aufbereitet werden, um zu garantieren, dass das Ende der Spülnadel weit genug in den Wurzelkanal eindringen könne und somit die Spüllösung im apikalen Bereich ihre Wirkung entfalte. Sei dies nicht der Fall, werde die Distanz der Spitze der Nadel zur apikalen Region zu groß und die Spüllösung werde dort nicht wirken (Hsieh *et al.* 2007). Bei der Aufbereitung müsse jedoch auch bedacht werden, dass der Wurzelkanal nicht zu weit in Anbetracht der Größe aufbereitet werde, da hierbei die Zahnschicht zu sehr geschwächt werde (Leo 2012). Die Aufbereitung solle den Raum, den die Pulpa während der Jugendzeit ausfüllte, nicht überschreiten (Clark 2005). Dies bedeute, der Wurzelkanal solle anatoförmig aufbereitet werden. Die Aufbereitung habe sich vom koronalen Zugang bis zur apikalen Spitze am ursprünglichen Verlauf zu orientieren (Himmel und Levitan 2003).

Bei der Aufbereitung per „Step-back“ sei es wichtig eine Feile zu wählen, mit der die apikale Region ausreichend aufbereitet werde. Im weiteren Verlauf werde mit zunehmender Feilengröße die Eindringtiefe reduziert. Werden Fehler bei dieser Art der Aufbereitung begangen, sei es später nicht möglich die Wurzelfüllung so zu

gestalten, dass alle Bereiche des Wurzelkanals dicht abgefüllt seien (Goodman et al. 1985).

Die Instrumente und deren Material müssten auf die Geometrie des Wurzelkanals abgestimmt werden. Werden dies nicht beachtet und seien die Instrumente zu steif, könne es mit zunehmender Instrumentengröße zu Perforationen, *via falsa*, Stufen, Zips und Elbows kommen (Schafer *et al.* 1995). Iatrogene Perforationen während der Gestaltung der Zugangskavität und Aufbereitung des Wurzelkanals seien nicht selten. In der Folge der Behandlung können diese künstlich geschaffenen Irregularitäten meist nur schwer behoben werden (Motamedi 2006). Die Möglichkeit der Versorgung einer Perforation sei abhängig von der Lokalisation, Größe und dem Zeitpunkt (Leo 2012). Die falsche Auswahl und Handhabung der Aufbereitungsinstrumente erhöhe die Gefahr für Frakturen der Instrumente. Zu den Gründen einer Fraktur des Wurzelkanalinstrumentes gehöre zudem die zu häufige Anwendung. Komme es zur Instrumentenfraktur, sei es von Bedeutung, ob eine Entfernung des Fragmentes und damit eine nachfolgende Wurzelfüllung möglich sei. Könne das Fragment entfernt werden, sei eine anschließende Wurzelfüllung möglich. Sei dies nicht möglich oder nur durch weitere Schwächung der Zahnschubstanz, werden die Erfolgsaussichten der Wurzelkanalbehandlung in jedem Fall durch die erschwerte Versorgung sinken. Im schlechtesten Fall, wenn das Fragment nicht entfernt oder die Perforation nicht suffizient gedeckt werden könne, könne es sogar zur Extraktion des Zahnes führen (Parashos und Messer 2006). Aus diesem Grund sei es wichtig, sich vor Behandlungsbeginn auf dem Röntgenbild den Verlauf der Wurzeln zu gegenwärtigen und dementsprechend die Aufbereitungsinstrumente auszuwählen (Li *et al.* 2011).

Die abgefüllte Kanallänge, dürfe in keinem Fall mehr als 2 mm vom röntgenologischen Apex entfernt sein (Wu et al. 2000). Durch die heutzutage zur Verfügung stehenden Methoden der Bestimmung der Arbeitslänge sei es möglich, den Wurzelkanal so nah wie möglich an die apikale Konstriktion aufzubereiten. Diese apikale Konstriktion sei in der Regel 0,5 bis 2 mm vom radiologischen Apex entfernt (European Society of Endodontology 2006). Erfolge die Aufbereitung und anschließend die Füllung des Wurzelkanals kürzer als 2 mm vom radiologischen Apex entfernt, sinke die Erfolgsrate um 20% im Vergleich zu Wurzelkanälen, die im Bereich von 0 bis 2 mm zum radiologischen Apex das Ende der Füllung fänden (Wu *et al.* 2000). Weiter solle vermieden werden, dass die abgefüllte Kanallänge länger

als der röntgenologische Apex sei. Auf Grundlage biologischer und klinischer Prinzipien solle sowohl die Aufbereitung als auch die Füllung nicht über das apikale Foramen hinaus erfolgen (Wu *et al.* 2000). Die Prognose sei deutlich schlechter, wenn Wurzelfüllmaterial über den Apex extrudiert würde. Es gäbe zudem eine Wechselbeziehung zwischen einer zu langen Wurzelfüllung und einer schlechten Prognose des wurzelkanalbehandelten Zahnes (Naito 2005). Andererseits wies Schilder jedoch darauf hin, apikale Puffs seien häufig bei einer Wurzelfüllung und seien ein Zeichen dafür, einen dichten Abschluss der Füllung erreicht zu haben (Schilder 2006). Bei einem apikalen Puff handelt es sich um überextendiertes Sealermaterial.

Die Hauptgründe für einen Misserfolg der Wurzelkanalfüllungen seien laut der Literatur keine adäquate Länge, mangelnde Randständigkeit, Inhomogenitäten oder Blasen der Füllung (Taschieri *et al.* 2011).

Guttapercha solle am Kanaleingang nachkondensiert werden und nachfolgend am Kanaleingang, sprich auf Schmelz-Zement-Grenze, abgetrennt werden. Dies erhöhe zum einen den Grad der Dichtigkeit und zum anderen die Möglichkeit der Kontrolle. Zudem ließen sich Undichtigkeiten durch die vertikale Kompaktion von lateral kondensierten Wurzelkanalfüllungen reduzieren (Yared *et al.* 1997).

Jede Behandlung müsse sorgfältig dokumentiert werden. Zu dieser Dokumentation gehörten, die Befundung, Diagnosestellung, Aufklärungen und Aufzeichnungen der Behandlung, Röntgenbilder und deren korrekte Beschriftung. Dies erleichtere es dem Zahnarzt die durchgeführte Wurzelkanalbehandlung nachzuvollziehen und abrechnen zu können, als auch sich im Rechtsstreit zu verteidigen (Cohen und Schwartz 1987; Cohen 1989).

1.3.2 Beschreibung der praktischen Aufgabenstellungen

Für jede Prüfung wurde eine Aufgabenstellung definiert, anhand derer die Zähne in den Prüfungen bearbeitet werden sollten (Huth, in Bearbeitung). Die erste Prüfung beinhaltete die Präparation einer auf die Karieslokalisation bezogenen Klasse-II-Kavität für direkte Kompositfüllungen. In Anhang 11.1 (Seite 114) ist die Aufgabenstellung angefügt.

In der zweiten Prüfung sollte eine anatomisch korrekte und ästhetisch hochwertige direkte Kompositfüllung mit Approximalkontakten im Seitenzahnbereich gestaltet werden. In Anhang 11.2 (Seite 114) findet man die Aufgabenstellung für die zweite Prüfung.

Die dritte Prüfung umfasste die Präparation einer klassischen Stufenteilkrone für eine indirekte Goldrestauration. In Anhang 11.3 (Seite 114) ist die Aufgabenstellung zu finden.

Die folgende vierte Prüfung umfasste die Präparation für eine Keramikteilkrone. Die Aufgabenstellung für die Präparation der Keramikteilkrone ist in Anhang 11.4 (Seite 115) zu finden.

Die fünfte Prüfung beschäftigte sich mit der Trepanation eines bleibenden Molaren und der Schaffung der Zugangskavität. Im Anhang unter 11.5 (Seite 115) ist die Aufgabenstellung aufgeführt.

In der letzten und sechsten Prüfung sollte an einem Frontzahn eine Wurzelkanalfüllung angefertigt werden. In Anhang 11.6 (Seite 115) findet man die Aufgabenstellung für die sechste Prüfung.

1.3.3 Nominal group technique

Es nahmen die folgenden vier Zentren teil: München, Freiburg, Frankfurt und Leipzig vertreten je durch den Kursleiter des Phantomkurses (Prof. Huth, Prof. Hahn, Prof. Gerhard Szep und Prof. Haak).

Die Erstellung der Bewertungskriterien ist Inhalt einer Projektarbeit von Fr. Prof. Huth im Rahmen des Masterstudiengangs für Medical Education (Huth, in Bearbeitung).

Durch das Zentrum München erfolgte die Auflistung der Bewertungskriterien für alle sechs Prüfungsarten (siehe Literaturrecherche Kapitel 1.3, S. 15-30). Die an der Studie teilnehmenden Zentren Freiburg, Leipzig und Frankfurt bekamen diese Auflistungen zugesandt, um ihrer Meinung nach fehlende Kriterien hinzuzufügen oder Kriterien zu streichen, die sich inhaltlich überschneiden oder eine zu geringe klinische Relevanz hatten. Die Ergebnisse wurden vom Zentrum München zusammen gefasst und für die nächste Phase der nominal group technique vorbereitet, die im Anschluss von allen vier Zentren separat durchgeführt wurde.

Im nächsten Schritt der nominal group technique (Horton 1980; Gallagher *et al.* 1993; Carney *et al.* 1996) wurden dabei pro Bewertungskriterium 3, 2 oder 1 Punkt, entsprechend der vom Bewerter eingeschätzten klinischen Wertigkeit und Wichtigkeit, vergeben. Dieser Arbeitsschritt wurde durchgeführt, um die Kriterien nach ihrer Wichtigkeit zu ordnen.

Auf Grundlage der nominal group technique ergab sich eine Gewichtung der Kriterien. Die Kriterien wurden gemäß der Punktebewertung aufgelistet. Grund für eine niedrige Bewertung war eine geringe klinische Bedeutung sowie geringe inhaltliche Abweichung zu anderen Kriterien. Gestrichen wurden im Folgenden die blau markierten Kriterien. Diese Kriterien erhielten die geringste Punktezahl. Auf diese Art und Weise hatte man pro Prüfung 10-15 Kriterien erhalten. Dies steigerte die Praktikabilität der Anwendung der Kriterien.

Die in dieser Studie verwendeten Bewertungskriterien dienten darüber hinaus einer weiteren bisher nicht veröffentlichten Studie der Poliklinik für Zahnerhaltung als Grundlage. Dabei ging es um die Übereinstimmung zwischen der Selbsteinschätzung der Studierenden und der Bewertung durch die Ausbilder in Anhängigkeit von verschiedenen Trainingsmethoden für die Studierenden (Huth *et al.*, in Bearbeitung).

Nachfolgend sind die Tabellen mit den Ergebnissen der nominal group technique aufgeführt.

Präparation einer auf die Karieslokalisierung bezogenen Klasse-II-Kavität für direkte Kompositfüllungen

1.	Zahn verletzt (zu präparierender Zahn und/oder Nachbarzähne)	12
2.	Kavität zu tief (tiefster Punkt der Zentralfissur richtung Pulpa >2,5 mm)	11
3.	Approximaler Kasten zu tief richtung zervikal (absolute Trockenlegung durch Kofferdam nicht möglich)	11
4.	Präparation im Schmelz nicht tief genug, keine vollständige Kariesentfernung	10
5.	Extension des approximalen Kastens in mesio-distaler Richtung zu breit (>2,0 mm)	10
6.	Präparationsgrenze und nachfolgender Füllungsrand im approximalen Kasten in oro-vestibulärer Richtung nicht sondierbar (nicht überprüfbar)	10
7.	Nicht patientenäquivalent gearbeitet (ohne Wasserkühlung, Handschuhe, Mundschutz, Schutzbrille)	10
8.	Isthmus zu breit in bucco-oraler Richtung (unnötige Substanzschwächung)	9
9.	Approximale Präparationsgrenze liegt im approximalen Kontaktpunktbereich	9
10.	Anschrägung zu stark (Kastenboden u. -flanken, Substanzschonung) (>1	9

	mm)	
11.	Präparationsgrenze undefiniert (klinische Relevanz: lose Schmelzprismen)	9
12.	Kontaktpunkt nicht vollständig aufgelöst	8
13.	Gleichmäßige approximale Separation zum Nachbarzahn (0,5 mm)	7
14.	Übergang pulpaaxial nach okklusal nicht gebrochen	5

Tabelle 1: NGT- Analyse für Prüfung Präparation Klasse II für Komposit. In der linken und mittleren Spalte sind die Bewertungskriterien aufgelistet. In der rechten Spalte ist pro Bewertungskriterium die Summe der von den 4 Zentren vergebenen Punkte der NGT-Analyse aufgelistet.

Anatomisch korrekte und ästhetisch hochwertige direkte Kompositfüllung mit Approximalkontakten im Seitenzahnbereich

1.	Zahn verletzt (zu präparierender Zahn und/oder Nachbarzähne)	12
2.	1 Approximalkontakt fehlt	12
3.	Beide Approximalkontakte fehlen	12
4.	Füllung locker	12
5.	Randspalt	11
6.	1 Approximalkontakt verblockt (mit Bonding)	10
7.	Beide Approximalkontakte verblockt (mit Bonding)	10
8.	Material inhomogen und/oder insuffiziente Inkrementadaptation	10
9.	Nicht patientenäquivalent gearbeitet (z.B. Ausarbeitung ohne Wasser, Mundschutz, Handschuhe, Schutzbrille)	10
10.	Unterschuss und/oder Überschuss des Füllungsmaterials am Füllungsrand	9
11.	Modellation Kaufläche insuffizient (siehe Aufgabenstellung)	9
12.	Keine korrekte Lage eines Approximalkontakts in inzisal-zervikaler Richtung	8
13.	Keine korrekte Lage beider Approximalkontakte in inzisal-zervikaler Richtung	8
14.	1 Approximalkontakt zu schwach und/oder zu punktförmig	8
15.	Beide Approximalkontakte zu schwach und/oder zu punktförmig	8
16.	Insuffiziente Ausarbeitung und/oder Politur unzureichend	7
17.	Nicht korrekte Farbauswahl	6

Tabelle 2: NGT- Analyse für Prüfung Legen einer Kompositfüllung Klasse II. In der linken und mittleren Spalte sind die Bewertungskriterien aufgelistet. In der rechten Spalte ist pro Bewertungskriterium die Summe der von den 4 Zentren vergebenen Punkte der NGT-Analyse aufgelistet.

Präparation einer klassischen Stufenteilkrone für eine indirekte Goldrestauration

1.	Zahn verletzt (zu präparierender Zahn und/oder Nachbarzähne)	12
2.	Gefahr Pulpaverletzung wegen zu starkem Substanzabtrag (approximal in mesio-distaler Richtung; okklusal: Isthmus und Höckerspitzenbereich)	12
3.	Präparationsgrenze und nachfolgender Restaurationsrand nicht sondierbar (nicht überprüfbar)	12
4.	Präparation untersichgehend	12
5.	Undefinierte Präparationsgrenze	12

6.	Mindestschichtstärke der Goldrestauration nicht eingehalten (okklusal) (0,8 mm)	11
7.	Öffnungswinkel der Präparation zu gering	11
8.	Kontinuierliche Übergänge der Anschrägung	11
9.	Scherhöcker Kaukantenschutz: Außenschliff zu schmal (Kontaktpunkt am Präparationsrand)	10
10.	Approximal (Kastenboden und Flanken) Federrand zu schmal oder fehlend	10
11.	Nicht patientenäquivalent gearbeitet (z.B. Ausarbeitung ohne Wasser)	10
12.	Keine proximale Stufe präpariert	9
13.	Öffnungswinkel der Präparation zu groß (Retention)	9
14.	Scharfe Ecken am Übergang approximaler Federrand-Höckerschutz/Kaukantenschutz (Problem: Gold-Guß)	9
15.	Isthmus zu breit	8
16.	Genügende Außerkontaktstellung (für die Abformung)	8
17.	Stützhöcker: Stufe zu breit	8
18.	Isthmus-, Kastenböden fallen ab	7
19.	Stützhöcker: Stufe abfallend (mesio-distal und/oder richtung zervikal)	7
20.	Innenkanten zu scharf	7
21.	Präparation nicht zahnachsengerecht	7
22.	Kavität schlecht finiert (Riefen)	7
23.	Stützhöcker: Stufe zu hoch (zu wenig Retentionsfläche)	6
24.	Stützhöcker: Stufe zu tief (maximal auf Höhe des Isthmusbodens)	6
25.	Stützhöcker: Stufe zu schmal	6
26.	Palatinal Anschrägung zu schmal	6
27.	Palatinal Anschrägung zu breit	6
28.	Falsch liegende okklusale Kontaktpunkte	6
29.	Approximal Federrand zu breit	6

Tabelle 3: NGT- Analyse für Prüfung Präparation Gold TK. In der linken und mittleren Spalte sind die Bewertungskriterien aufgelistet. In der rechten Spalte ist pro Bewertungskriterium die Summe der von den 4 Zentren vergebenen Punkte der NGT-Analyse aufgelistet.

Präparation für eine Keramikteilkrone

1.	Zahn verletzt (zu präparierender Zahn und/oder Nachbarzähne)	12
2.	Gefahr Pulpaverletzung durch zu großen Substanzabtrag (Höckerspitzen, Isthmus, approximaler Kasten in mesio-distaler Richtung)	12
3.	Präparationsgrenze zu tief in Richtung zervikal (Trockenlegung nicht möglich)	12
4.	Undefinierte Präparationsgrenze	12
5.	Fälschliche Randanschrägung	12
6.	Mindestschichtstärke für Keramik nicht eingehalten (approximal, okklusal, Höckerbereich) (<1,5 mm)	11
7.	Präparationsgrenze und nachfolgender Restaurationsrand nicht sondierbar (nicht überprüfbar)	11
8.	Scharfe Ecken in der Außenlinie der Präparation	11
9.	Keine eindeutige Endposition der späteren Restauration möglich	11
10.	Präparation untersichgehend	11
11.	90° Winkel nicht eingehalten zwischen Keramik und Zahnschicht	10

12.	Nicht patientenäquivalent gearbeitet (z.B. ohne Wasserkühlung)	10
13.	Isthmus zu schmal (<2 mm)	9
14.	Öffnungswinkel der Präparation zu gering	9
15.	Innenkanten zu scharf	9
16.	Kavität schlecht finiert (Riefen)	8

Tabelle 4: NGT- Analyse für Prüfung Präparation Keramik TK. In der linken und mittleren Spalte sind die Bewertungskriterien aufgelistet. In der rechten Spalte ist pro Bewertungskriterium die Summe der von den 4 Zentren vergebenen Punkte der NGT-Analyse aufgelistet.

Trepanation und Zugangskavität

1.	Perforation	12
2.	Zugangskavität zu klein (keine Übersicht, Gefahr der unvollständiger Entfernung von Pulpagewebe, kein gerader Feilenzugang)	11
3.	Kavität nicht vollständig kariesfrei	11
4.	Zugangskavität zu groß (Schwächung der Restzahnschubstanz)	10
5.	Lage oder Form der Zugangskavität nicht korrekt bzgl. des gewählten Zahnes	10
6.	Pulpdach nicht komplett abgetragen	10
7.	Kanaleingänge nicht ausreichend erweitert (kein geradliniger Feilenzugang)	10
8.	Nicht alle Kanäle gefunden	10
9.	Nicht patientenäquivalent gearbeitet (z.B. ohne Wasserkühlung, ohne Kofferdam, Schutzbrille)	10
10.	Pulpakammerboden anpräpariert	9
11.	Dentinüberhänge an den Kanalzugängen belassen	9
12.	Kanaleingänge zu stark erweitert (Schwächung Restzahnschubstanz)	9
13.	Zugangskavität nicht gereinigt (Kavitätoilette)	6

Tabelle 5: NGT- Analyse für Prüfung Trepanation. In der linken und mittleren Spalte sind die Bewertungskriterien aufgelistet. In der rechten Spalte ist pro Bewertungskriterium die Summe der von den 4 Zentren vergebenen Punkte der NGT-Analyse aufgelistet.

Wurzelkanalfüllung

1.	Abgefüllte Kanallänge ist mehr als 2 mm vom röntgenologischen Apex entfernt (zu kurz)	12
2.	Abgefüllte Kanallänge ist länger als der röntgenologische Apex (Guttapercha)	12
3.	Kanal nicht anatoförmig aufbereitet (Stufen im Kanal, ungenügendes circumferential filing, Aufbereitung entspricht nicht Kanalverlauf)	12
4.	Unvollständige Wurzelfüllung (mangelhafte Randständigkeit, Inhomogenität, Blasen)	11
5.	Via falsa	11
6.	Nicht patientenäquivalent gearbeitet (z.B. ohne Wasserkühlung, ohne Kofferdam, Mundschutz, Handschuhe)	10
7.	Kanal zu weit/zukonisch aufbereitet (Substanzschonung)	9
8.	Kanal in Rö-Kontrollaufnahme zeigt Zip-Elbow	9
9.	Kanal nicht konisch im Sinne der step-back-Technik aufbereitet	8
10.	Kanal in Anbetracht der Wurzelstärke zu gering oder zu weit aufbereitet	8

	(klinische Relevanz: ungenügende mechanische Aufbereitung oder Wurzelperforation)	
11.	Instrumentenfraktur (Entfernung des Fragmentes und/oder nachfolgende Wurzelfüllung nicht möglich)	8
12.	Guttapercha nicht am Kanaleingang abgetrennt (Schmelz-Zement-Grenze)	7
13.	Beschriftung der Arbeitsunterlagen falsch oder unvollständig (Dokumentationspflicht)	7
14.	Guttapercha ist nicht nachkondensiert am Kanaleingang	5

Tabelle 6: NGT- Analyse für Prüfung Wurzelfüllung. In der linken und mittleren Spalte sind die Bewertungskriterien aufgelistet. In der rechten Spalte ist pro Bewertungskriterium die Summe der von den 4 Zentren vergebenen Punkte der NGT-Analyse aufgelistet.

1.3.4 Matrixanalyse

Im Anschluss wurde von allen vier Zentren getrennt eine Matrixanalyse durchgeführt (Huth, in Bearbeitung). Diese Analyse diente dazu, die Kriterien nach ihrer inhaltlichen Relevanz untereinander zu vergleichen. Zu diesem Zweck wurde eine Tabelle erstellt, in der die Kriterienliste sowohl in der Horizontalen als auch Vertikalen eingetragen wurde. Somit war es möglich die einzelnen Bewertungskriterien miteinander zu vergleichen. Es wurden pro Bewertungskriterium 10, 9, 8, 7, 6, 5, 4, 3, 2, 1 Punkte im Verhältnis zum anderen Kriterium in die vorbereitete Tabelle eingetragen, wobei die Summe der beiden zu vergleichenden Kriterien immer 10 ergeben musste. Im Anschluss wurden die Punkte in horizontaler Richtung zusammengezählt. Ausgehend von der Gesamtpunktzahl aller Kriterien (100%) wurde für jedes Kriterium ein inhaltlicher Prozentwert an der Gesamtbewertung berechnet. Um nun einen Gesamtwert und eine endgültige Gewichtung zu erhalten, wurden die vier Werte der Experten für jedes Bewertungskriterium zusammengezählt und prozentual, in Bezug auf den Punktwert aller Kriterien, gewichtet. Die nachfolgenden Tabellen zeigen die zusammengezählten Punktwerte und dazugehörigen Prozentwerte der jeweiligen Bewertungskriterien aus den Matrixanalysen für die sechs verschiedenen Prüfungen (Huth, in Bearbeitung).

Präparation einer auf die Karieslokalisierung bezogenen Klasse-II-Kavität für direkte Kompositfüllungen

Die Matrix-Analyse für die erste Prüfung ergab nachfolgend aufgeführtes Ergebnis: Die „Verletzung des Nachbarzahnes“ sowie „zu tiefe Präparation und damit fehlende Möglichkeit der absoluten Trockenlegung“ wurde am stärksten bewertet. In der nachfolgenden Tabelle sieht man nun die Daten für jedes Kriterium.

Bewertungskriterium	Summe	Gesamt in %
Zahn verletzt (zu präparierender Zahn und/oder Nachbarzähne)	252	11
Approximaler Kasten zu tief Richtung zervikal (absolute Trockenlegung durch Kofferdam nicht möglich)	239	11
Kavität zu tief (tiefster Punkt der Zentralfissur richtung Pulpa >2,5 mm))	225	10
Extension des approximalen Kastens in mesio-distaler Richtung zu breit (>2,0 mm)	216	10
Präparationsgrenze undefiniert (klinische Relevanz: lose Schmelzprismen)	195	9
Nicht patientenäquivalent gearbeitet (ohne Wasserkühlung, Handschuhe, Mundschutz, Schutzbrille)	194	9
Präparation im Schmelz nicht tief genug, keine vollständige Kariesentfernung	192	9
Präparationsgrenze und nachfolgender Füllungsrand im approximalen Kasten in oro-vestibulärer Richtung nicht sondierbar (nicht überprüfbar)	175	8
Isthmus zu breit in bukkal-oraler Richtung (unnötige Substanzschwächung)	174	8
Anschrägung zu stark (Kastenboden u. -flanken, Substanzschonung) (>1 mm)	173	8
Approximale Präparationsgrenze liegt im approximalen Kontaktpunktbereich	165	7

Tabelle 7: Matrix-Analyse für Prüfung 1. In der linken Spalte ist das Kriterium, in der mittleren Spalte die gesamte Summe der Bewertungspunkte aller 4 Zentren und rechts der Gesamtprozentwert zu sehen.

Anatomisch korrekte und ästhetisch hochwertige direkte Kompositfüllung mit Approximalkontakten im Seitenzahnggebiet

Die Matrix-Analyse für die zweite Prüfung hatte die unten aufgeführte Tabelle zur Folge. Es ist zu sehen, dass auch bei dieser Prüfung die „Schädigung des Nachbarzahns“ am stärksten bewertet wurde. Eine hohe Wertigkeit erfuhr auch die „lockere Füllung“, „ein Randspalt“ sowie die „falsche Gestaltung der Approximalkontakte“.

Bewertungskriterium	Summe	Gesamt in %
Zahn verletzt (zu präparierender Zahn und/oder Nachbarzähne)	267	12
Füllung locker	259	12
Beide Approximalkontakte fehlen	239	11
Randspalt	209	10
1 Approximalkontakt fehlt	205	9
Unterschuss und/oder Überschuss des Füllungsmaterials am Füllungsrand	205	9
Beide Approximalkontakte verblockt (mit Bonding)	184	9
1 Approximalkontakt verblockt (mit Bonding)	163	7
Materialoberfläche inhomogen und/oder insuffiziente Inkrementadaptation	160	7
Modellation Kaufläche insuffizient (siehe Aufgabenstellung)	160	7
Nicht patientenäquivalent gearbeitet (z.B. Ausarbeitung ohne Wasser, Mundschutz, Handschuhe, Schutzbrille)	149	7

Tabelle 8: Matrix-Analyse für Prüfung 2. In der linken Spalte ist das Kriterium, in der mittleren Spalte die gesamte Summe der Bewertungspunkte aller 4 Zentren und rechts der Gesamtprozentwert zu sehen.

Präparation einer klassischen Stufenteilkrone für eine indirekte Goldrestauration

Die Bewertungskriterien wurden auch für diese Prüfung einer Matrix-Analyse unterzogen. Die nachfolgend gezeigte Tabelle beinhaltet die Ergebnisse der Matrix-Analyse für die dritte Prüfung. Man sieht, dass die maximalen Prozentwerte 9 und die minimalen 7 betragen. Die „Gefahr der Verletzung der Pulpa und des Nachbarzahnes“ während der Präparation wurden als am schwerwiegendsten eingestuft. Weiter von Interesse ist die „Form und Gestaltung der Präparation“.

Bewertungskriterium	Summe	Gesamt in %
Gefahr Pulpaverletzung wegen zu starkem Substanzabtrag (approximal in mesio-distaler Richtung; okklusal: Isthmus und Höckerspitzenbereich)	318	9
Zahn verletzt (zu präparierender Zahn und/oder Nachbarzähne)	317	9
Öffnungswinkel der Präparation zu groß (Retention)	290	8
Präparation untersichgehend	273	8
Approximal (Kastenboden und Flanken) Federrand und Anschrägung Stufe zu schmal oder fehlend	273	8
Undefinierte Präparationsgrenze	266	7
Mindestschichtstärke der Goldrestauration nicht eingehalten (okklusal) (0,8 mm)	253	7
Nicht patientenäquivalent gearbeitet (z.B. Ausarbeitung ohne Wasser)	253	7
Scherhöcker Kaukantenschutz: Außenschliff zu schmal (Kontaktpunkt am Präparationsrand)	251	7
Keine approximale Stufe präpariert	246	7
Präparationsgrenze und nachfolgender Restaurationsrand nicht sondierbar (nicht überprüfbar)	245	6
Kontinuierliche Übergänge der Anschrägung	220	6
Scharfe Ecken am Übergang approximaler Federrand-Höckerschutz/Kaukantenschutz (Problem: Gold-Guß)	220	6
Öffnungswinkel der Präparation zu gering	215	5

Tabelle 9: Matrix-Analyse für Prüfung 3. In der linken Spalte ist das Kriterium, in der mittleren Spalte die gesamte Summe der Bewertungspunkte aller 4 Zentren und rechts der Gesamtprozentwert zu sehen.

Präparation Keramik Teilkrone

Die Matrix-Analyse für Prüfung 4 hat folgendes Ergebnis erbracht. Wie bei der Präparation der Goldteilkrone erstrecken sich die Werte über zwei Prozentpunkte. Auch bei dieser Prüfung wurde die „Verletzung der Zahnes und der Pulpa“ mit einer hohen Gewichtung versehen. Des Weiteren sind die „Trockenlegung und Präparationsgeometrie“ mit hoch gewichtet worden. Die folgende Tabelle gibt die genauen Werte für die Präparation einer Keramikteilkrone wieder.

Bewertungskriterium	Summe	Gesamt in %
Zahn verletzt (zu präparierender Zahn und/oder Nachbarzähne)	372	8
Präparationsgrenze zu tief in Richtung zervikal (Trockenlegung nicht möglich)	358	7
Fälschliche Randanschlagung	358	7
Gefahr Pulpaverletzung durch zu großen Substanzabtrag (Höckerspitzen, Isthmus, approximaler Kasten in mesio-distaler Richtung)	354	7
Präparation untersichgehend	316	7

Mindestschichtstärke für Keramik nicht eingehalten (approximal, okklusal, Höckerbereich) (<1,5 mm)	311	7
Nicht patientenäquivalent gearbeitet (z.B. ohne Wasserkühlung)	311	7
Keine eindeutige Endposition der späteren Restauration möglich	310	6
Undefinierte Präparationsgrenze	303	6
90° Winkel nicht eingehalten zwischen Keramik und Zahnschmelze	286	6
Scharfe Ecken in der Außenlinie der Präparation	284	6
Präparationsgrenze und nachfolgender Restaurationsrand nicht sondierbar (nicht überprüfbar)	275	6
Öffnungswinkel der Präparation zu gering	266	6
Isthmus zu schmal (<2 mm)	254	5
Innenkanten zu scharf	248	5
Kavität schlecht finiert (Riefen)	194	4

Tabelle 10: Matrix-Analyse für Prüfung 4. In der linken Spalte ist das Kriterium, in der mittleren Spalte die gesamte Summe der Bewertungspunkte aller 4 Zentren und rechts der Gesamtprozentwert zu sehen.

Trepanation und Zugangskavität

Die unten aufgeführte Tabelle beinhaltet die Ergebnisse der Matrix-Analyse für die fünfte Prüfung. Das Ergebnis zeigt, dass die „Perforation“ am stärksten, gefolgt von „nicht alle Kanäle gefunden“, bewertet wurde. Am Ende der Tabelle und damit am schwächsten bewertet ist das Kriterium, dass der „Pulpaboden anpräpariert“ wurde.

Bewertungskriterium	Summe	Gesamt in %
Perforation	322	12
Nicht alle Kanäle gefunden	252	10
Kanaleingänge zu stark erweitert (Schwächung Restzahnschmelze)	221	8
Nicht patientenäquivalent gearbeitet (z.B. ohne Wasserkühlung, ohne Kofferdam, Schutzbrille)	215	8
Lage oder Form der Zugangskavität nicht korrekt bzgl. des gewählten Zahnes	213	8
Kavität nicht vollständig kariesfrei	212	8
Zugangskavität zu groß (Schwächung der Restzahnschmelze)	209	8
Pulpadach nicht komplett abgetragen	207	8
Dentinüberhänge an den Kanalzugängen belassen	200	8
Kanaleingänge nicht ausreichend erweitert (kein geradliniger Feilenzugang)	199	8

Zugangskavität zu klein (keine Übersicht, Gefahr der unvollständiger Entfernung von Pulpagewebe, kein gerader Feilenzugang)	196	7
Pulpakammerboden anpräpariert	194	7

Tabelle 11: Matrix-Analyse für Prüfung 5. In der linken Spalte ist das Kriterium, in der mittleren Spalte die gesamte Summe der Bewertungspunkte aller 4 Zentren und rechts der Gesamtprozentwert zu sehen.

Wurzelkanalfüllung

Die letzte Prüfung beschäftigt sich mit der Wurzelfüllung an einem Frontzahn durch Step-back-Technik. Die Matrix-Analyse hat ergeben, dass die „Instrumentenfraktur“, „zu lange Wurzelfüllung“ und „Via falsa“ die größten Werte bekommen haben. Am Ende der Tabelle finden sich die Kriterien „Beschriftung der Arbeitsunterlagen“ fehlerhaft und „Guttapercha nicht am Kanaleingang abgetrennt“.

Bewertungskriterium	Summe	Gesamt in %
Instrumentenfraktur (Entfernung des Fragmentes und/oder nachfolgende Wurzelfüllung nicht möglich)	312	10
Abgefüllte Kanallänge ist länger als der röntgenologische Apex (Guttapercha)	303	10
Via falsa	302	10
Abgefüllte Kanallänge ist mehr als 2 mm vom röntgenol. Apex entfernt (zu kurz)	265	8
Kanal in Anbetracht der Wurzelstärke zu gering/ oder zu weit aufbereitet (klinische Relevanz: ungenügende mechanische Aufbereitung/ oder Wurzelfraktur)	252	8
Unvollständige Wurzelfüllung (mangelhafte Randständigkeit, Inhomogenität, Blasen)	248	8
Kanal in Rö-Kontrollaufnahme zeigt Zip-Elbow	247	8
Kanal zu weit/zu konisch aufbereitet (Substanzschonung)	231	7
Kanal nicht anatoförmig aufbereitet (Stufen im Kanal, ungenügendes circumferential filling, Aufbereitung entspricht nicht Kanalverlauf)	230	7
Nicht patientenäquivalent gearbeitet (z.B. ohne Wasserkühlung, ohne Kofferdam, Mundschutz, Handschuhe)	223	7
Kanal nicht konisch im Sinne der step-back-Technik aufbereitet	206	7
Beschriftung der Arbeitsunterlagen falsch oder unvollständig (Dokumentationspflicht)	156	5
Guttapercha nicht am Kanaleingang abgetrennt (Schmelz-Zement-Grenze)	145	5

Tabelle 12: Matrix-Analyse für Prüfung 6. In der linken Spalte ist das Kriterium, in der mittleren Spalte die gesamte Summe der Bewertungspunkte aller 4 Zentren und rechts der Gesamtprozentwert zu sehen.

1.3.5 Definition von Durchfallkriterien

Sollte es während der Behandlung zu einer Gefährdung des Patienten oder des Zahnes und in der Folge dessen zum Verlust des Zahnes kommen, führt dies zum sofortigen Durchfallen. Dies konnte durch die prozentuale Gewichtung aus der Matrixanalyse jedoch nicht dargestellt werden.

Daher wurden die folgenden Durchfallkriterien definiert (Huth, in Bearbeitung):

Ein wesentliches Durchfallkriterium war die **Verletzung von Nachbarzähnen** während der Therapie. So wurde beschrieben, dass das Risiko für die Entstehung von Karies in Folge einer iatrogenen Verletzung des Nachbarzahnes deutlich ansteige und damit die restaurative Therapie des Nachbarzahnes notwendig mache (Lussi 1995). Dies konnte im Rahmen der studentischen Ausbildung nicht toleriert werden.

Ein weiteres Durchfallkriterium war die Traumatisierung in Folge **nicht patientenäquivalenten Arbeitens**. Eine Wasserkühlung des Schleifkörpers und des Zahnes sei notwendig, da die Pulpa des Zahnes einen Temperaturanstieg von mehr als 5,5°C nicht tolerierte. Die Folge könnte eine temporäre oder dauerhafte Schädigung der dentalen Pulpa sein (Cavalcanti *et al.* 2002).

Ein zusätzliches Durchfallkriterium war eine **lockere Füllung** in Folge der Anfertigung einer Füllung, da daraus eine Neuanfertigung der Füllung resultierte. Die Konsequenz aus einer lockeren Füllung sei Sekundärkaries. Zur Vermeidung von Sekundärkaries müssten bei der Herstellung von Füllungen einige Kriterien erfüllt werden. Zu diesen gehöre, dass der Rand der Füllung dicht und die Füllung fest sei, damit keine kariesbildenden Prozesse zwischen Zahn und Füllungen stattfinden könnten (Burke *et al.* 1999). Sekundärkaries sei der Hauptgrund für die Neuanfertigung einer Füllung, unabhängig vom Füllungsmaterial. Dies konnte im Rahmen der studentischen Ausbildung nicht akzeptiert werden.

Ebenfalls war ein Durchfallkriterium, wenn **die Gefahr der Pulpaverletzung durch zu großen Substanzabtrag** während der Präparation des Zahnes gegeben war. Bei der Präparation eines Zahnes sei es wichtig die Anatomie des Zahnes und deren Strukturen zu kennen, um diesen nicht dauerhaft zu traumatisieren. Die weitreichendste Traumatisierung des Zahnes sei die Verletzung der Pulpa, die bis zur Wurzelkanalbehandlung führen könnte. An Hand von Untersuchungen wurde

dargelegt, dass eine Restdentindicke von 0,5 mm notwendig sei, um Pulpaverletzungen zu vermeiden (Murray *et al.* 2003). Sollte es zu einer Verletzung der Pulpa in Folge der Präparation des Zahnes kommen, führte dies zum sofortigen Durchfallen.

Ein weiteres Durchfallkriterium war die **fehlende Möglichkeit der Trockenlegung** beim Eingliedern keramischer Arbeiten, da die **Präparationsgrenze zu tief zervikal** lag. Eine ausreichende Trockenlegung sei notwendig, um eine zahntechnische Arbeit aus Keramik einzusetzen. Um dies gewährleisten zu können, müsse die Präparation supra- bis isogingival enden. Die Ränder der Präparation sollten auf jeden Fall an der Schmelz-Zement-Grenze enden. Wenn die Präparationsgrenze tiefer läge, müsste von einer Verminderung der Adhäsion ausgegangen werden (Federlin *et al.* 2004).

Ein Durchfallkriterium zudem war es, wenn bei der Trepanation **nicht alle Wurzelkanäle gefunden** wurden. Das Ziel der Wurzelkanalbehandlung sei es, den Zahn zu erhalten und eine Beschwerdefreiheit für den Patienten zu erreichen. Beschwerdefreiheit sei nur zu erlangen, wenn das gesamte Pulpagewebe entfernt werde. Dies bedeute, dass alle Wurzelkanäle aufgesucht und therapiert werden müssten. Eine gute Kenntnis über die Anatomie des Zahnes und des Wurzelkanalsystems sei Voraussetzung für eine erfolgreiche Wurzelkanalbehandlung (Li *et al.* 2011).

Zusätzlich wurde zu den Durchfallkriterien gezählt, wenn es bei der Trepanation oder der Aufbereitung der Wurzelkanäle zu einer **Perforation** kam. Bei einer Perforation komme es zu einer vom Behandler geschaffenen Kontinuitätsunterbrechung der Wände oder des Bodens der Pulpakammer (Cohen 1989). Der Autor beschreibt, dass die Perforation im Bereich der Furkation ein Hauptgrund für einen Misserfolg der Wurzelkanalbehandlung sei. Sollte es im Rahmen der Wurzelkanalbehandlung zu einer Perforation kommen, bedeutete dies, dass der Studierende die Prüfung nicht bestanden hatte, da dies nicht während der studentischen Ausbildung toleriert werden konnte.

Darüber hinaus war ein Kriterium, das zum Durchfallen führte, wenn die angefertigte **Wurzelfüllung länger als der radiologische Apex** war. Die Wurzelfüllung sollte der Anatomie des Wurzelkanals entsprechen, diesen komplett ausfüllen und 0-2 mm vor dem radiologischen Apex enden (Naito 2005). Laut dieser Arbeit war die Prognose schlechter und es könnte zum Misserfolg führen, sollte Wurzelfüllungsmaterial über

den radiologischen Apex extrudiert werden.

Ein weiteres Durchfallkriterium bei der endodontischen Behandlung war die **Via falsa**. Bei einer Via falsa wird der korrekte Verlauf des Wurzelkanals verlassen. Dies geschehe iatrogen durch die Anwendung des eingebrachten Instrumentes. Die Prognose für den Erfolg der Wurzelkanalbehandlung sinke in diesem Fall deutlich ab und hänge davon ab in welchem Bereich des Wurzelkanals die Via falsa zu Stande käme und ob diese gedeckt werden könnte. Eine Via falsa sei bei der Schaffung einer Zugangskavität oder der Präparation des Wurzelkanals nicht unüblich und reduziere die Erfolgsaussichten (Motamedi 2006). Dies könnte dazu führen, dass der Zahn nicht zu erhalten sei und wurde deshalb in der studentischen Ausbildung nicht gestattet.

Ein letztes Durchfallkriterium war die **Instrumentenfraktur**, wenn daraus resultierte, dass eine adäquate Wurzelfüllung nicht möglich war, da das Fragment nicht entfernt werden konnte. Sollte es nicht möglich sein eine adäquate Wurzelfüllungen anzufertigen in Folge der Instrumentenfraktur führe dies zum Misserfolg der Behandlung (Parashos und Messer 2006) und wurde daher während der Prüfung nicht akzeptiert.

Im Folgenden sind pro Prüfung die Durchfallkriterien aufgelistet:

Präparation einer auf die Karieslokalisation bezogenen Klasse-II-Kavität für direkte Kompositfüllungen:

- Zahn verletzt (zu präparierender Zahn und/oder Nachbarzähne)
- Approximaler Kasten zu tief Richtung zervikal (absolute Trockenlegung durch Kofferdam nicht möglich)

Anatomisch korrekte und ästhetisch hochwertige direkte Kompositfüllung mit Approximalkontakten im Seitenzahnbereich:

- Füllung locker
- Zahn verletzt (zu präparierender Zahn und/oder Nachbarzähne)

Präparation einer klassischen Stufenteilkrone für eine indirekte Goldrestauration:

- Gefahr Pulpaverletzung wegen zu starkem Substanzabtrag (approximal in mesio-distaler Richtung; okklusal: Isthmus und Höckerspitzenbereich)
- Zahn verletzt (zu präparierender Zahn und/oder Nachbarzähne)

Präparation einer Keramikteilkrone:

- Zahn verletzt (zu präparierender Zahn und/oder Nachbarzähne)
- Präparationsgrenze zu tief in Richtung zervikal (Trockenlegung nicht möglich)

Trepanation und Zugangskavität:

- Perforation
- Nicht alle Kanäle gefunden

Wurzelfüllung:

- Instrumentenfraktur (Entfernung des Fragmentes und/ oder nachfolgende WF nicht möglich)
- Abgefüllte Kanallänge ist länger als der röntgenologische Apex (Guttapercha)
- Via falsa

1.3.6 Erstellung eines Notenschlüssels

Anhand der Approbationsordnung für Ärzte (ÄAppO 2002) wurde pro Arbeit ein Notenschlüssel erstellt (Huth, in Bearbeitung). Dieser bezieht sich auf die prozentuale Gewichtung, die sich aus der Matrixanalyse ergeben hat. Die Fehlerkriterien wurden von 100% abgezogen. Zum Beispiel wurde die Note „sehr gut“ vergeben, wenn eine Arbeit nach Abzug der gemachten Fehler noch mehr als 75% über die Bestehensgrenze von 60% hinaus erreichte. Dies entspricht einer Gesamtbenotung von mehr als 90% der erreichbaren 100%.

Die Bestehensgrenze liegt bei mindestens 60%, das heißt < 40% an Fehlerkriterien.

Berechnung der Noten anhand der Gewichtung der Matrixanalyse:

Sehr gut (1):	>75% über 60% hinaus	=	>90%
Gut (2):	50-75% über 60% hinaus	=	80-90%
Befriedigend (3)	25-50% über die 60% hinaus	=	70-80%
Ausreichend (4)	0-25% über die 60% hinaus	=	60-70%
Mangelhaft (5)		=	nicht bestanden

2 Fragestellung

Das Ziel dieser Studie war es, die für die praktischen Studentarbeiten im Phantomkurs der Zahnerhaltung erstellten Bewertungskriterien hinsichtlich ihrer Reproduzierbarkeit durch Studierende und Ausbilder im Vergleich mit einer Bewertung per Augenschein zu beurteilen.

Aus dieser Fragestellung wurde die Hypothese formuliert, dass die Ausbilder und die Studierenden unter Verwendung der Kriterien eine signifikant bessere intraexaminer Reliabilität als durch die Bewertung per Augenschein haben.

Von Interesse war außerdem, wie stark die Abweichung zwischen der Bewertung durch die Studierenden und die Bewertung durch die Ausbilder unter Verwendung der definierten Bewertungskriterien bzw. die Bewertung per Augenschein ist. Dies wurde deskriptiv beschrieben. Als Hypothese wurde formuliert, dass die Abweichung der Bewertungen zwischen Ausbildern und Studierenden unter Verwendung der Bewertungskriterien geringer als mithilfe der Bewertung per Augenschein ist.

Ein Fragebogen sollte klären wie die Ausbilder und Studierenden die Bewertungskriterien einschätzten. Zu diesem Zweck wurden fünf Themenkomplexe erstellt. Diese Themenkomplexe beinhalteten den Einfluss der vordefinierten Bewertungskriterien auf das praktische Arbeiten, die Güte der Bewertung, die Kommunikation und den Zeitaufwand. Außerdem sollte geklärt werden, welche der praktischen Arbeiten mehr von der Verwendung der Kriterien profitierten als andere aus Sicht der Ausbilder.

3 Material und Methode

3.1 Ort und Zeitraum der Studiendurchführung

Die Studie wurde im Phantomkurs der Poliklinik für Zahnerhaltung und Parodontologie der LMU unter Leitung von Fr. Prof. Huth durchgeführt.

Es nahmen an der Überprüfung der Verständlichkeit zwei Zahnärzte (Dr. Michael Nemecek und Dr. Hans-Georg Kirchner) sowie ein Studierender (Maximilian Kollmuß) teil.

An der Anwendung der Bewertungskriterien nahmen zwei Zahnärzte (Dr. Michael Nemecek und Dr. Hans-Georg Kirchner) und zwei Studierende (Christopher Schmid und Jana Kleinschmidt) teil.

Die praktischen Arbeiten wurden im Wintersemester 2012/2013 angefertigt und anschließend bewertet.

Im Anschluss an die praktischen Prüfungen wurde die Befragung durchgeführt. An der Befragung nahmen auf der einen Seite die drei regelhaft den Phantomkurs betreuenden Zahnärzte (Dr. Franz Michael Jakob, Dr. Hans-Georg Kirchner, Dr. Michael Nemecek), zwei weitere Zahnärzte Dr. Ann-Christin von Köckritz, ZA Moritz Baumann), die den Kurs in der Vergangenheit betreut hatten, sowie die Kursleiterin (Fr. Prof. Huth) teil. Auf der anderen Seite nahmen 42 Studierende des Sommersemesters 2013 an der Befragung teil.

Die Studiendauer erstreckte sich über zwei Jahre, von September 2011 bis Juli 2013.

3.2 Überprüfung der Verständlichkeit der Bewertungskriterien

Es wurde die Verständlichkeit der Kriterien für Ausbilder und Studierende anhand von je zwei exemplarischen Studentendarbeiten früherer Semester pro Notenstufe überprüft.

Es wurden aus bereits bewerteten, archivierten praktischen Studentendarbeiten der früheren Semester pro Notenstufe (1-5) je zwei Arbeiten für alle Prüfungsthemen herausgesucht. Diese Arbeiten wurden blind von zwei Ausbildern und einem Studierenden anhand der Kriterien benotet. Im Anschluss an die Bewertung erfolgte

ein Gespräch mit den beiden Ausbildern und dem Studierenden um herauszufinden, ob Verständnisprobleme oder Unklarheiten in der Formulierung bestanden. Der Projektleiter analysierte anhand der Berechnung des Intraklassen Korrelationskoeffizienten (ICC) und der gemittelten Abweichung der Bewertungsdurchgänge die Daten. Wären Verständnisprobleme in der Formulierung aufgetreten, die sich in den Ergebnissen widerspiegelt hätten, wären die jeweiligen Formulierungen entsprechend angepasst worden. Dies war nicht der Fall.

3.3 Vergleich zwischen der Bewertung mit den erstellten Kriterien und der Bewertung per Augenschein

Während des Phantomkurses der Zahnerhaltung wurden die zuvor erwähnten sechs verschiedenen Arbeiten während der Prüfungen von den Studierenden angefertigt.

Es wurden pro Themenkomplex 20 Studentenarbeiten zufällig ausgewählt und anonymisiert. Zunächst benoteten zwei Ausbilder nach dem bisher verwendeten Bewertungssystem per Augenschein (Vann *et al.* 1983; Jenkins *et al.* 1998) in ganzen Noten von 1-4, bestanden, und Note 5 durchgefallen. Nach zwei Wochen wurde die Benotung nach dem Bewertungssystem per Augenschein wiederholt. Nach erneut zwei Wochen erfolgte die Benotung nach den erstellten Bewertungskriterien, was wiederum nach zwei Wochen wiederholt wurde. Dieses Vorgehen sollte verhindern, dass sich die bewertenden Studienteilnehmer an die bereits benoteten Arbeiten erinnerten.

Dies wurde ebenso durch zwei Studierende mit denselben 20 Studentenarbeiten durchgeführt.

Sowohl die Ausbilder als auch die Studierenden bewerteten die Prüfungsarbeiten allein ohne sich beraten zu können.

3.4 Fragebogen für Studierende und Ausbilder zum Thema Evaluation objektiver Bewertungskriterien für praktische Studentenarbeiten

Im weiteren Verlauf wurde ein Fragebogen entwickelt. Es sollte über die Befragung der beteiligten Studierenden und Ausbildern die subjektive Wahrnehmung der

Bewertungskriterien bezüglich des Lernfortschritts, Transparenz und zeitlichen Bedarf eruiert werden.

Die Fragebögen für Studierende und Ausbilder unterschied sich in einigen Fragen, um gezielt auf die jeweiligen Gruppen einzugehen.

Es gab verschiedene Fragenkomplexe, die sich mit einzelnen Themenschwerpunkten auseinandersetzten. Der erste Themenkomplex beinhaltete Aussagen, die sich mit dem **Einfluss auf das praktische Arbeiten** auseinandersetzten. Der zweite Themenkomplex befasste sich mit Aussagen über den **Einfluss der Kriterien auf die Güte der Bewertung**. Der dritte Komplex fasste Aussagen zusammen, die sich mit dem **Einfluss der Kriterien auf die Kommunikation** beschäftigten. Im vierten und letzten Komplex wurden die beiden Bewertungsverfahren miteinander verglichen. Die Fragen waren mit einer Likert Skala (Norman 2010) zu beantworten. Für die ersten Themenkomplexe gab es Auswahlmöglichkeit aus 4 Intervallen (1 trifft gar nicht zu, 2 trifft eher nicht zu, 3 trifft eher zu, 4 trifft voll zu). Im letzten Fragenkomplex, in dem die verschiedenen Prüfungsthemen miteinander verglichen wurden, gab es ein zusätzliches fünftes Auswahlfeld (klare Verbesserung, Verbesserung, Neutral, Verschlechterung, klare Verschlechterung). Am Schluss der Befragung gab es die Möglichkeit die positiven Aspekte und den Verbesserungsbedarf in einem Freitext aufzuführen.

Die Befragung durch die Fragebögen wurde im Anschluss an die praktischen Arbeiten des Phantomkurses der Zahnerhaltung durchgeführt. An der Befragung nahmen 42 Studierende und 6 Ausbilder teil. Jeder der Befragten hat den Fragebogen für sich allein bearbeitet.

In Anhang 7 (Seite 117-120) sind der Fragebogen der Studierenden, in Anhang 8 (Seite 121-126) der Fragebogen der Ausbilder angefügt.

3.5 Datenanalyse

Die statistische Auswertung der Daten erfolgte mit der Software SPSS, Version 21 (Stakenas und Merrick 1982). Zur Überprüfung der Verständlichkeit und der Bestimmung der intrarater Reliabilität wurde der Intraklassenkorrelationskoeffizient (ICC) gebildet.

Die Intraklassenkorrelation ist ein parametrisches statistisches Verfahren zur Quantifizierung der Übereinstimmung zwischen mehreren Beurteilern in Bezug auf mehrere Beobachtungsobjekte (Shrout 1979; Wikipedia 2013).

Von einer reliablen Beobachtung kann ausgegangen werden, wenn die Unterschiede zwischen den Messobjekten relativ groß sind und gleichzeitig die Varianz zwischen den Bewertungsdurchgängen in Bezug auf die Messobjekte klein. Bei großer Urteils Konkordanz, also geringer Varianz zwischen den Einschätzungswerten, liegt der ICC dann hoch (Wikipedia 2013).

Wie bei anderen Korrelationskoeffizienten kann der ICC Werte zwischen -1,0 und +1,0 annehmen. Da Reliabilitätsmaße definitionsgemäß auf einen Wertebereich von 0 bis 1 beschränkt sind, indizieren negative ICCs eine Reliabilität von 0 (Wirtz 2002; Wikipedia 2013).

Graphisch wurden die gemittelten Abweichungen der Bewertungen von je 20 Arbeiten pro Prüfungsart mit Microsoft Excel in Balkendiagrammen dargestellt. Dazu wurden die Differenzen der beiden Durchgänge per Augenschein und Kriterien gesondert gebildet und daraus der Mittelwert errechnet. Dies wurde für alle sechs Prüfungen und alle vier an der Studie teilnehmenden Bewerter angewandt. Die nun für die jeweiligen Fragenstellungen relevanten Werte wurden gegenüber gestellt.

Die Auswertung der Fragebögen erfolgte zunächst durch die Software SPSS, Version 21 zur Erstellung der Balkendiagramme. Im Anschluss wurden diese deskriptiv beschrieben.

4 Ergebnisse

4.1 Überprüfung der Verständlichkeit der Bewertungskriterien

Von 42 Studierenden wurden pro Themenkomplex 42 Arbeiten im Rahmen der Prüfungen angefertigt. Zur Überprüfung der Verständlichkeit wurden von diesen 42 Arbeiten pro Thema 12 Arbeiten gleichmäßig verteilt auf die Notenstufen zufällig ausgewählt. Diese Arbeiten wurden durch „Random Sampling“ mit der Software SPSS, Version 21 anonymisiert und mittels Kriterienbewertung durch zwei Ausbilder und einen Studierenden begutachtet.

Der Intraklassenkorrelationskoeffizient wurde für die drei möglichen Konstellationen erstellt. Die Konstellationen lauteten: A: Ausbilder 1 und Ausbilder 2, B: Ausbilder 1 und Studierender und als letztes C: Ausbilder 2 und Studierender. Die Berechnung der Übereinstimmung der Beurteilung der Arbeiten zeigte, dass die Werte zwischen 0,66 und dem maximal zu erreichenden Wert von 1 lagen. Im Durchschnitt ergab sich somit ein Wert von 0,925 ($p < 0,05$) für die Größenordnung der Übereinstimmung. Dies entsprach einer sehr guten Übereinstimmung und nur minimalen Abweichungen. Die nachfolgende Tabelle zeigt ICC-Werte für alle sechs verschiedenen Prüfungen und die jeweils zu vergleichende Konstellation. Der erste Wert beschreibt den Koeffizienten. Der Wert in der zweiten Zeile in der Klammer gibt das 95% - Konfidenzintervall an.

Konstellation	Komposit Präp MOD	Komposit Füllung MOD	Präp Gold TK
Ausbilder 1 versus Ausbilder 2	0,971 (0,907 - 0,991)	0,894 (0,628 - 0,969)	0,973 (0,921 - 0,991)
Ausbilder 1 versus Studierender	0,973 (0,913 - 0,992)	0,872 (0,569 - 0,962)	0,972 (0,918 - 0,990)
Ausbilder 2 versus Studierender	0,972 (0,912 - 0,991)	0,866 (0,615 - 0,957)	0,972 (0,921 - 0,991)

Konstellation	Präp Keramik TK	Trepanation	Wurzelfüllung
Ausbilder 1 versus Ausbilder 2	0,972 (0,920 - 0,991)	0,66 (0,188 - 0,887)	0,968 (0,895 - 0,991)
Ausbilder 1 versus Studierender	1 (1 - 1)	0,771 (0,4 - 0,927)	0,967 (0,891 - 0,990)
Ausbilder 2 versus Studierender	0,972 (0,920 - 0,991)	0,904 (0,714 - 0,971)	0,969 (0,897 - 0,991)

Tabelle 13: Überprüfung der Verständlichkeit. 1. Spalte: Angabe der Konstellation. 2.-7. Spalte: Prüfungsart. In den drei Zeilen werden die ICC-Werte mit den Konvidenzintervallen für den jeweiligen Vergleich angegeben.

Im Anschluss an die Bestimmung des Intraklassenkoeffizienten wurden die gemittelten Abweichungen zwischen den drei Bewertern pro Prüfungsart gebildet und mittels Microsoft Excel graphisch mit einem Balkendiagramm veranschaulicht.

Die gemittelte Abweichung der Bewertung der drei Bewerter (1-2, 1-3, 2-3) der jeweiligen Arbeiten lag zwischen 0,13 im besten Fall und im schlechtesten Fall bei 0,25 Notenpunkten. Das nachfolgende Diagramm beschreibt graphisch die Abweichungen, die für die sechs Prüfungen ermittelt wurden.

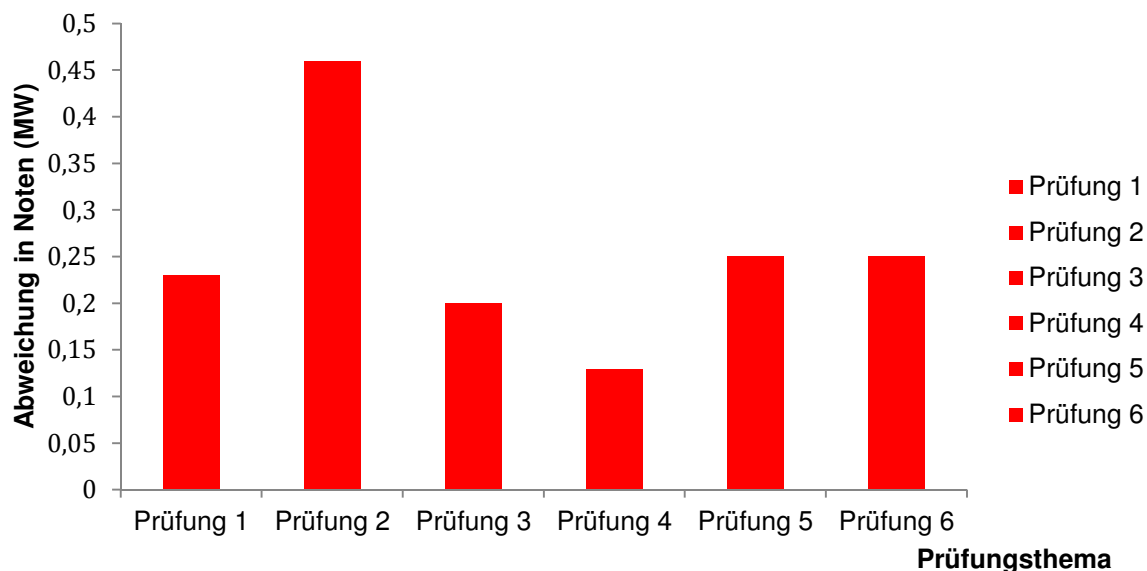


Abbildung 1: Graphische Darstellung der gemittelten Abweichung in Noten für 12 Arbeiten pro Prüfungsart der Überprüfung der Verständlichkeit pro Prüfung. MW = Mittelwert

Auf Grundlage der Ergebnisse dieser Analyse konnte davon ausgegangen werden, dass sowohl für Ausbilder als auch für Studierende von einer ausreichenden Verständlichkeit der Kriterien ausgegangen werden konnte. Es erfolgte ein Gespräch mit den beiden Ausbildern und dem Studierenden, ob von Ihrer Seite Verständnisprobleme bestanden. Dies war nicht der Fall, daher war eine Anpassung der Formulierung von Kriterien nicht notwendig.

4.2 Reliabilität der Bewertungen per Augenschein und per Kriterien durch die Ausbilder

Für diesen Auswertungsansatz standen die gleichen, wie in 4.1 beschriebenen, studentischen Arbeiten zur Verfügung. Aus den 42 Arbeiten pro Themenkomplex wurden nun 20 Arbeiten durch die Methode des „Random Sampling“ mit SPSS, Version 21 zufällig ausgewählt und anonymisiert.

Diese 20 Arbeiten pro Themenkomplex wurden wie in Material und Methode beschrieben von zwei Ausbildern und zwei Studierenden durch die zwei verschiedenen Beurteilungsmethoden bewertet.

Der erste Auswertungsansatz bestand darin den Intraklassen Korrelationskoeffizienten (ICC) pro Bewerter zwischen den beiden Bewertungsdurchgängen einer Methode zu bestimmen. Der ICC wurde mit absoluter Übereinstimmung, 2fach zufällig mit einem Konfidenzintervall von 95 % berechnet. Somit ergaben sich pro Prüfung für jeden der Bewerter je ein Wert von -1 bis 1 mit einem Konfidenzintervall für die Beurteilungsmethoden. Ab einem Wert von 0,8 für den ICC konnte man von einer guten Übereinstimmung ausgehen. Zur Beurteilung der Signifikanz wurde das Konfidenzintervall herangezogen. Ausgehen konnte man von einer Signifikanz, wenn die beiden zu vergleichenden Intervalle sich nicht überlappten. In den nachfolgenden Tabellen kann man die ICC - Werte mit den dazugehörigen Konfidenzintervallen pro Prüfungsart, Bewerter und Beurteilungsmethode sehen.

Es wurde sehr deutlich, dass die Dimension der Übereinstimmung der Bewertung der Prüfungen 1 – 4 für beide Ausbilder im Rahmen der Bewertung per Augenschein schlecht ist. Die Werte hatten eine Spanne für Ausbilder 1 von 0,22 bis 0,506, für

Ausbilder 2 von -0,02 bis 0,496. Für die Prüfung 5 (Trepanation) und Prüfung 6 (Wurzelfüllung) wurden von beiden Ausbildern mäßig bis gute Ergebnisse durch diese Methode erzielt (0,499- 0,919)

Die Beurteilung der Ergebnisse der Bewertung durch die vordefinierten Kriterien zeigt, dass bei allen sechs Prüfungen, bei beiden Ausbildern eine Steigerung des Intraklassenkorrelationskoeffizienten zu verzeichnen war. Alle Werte lagen in einem Bereich, der bedeutete, dass eine hohe Übereinstimmung vorlag (0,808- 1). Auch die bereits guten Werte der Prüfungen fünf und sechs aus der Bewertung per Augenschein konnten durch diese Beurteilungsmethode nochmals erhöht werden. Der Vergleich der Konfidenzintervalle zeigte, dass für alle sechs Prüfungen von einem signifikanten Ergebnis ausgegangen werden konnte, da bis auf in Prüfung sechs bei Ausbilder 2 keine Überschneidungen der Intervalle zu verzeichnen waren. Die Ergebnisse zeigen somit signifikant, dass im Gegensatz zur Beurteilung per Augenschein die Bewertung durch vordefinierte Kriterien ein hohes Maß an inter- und intrapersoneller Übereinstimmung hervorbrachte.

	Komposit Präp MOD	Komposit Füllung MOD	Präp Gold TK	Präp Keramik TK	Trepanation	Wurzelfüllung
Ausb. 1 Augenschein	0,323 (-0,88-0,654)	0,417 (-0,026-0,721)	0,506 (0,109-0,768)	0,22 (-0,170-0,576)	0,674 (0,351-0,856)	0,919 (0,809-0,967)
Ausb. 1 Kriterien	0,982 (0,955-0,993)	0,943 (0,863-0,977)	1 (1,000-1000)	0,974 (0,935-0,989)	0,905 (0,779-0,961)	0,991 (0,978-0,996)
Ausb 2 Augenschein	0,496 (0,098-0,762)	0,369 (-0,111-0,724)	-0,02 (-0,374-0,382)	0,044 (-0,393-0,466)	0,499 (0,107-0,762)	0,898 (0,666-0,964)
Ausb. 2 Kriterien	0,985 (0,962-0,994)	0,922 (0,814-0,969)	0,991 (0,977-0,996)	0,865 (0,695-0,944)	0,808 (0,582-0,919)	0,967 (0,92 -0,987)

Tabelle 14: Bestimmung der Intraklassenkorrelationskoeffizienten zum Vergleich der Reliabilität der Bewertung zwischen Augenschein und Kriterien der sechs Prüfungen durch die Ausbilder. 1 Spalte: Bewertungsart. 2.-7. Spalte Prüfungsart. In den vier Zeilen sind die ICC Werte mit den Konfidenzintervallen zu sehen.

Der zweite Auswertungsansatz bestand darin, die gemittelte Abweichung der Beurteilungsdurchgänge pro Ausbilder pro Prüfung (20 Arbeiten pro Prüfung) zu berechnen und graphisch darzustellen. Die nachfolgend aufgeführte Grafik zeigt sehr

deutlich, dass die Abweichung der Benotungen bei beiden Ausbildern für alle Prüfungen deutlich höher ist, wenn die Bewertung per Augenschein durchgeführt wird im Vergleich zur Kriterienbewertung. Für die Bewertung per Augenschein lag die Abweichung über alle sechs Prüfungen hinweg zwischen 0,3 und 1,1 Notenpunkten, im Gegensatz dazu für die Bewertung durch Kriterien zwischen 0,0 und 0,2 Notenpunkten.

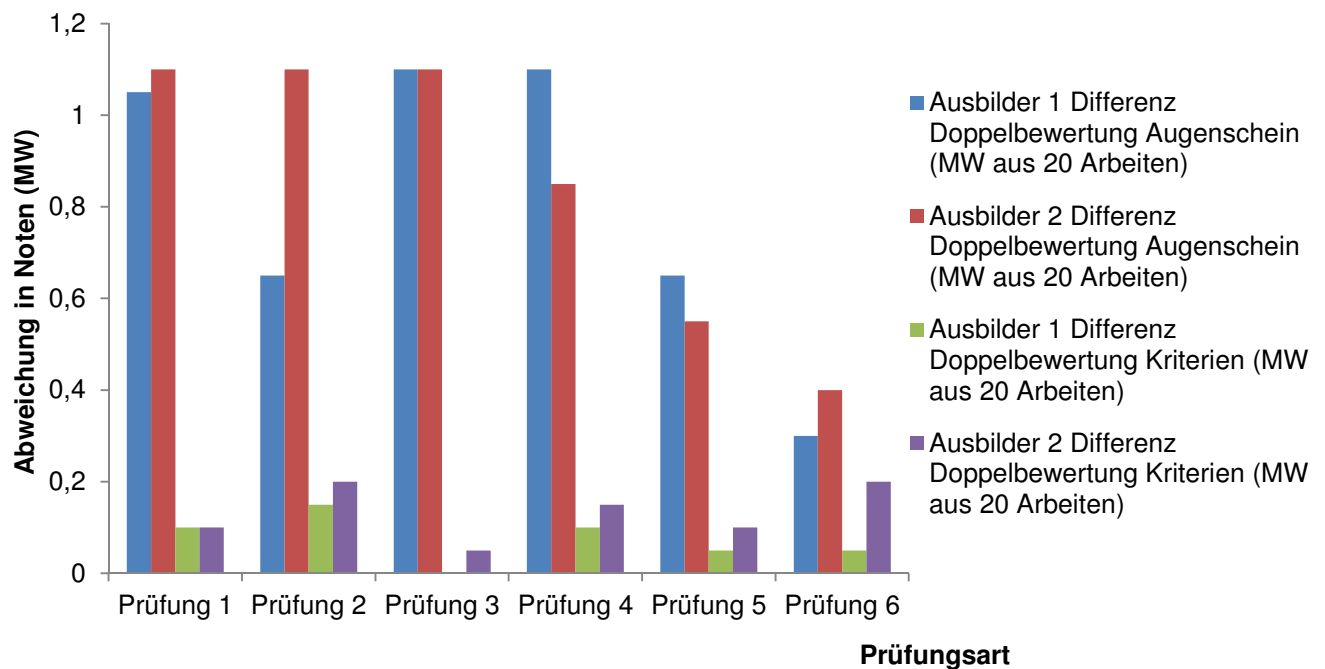


Abbildung 2: Graphische Darstellung der gemittelte Abweichung der Bewertung der praktischen Arbeiten durch beide Methoden pro Prüfung und Bewerter. MW = Mittelwert

4.3 Reliabilität der Bewertungen per Augenschein und per Kriterien durch die Studierenden

Für den ersten Auswertungsansatz wurde in diesem Fall verfahren wie in 4.2 beschrieben. In diesem Fall wurde die Übereinstimmung der Bewertung der beiden Studierenden für die beiden Beurteilungssysteme ermittelt. Die Bestimmung des Intraklassenkorrelationskoeffizienten (ICC) zeigte ein ähnliches Ergebnis wie bei den Ausbildern in 4.2. Auch in diesem Fall lag eine schlechte bis mäßige Übereinstimmung für beide Studierenden in Bezug auf die Bewertung per Augenschein für alle sechs Prüfungen vor. Die ICC-Werte lagen zwischen -0,214 und 0,691. Die Werte für die Berechnung des Ausmaßes der Übereinstimmung für die

Bewertung durch Kriterien lagen im Gegensatz dazu zwischen 0,808 und 1. Diese Werte drückten eine hohe Übereinstimmung aus. Der Vergleich der Konfidenzintervall machte deutlich, dass es bis auf die Trepanation bei Student 2 keine Überschneidungen gab und somit auch in diesem Fall von einer Signifikanz ausgegangen werden konnte.

	Komposit Präp MOD	Komposit Füllung MOD	Präp Gold TK	Präp. Keramik TK	Trepanation	Wurzelfüllung
Stud. 1 Augenschein	0,576 (0,189-0,808)	0,387 (-0,098-0,726)	0,403 (-0,041-0,713)	0,422 (-0,013-0,722)	0,163 (-0,261-0,547)	0,691 (-0,024-0,902)
Stud. 1 Kriterien	0,974 (0,935-0,989)	0,937 (0,842-0,975)	1 (1,000-1,000)	0,958 (0,897-0,983)	0,905 (0,779-0,961)	0,991 (0,977-0,996)
Stud. 2 Augenschein	-0,06 (-0,471-0,440)	0,367 (-0,093-0,693)	0,071 (-0,383-0,493)	-0,214 (-0,551-0,217)	0,314 (-0,147-0,660)	0,668 (0,332-0,854)
Stud. 2 Kriterien	0,967 (0,918-0,987)	0,939 (0,854-0,975)	0,971 (0,924-0,988)	0,958 (0,898-0,983)	0,808 (0,582-0,919)	0,983 (0,957-0,993)

Tabelle 15: Bestimmung der Intraklassenkorrelationskoeffizienten zum Vergleich der Reliabilität der Bewertung zwischen Augenschein und Kriterien der sechs Prüfungen durch die Studierenden. 1 Spalte: Bewertungsart. 2.-7. Spalte Prüfungsart. In den vier Zeilen sind die ICC Werte mit den Konfidenzintervallen zu sehen.

Der zweite Auswertungsansatz bestand auch in diesem Fall darin, die Standardabweichungen der Bewertungsdurchgänge pro Studierenden pro Prüfung zu errechnen und graphisch darzustellen. Die unten aufgeführte Grafik zeigte ein vergleichbares Ergebnis wie bei den Ausbildern unter 4.2. Die gemittelten Notenabweichung bei den Doppelbewertungen war bei beiden Studierenden für alle Prüfungen deutlich geringer, wenn die Beurteilung durch Kriterien erfolgte im Vergleich zur Bewertung per Augenschein. In diesem Fall lag die Abweichung bei der Bewertung per Augenschein zwischen 0,45 und 1,35 Notenpunkten. Im Gegensatz dazu lag sie bei der Bewertung durch Kriterien lediglich zwischen 0 und 0,2 Notenpunkten.

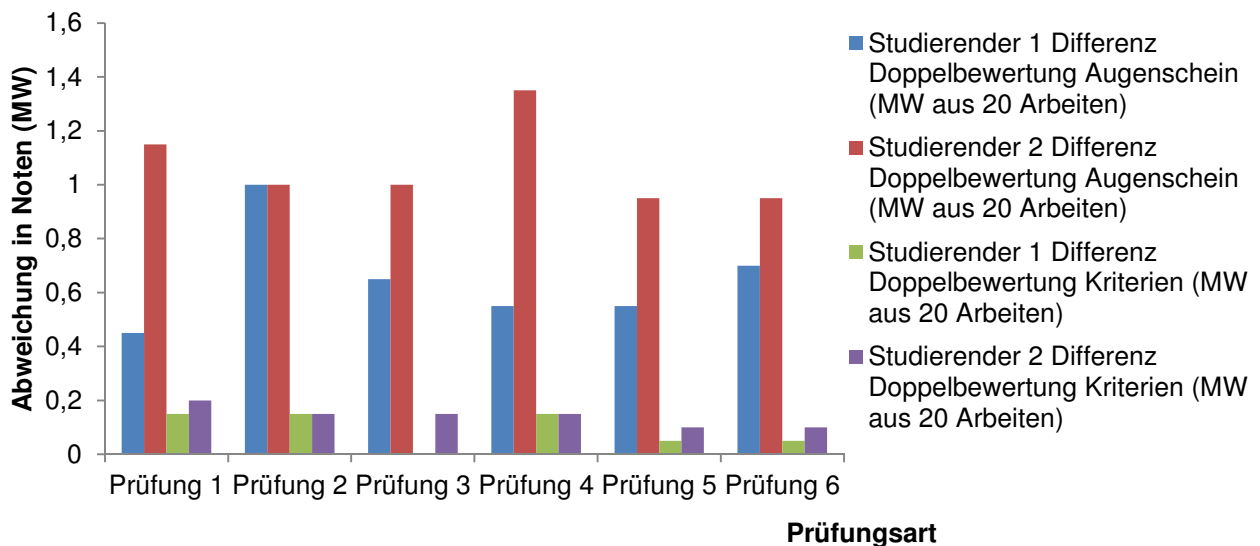


Abbildung 3: Graphische Darstellung der gemittelte Abweichung der Bewertung der praktischen Arbeiten durch beide Methoden pro Prüfung und Studierender. MW = Mittelwert

4.4 Gegenüberstellung der Abweichungen von Ausbildern und Studierenden

Der erste Auswertungsansatz bestand darin, anhand des Intraklassen Korrelationskoeffizienten (ICC) zu untersuchen, ob die Abweichung der Bewertungen zwischen Ausbildern und Studierenden unter Verwendung der Bewertungskriterien signifikant geringer ist als mithilfe der Bewertung per Augenschein.

Zu diesem Zweck wurden für jede Prüfung zuerst die intraexamen ICC-Werte für die Augenscheinbewertung und anschließend für die Kriterienbewertung gegenübergestellt. Bei dieser Gegenüberstellung wurde deutlich, dass die Abweichungen geringer waren zwischen Ausbildern und Studierenden unter Verwendung der Kriterien im Vergleich zur Augenscheinbewertung. Exemplarisch wurde dies anhand der ICC-Werte für die Präparation der Goldteilkrone sehr deutlich. Die Spannweite der Werte für die Bewertung per Augenschein lag zwischen -0,02 und 0,506. Im Gegensatz dazu lag die Spannweite für die Kriterienbewertung zwischen 0,971 und 1. Die übrigen fünf Prüfungen zeigten eine ähnliche Tendenz.

	Komposit Präp MOD	Komposit Füllung MOD	Präp Gold TK	Präp Keramik TK	Trepanation	Wurzelfüllung
Ausb. 1 Augen- schein	0,323 (-0,88-0,654)	0,417 (-0,026-0,721)	0,506 (0,109-0,768)	0,22 (-0,170-0,576)	0,674 (0,351-0,856)	0,919 (0,809-0,967)
Ausb. 2 Augen- schein	0,496 (0,098-0,762)	0,369 (-0,111-0,724)	-0,02 (-0,374-0,382)	0,044 (-0,393-0,466)	0,499 (0,107-0,762)	0,898 (0,666-0,964)
Stud. 1 Augen- schein	0,576 (0,189-0,808)	0,387 (-0,098-0,726)	0,403 (-0,041-0,713)	0,422 (-0,013-0,722)	0,163 (-0,261-0,547)	0,691 (-0,024-0,902)
Stud. 2 Augen- schein	-0,06 (-0,471-0,440)	0,367 (-0,093-0,693)	0,071 (-0,383-0,493)	-0,214 (-0,551-0,217)	0,314 (-0,147-0,660)	0,668 (0,332-0,854)

Tabelle 16: Bestimmung der Intraklassenkorrelationskoeffizienten zur Gegenüberstellung der Reliabilität der Bewertung per Augenschein der sechs Prüfungen durch die Ausbilder und Studierenden. 1. Spalte: Bewerter. 2.-7. Spalte Prüfungsart. In den vier Zeilen sind die ICC Werte mit den Konvidenzintervallen zu sehen.

	Komposit Präp MOD	Komposit Füllung MOD	Präp Gold TK	Präp Keramik TK	Trepanation	Wurzelfüllung
Ausb. 1 Kriterien	0,982 (0,955-0,993)	0,943 (0,863-0,977)	1 (1,000-1000)	0,974 (0,935-0,989)	0,905 (0,779-0,961)	0,991 (0,978-0,996)
Ausb. 2 Kriterien	0,985 (0,962-0,994)	0,922 (0,814-0,969)	0,991 (0,977-0,996)	0,865 (0,695-0,944)	0,808 (0,582-0,919)	0,967 (0,921-0,987)
Stud. 1 Kriterien	0,974 (0,935-0,989)	0,937 (0,842-0,975)	1 (1,000-1,000)	0,958 (0,897-0,983)	0,905 (0,779-0,961)	0,991 (0,977-0,996)
Stud. 2 Kriterien	0,967 (0,918-0,987)	0,939 (0,854-0,975)	0,971 (0,924-0,988)	0,958 (0,898-0,983)	0,808 (0,582-0,919)	0,983 (0,957-0,993)

Tabelle 17: Bestimmung der Intraklassenkorrelationskoeffizienten zur Gegenüberstellung der Reliabilität der Bewertung per Kriterien der sechs Prüfungen durch die Ausbilder und Studierenden. 1. Spalte: Bewerter. 2.-7. Spalte Prüfungsart. In den vier Zeilen sind die ICC Werte mit den Konvidenzintervallen zu sehen.

Im Rahmen des zweiten Auswertungsansatzes sollte mithilfe der graphischen Darstellung der gemittelten Notenabweichungen aus 20 Arbeiten pro Prüfungsart gezeigt werden, dass unter Verwendung der Bewertungskriterien die Abweichung der Beurteilungen deutlich geringer war als unter Verwendung des Systems per Augenschein. Dies konnte sehr eindrucksvoll dargestellt werden. Während die Abweichungen bei der Bewertung per Augenschein zwischen 0,3 und 1,35 lagen und einen durchschnittlichen Wert von 0,84 Notenpunkten über alle sechs Prüfungsarten hinweg ergaben, lag diese bei der Bewertung durch Kriterien zwischen 0 und 0,2 mit einem durchschnittlichen Wert von 0,11 Notenpunkten. Die nachfolgende Grafik illustriert die zuvor beschriebenen Daten.

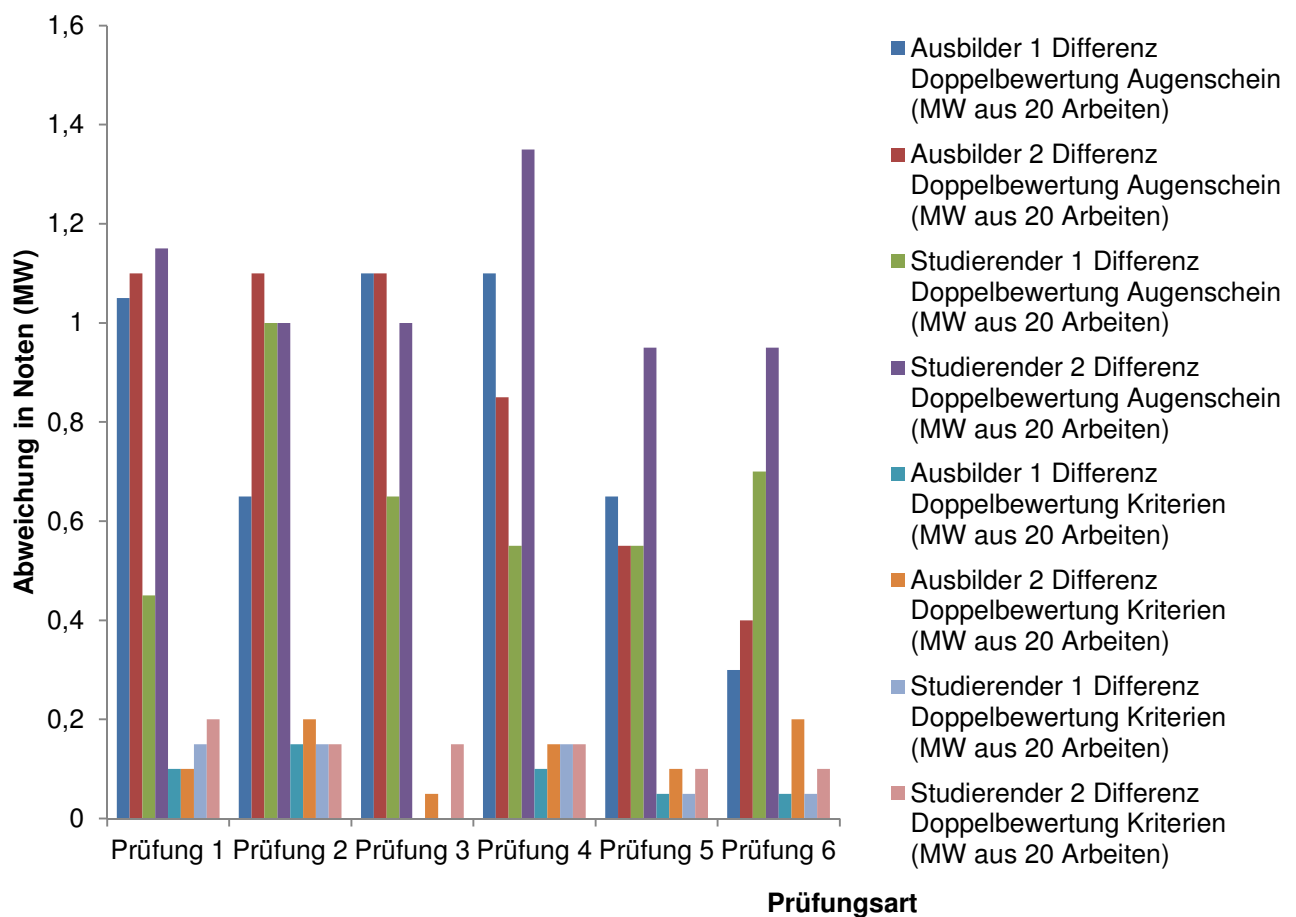


Abbildung 4: Graphische Darstellung der mittelwertigen Notenabweichungen aller Bewerter pro Prüfung und Beurteilungssystem für 20 Arbeiten pro Prüfungsart. MW = Mittelwert

4.5 Fragebogenauswertung

Die nachfolgenden Diagramme beschreiben die Aussagen der 42 an der Befragung teilnehmenden Studierenden und sechs Ausbilder in Bezug auf die gewählten Aussagen. Hierbei werden die absoluten Häufigkeiten dargestellt.

4.5.1 Einfluss der Kriterien auf das praktische Arbeiten

Sowohl der Fragebogen der Studierenden als auch der der Ausbilder gliederte sich in mehrere Fragenkomplexe. Der erste Fragenkomplex beinhaltete bei den Studierenden und Ausbildern die Fragen zum Thema: „Einfluss der Kriterien auf das praktische Arbeiten“.

Das erste Diagramm zeigt die Verteilung zur Aussage, dass die Bewertungskriterien korrekt seien. Die Mehrzahl der Studierenden ist der Meinung, dass dies auf die Kriterien zumindest eher zutrifft. Für die Aussage, dass dies gar nicht zutrifft, hatte sich ein Studierender entschieden.

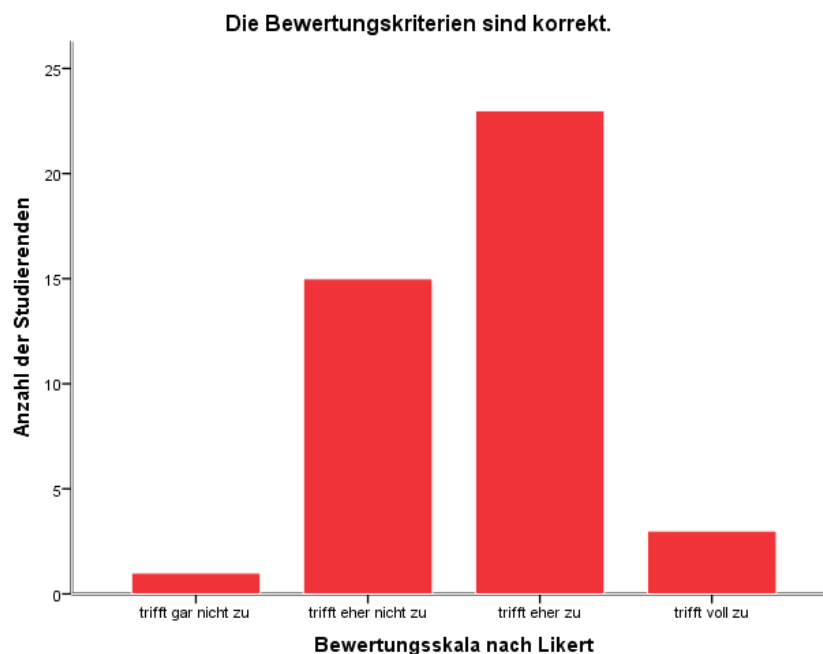


Abbildung 5: Graphische Darstellung der Aussage, dass die Bewertungskriterien korrekt seien aus Sicht der Studierenden.

Das folgende Diagramm zeigt die Aussagenverteilung der Ausbilder. Man sieht eine vergleichbare Verteilung. Die Mehrzahl entschied sich für die Aussage, dass dies eher zutrefte.



Abbildung 6: Graphische Darstellung der Aussage, dass die Bewertungskriterien korrekt seien aus Sicht der Ausbilder

Die nächsten Diagramme zeigen die gewählten Antworten auf die Aussage, dass die Bewertungskriterien vollständig seien. Es ist zu erkennen, dass 16 der Studierenden sich für die Antwort, dass dies eher nicht zutreffe und 20, dass dies eher zutreffe, entschieden hatten. Die verbleibenden 6 Studierenden hatten sich für die beiden übrigen Antwortmöglichkeiten entschieden.

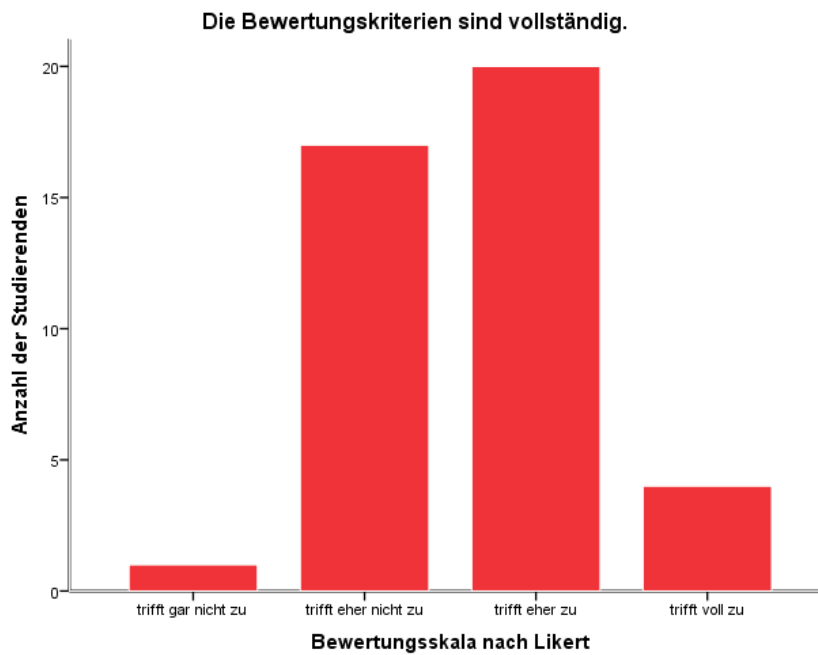


Abbildung 7: Graphische Darstellung der Aussage, dass die Bewertungskriterien vollständig seien aus Sicht der Studierenden

Die Ausbilder hatten die gleiche Frage gestellt bekommen. Nachfolgend sind die Ergebnisse zu sehen auf die Frage, ob die Bewertungskriterien vollständig seien. Die Ausbilder zeigten eine ähnliche Tendenz wie die Studierenden, jedoch deutlicher, dass dies sogar voll zutreffe.

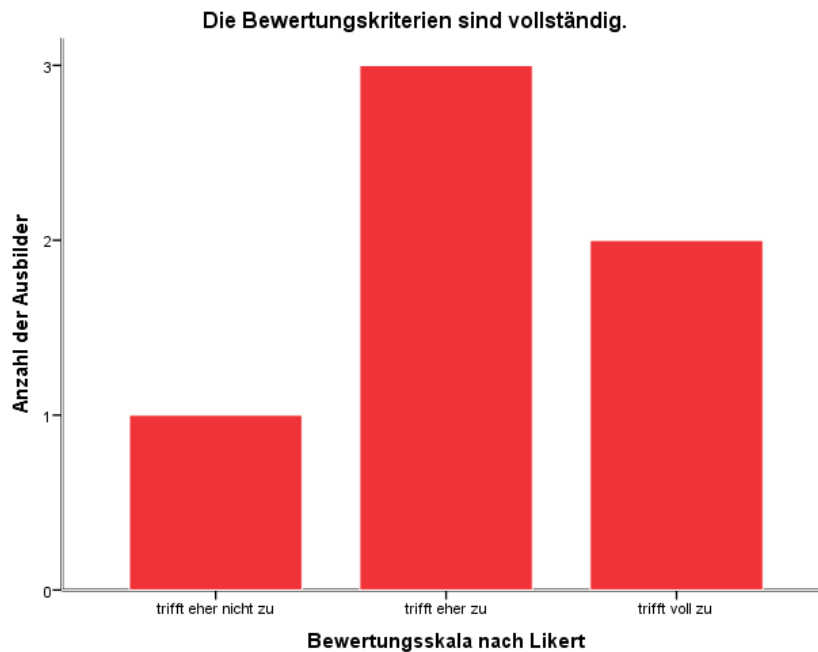


Abbildung 8: Graphische Darstellung der Aussage, dass die Bewertungskriterien vollständig seien aus Sicht der Ausbilder.

Das folgende Balkendiagramm illustriert die gewählten Antworten auf die Aussage, dass die Bewertungskriterien zu umfangreich seien. Die Mehrzahl der Studierenden entschied sich für Aussage, dass dies eher nicht zutreffe. Jedoch entschieden sich auch einige der Studierenden für die Aussage, dass dies eher zutreffe

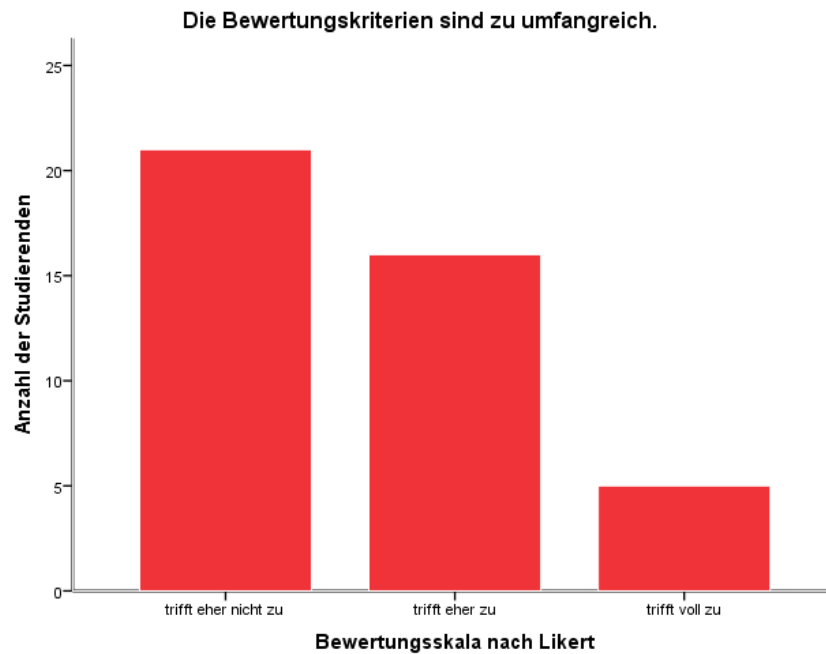


Abbildung 9: Graphische Darstellung der Aussage, dass die Bewertungskriterien zu umfangreich seien aus Sicht der Studierenden.

Dahingegen zeigt das Ergebnis der gewählten Aussagen der Ausbilder, dass diese der Meinung waren, dass die Bewertungskriterien nicht zu umfangreich sein. Dies ist deutlich in der unten aufgeführten Abbildung zu erkennen. Alle Ausbilder verteilten sich auf die beiden Antworten, dass es gar nicht zutreffe und eher nicht zutreffe.

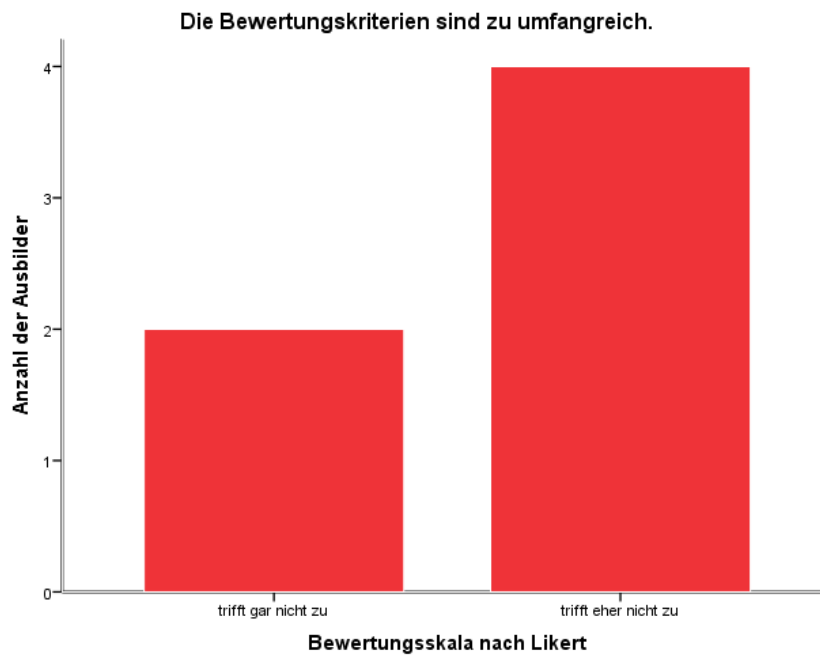


Abbildung 10: Graphische Darstellung der Aussage, dass die Bewertungskriterien zu umfangreich seien aus Sicht der Ausbilder.

Die nachfolgende Grafik beschäftigt sich mit der Aussage, ob die Bewertungskriterien für die Studierenden klar sind. Der Hauptanteil der Studierenden gab an, dass die Kriterien eher und voll zu treffen. Die Minderheit war der Meinung, dass dies nicht zutreffe.



Abbildung 11: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien klar für den Studierenden seien.

Die Ausbilder sollten sich auch zu dieser Aussage äußern. Das nächste Diagramm beschreibt die Ergebnisse der Antworten der Ausbilder auf die Aussage, dass die Bewertungskriterien für den Ausbilder klar wären. Zwei Drittel der Ausbilder waren der Meinung, dass dies voll zutreffe.



Abbildung 12: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien klar für den Ausbilder seien.

Das nächste Diagramm zeigt die Ergebnisse auf die Frage ob die Anwendung der Kriterien es dem Studierenden erleichtert, das praktische Arbeiten zu strukturieren. Der überwiegende Anteil der Studierenden gab an, dass dies eher auf sie zutrefe. Kein studierender gibt an, dass dies gar nicht zutrefe.



Abbildung 13: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Studierenden erleichtere das praktische Arbeiten zu strukturieren.

Die Ausbilder wurden zu diesem Thema mit der Aussage konfrontiert, ob die Anwendung der Bewertungskriterien, es dem Ausbilder erleichtere das praktische Arbeiten zu strukturieren. Die eindeutige Mehrheit empfand, dass dies eher oder voll zutreffe. Kein Ausbilder war der Meinung, dass dies gar nicht zutreffe.

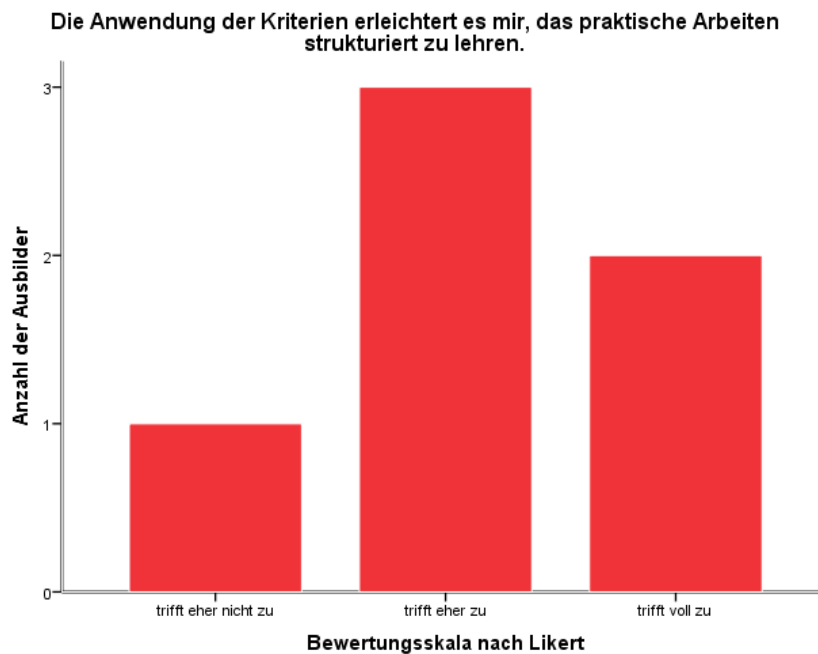


Abbildung 14: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Ausbilder erleichtere das praktische Arbeiten strukturiert zu lehren.

Das folgende Diagramm zeigt die Verteilung der gegebenen Antworten auf die Aussage, dass die Anwendung der Kriterien die Vorbereitung des Studierenden erleichtere. Die deutliche Mehrheit sah dies so und gab an, dass dies eher zutreffe oder voll zutreffe.



Abbildung 15: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien die Vorbereitung des Studierenden erleichtere.

Das unten gezeigte Diagramm zeigt die von den Studierenden getroffenen Antworten auf die Aussage, dass die Anwendung der Kriterien es ihnen erleichtere, die aufeinanderfolgenden Arbeitsschritte selbstständig durchzuführen. Knapp die Hälfte der Teilnehmer des Kurses war der Ansicht, dass dies eher nicht zutreffe.



Abbildung 16: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Studierenden erleichtere die Arbeitsschritte selbstständig durchzuführen.

Das nächste Diagramm zeigt die Verteilung der Antworten auf die Feststellung, dass die Anwendung der Kriterien es dem Studierenden erleichtere eigene Defizite zu erkennen und gezielt zu üben. Der Hauptteil der Studierenden empfand es so, dass dies zutreffe.

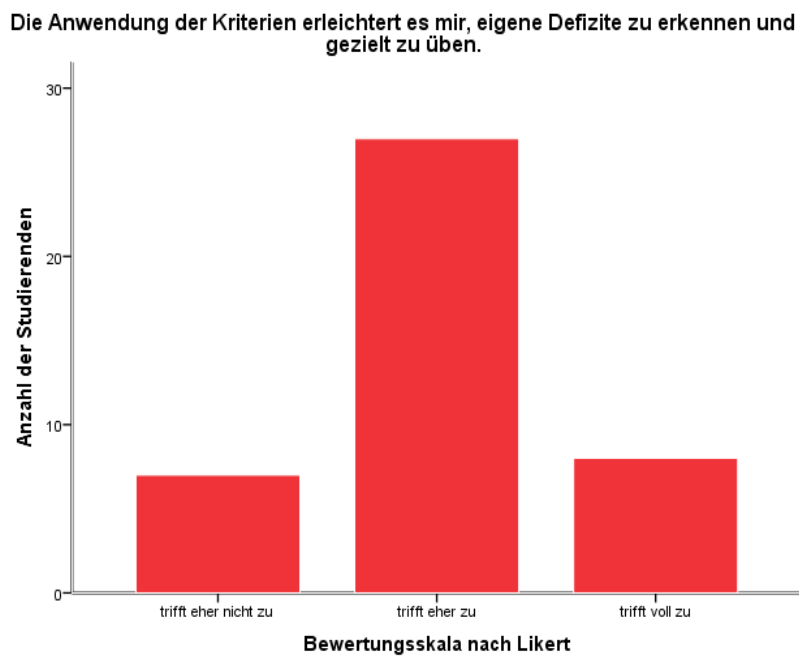


Abbildung 17: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Studierenden erleichtern eigene Defizite zu erkennen und zu üben.

Die Ausbilder sollten einschätzen, ob die Anwendung der Bewertungskriterien es dem Ausbilder erleichtere die Unterschiede der Präparationsformen für Gold, Keramik und Komposit aufzuzeigen. Die Mehrheit der Ausbilder war der Meinung, dass die Anwendung der Bewertungskriterien es erleichtere die Unterschiede der Präparationsformen aufzuzeigen.



Abbildung 18: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Ausbilder erleichtere, die Unterschiede der Präparationsformen Gold, Keramik und Komposit aufzuzeigen.

Die Ausbilder sollten Stellung beziehen, ob sie den Eindruck hätten, dass durch die Anwendung der Bewertungskriterien das Wesentliche der praktischen Arbeiten besser gelernt würde als ohne die Anwendung der Kriterien. Der Hauptanteil der Ausbilder empfindet, dass dies eher zutreffe oder voll zutreffe.



Abbildung 19: Graphische Darstellung der Aussage der Ausbilder, dass durch die Anwendung der Kriterien das Wesentliche der praktischen Arbeiten besser gelernt würde als ohne Kriterien.

Zum Abschluss des ersten Fragenkomplexes wurden die Ausbilder befragt, ob sie den Eindruck hätten, dass sie durch die Kenntnis der Kriterien die Lernziele des Kurses besser vermitteln könnten. Es wurde sehr deutlich, dass der Großteil der Ausbilder diese These unterstützte. Für die Antwortmöglichkeit, dass dies gar nicht zutreffe, entschied sich kein Ausbilder.

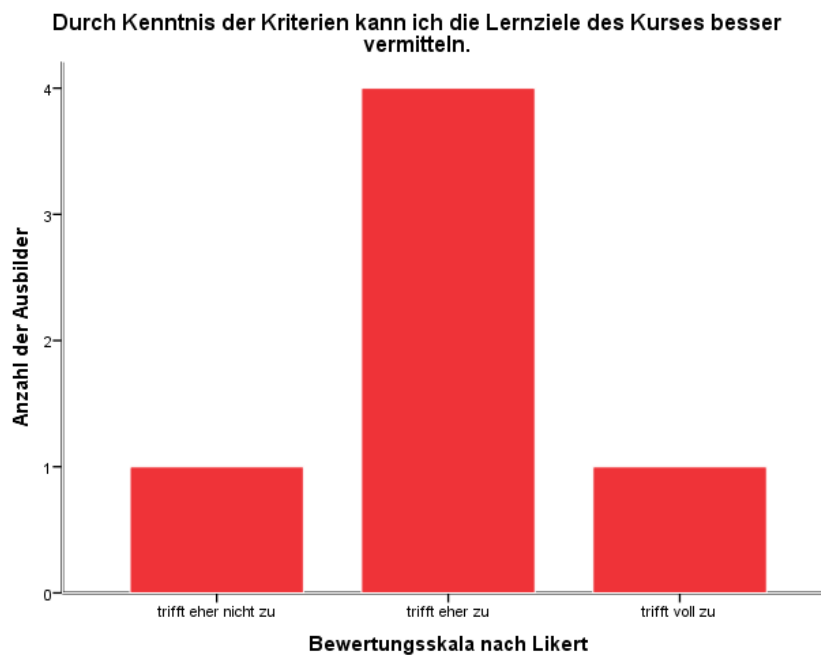


Abbildung 20: Graphische Darstellung der Aussage, dass der Ausbilder durch die Kenntnis der Kriterien die Lernziele des Kurses besser vermitteln könnte.

4.5.2 Einfluss der Kriterien auf die Güte der Bewertung

Der zweite Themenkomplex befasste sich mit der Einschätzung der Güte auf die Bewertung.

Das unten aufgeführte Diagramm zeigt die Ergebnisse auf die Aussage, dass die Anwendung der Kriterien es dem Studierenden erleichterte Stärken und Schwächen bei den einzelnen Arbeitsschritten isoliert zu erkennen. Die Eindeutige Mehrheit empfindet, dass dies eher zutreffe. Die klare Minderheit hat das Gefühl, dass dies gar nicht oder eher nicht zutreffe.



Abbildung 21: Graphische Darstellung der Aussage, dass die Bewertungskriterien es dem Studierenden erleichterten Stärken und Schwächen bei den einzelnen Arbeitsschritten isoliert zu erkennen.

Die nächsten Diagramme dieses Komplexes beschäftigten sich damit, ob der Studierende und der Ausbilder den Eindruck hätten, dass die Anwendung der Bewertungskriterien es erleichterte die Arbeiten im Laufe der jeweiligen Sequenz zu vergleichen. Das erste Diagramm zeigt die Ergebnisse der Studierenden. Die Mehrheit der Studierenden hatte den Eindruck, dass dies eher zutreffe. Dem entgegen stehen 15 Studierende, die der Meinung sind, dass dies eher nicht zutreffe.

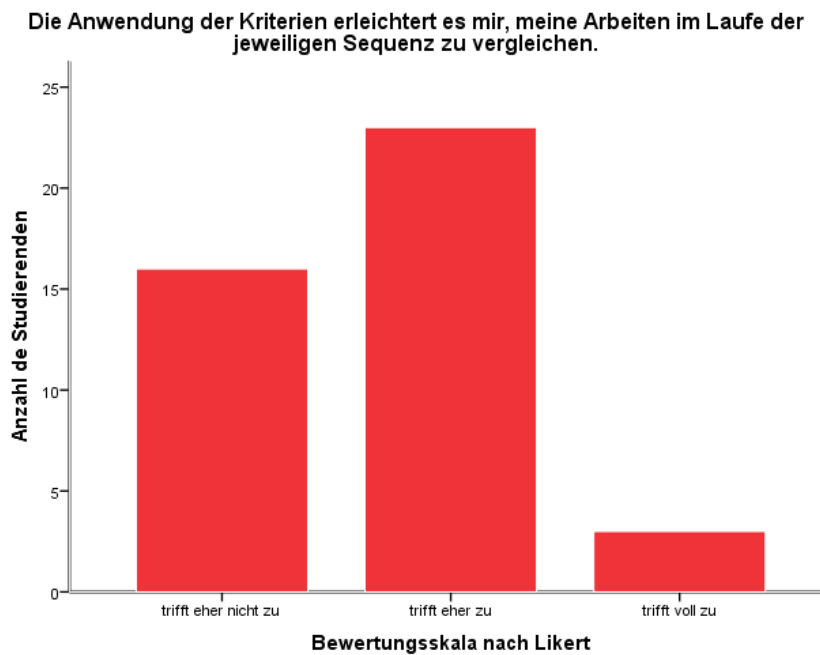


Abbildung 22: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es den Studierenden erleichtere die Arbeiten im Laufe der jeweiligen Sequenz zu vergleichen.

Dem entgegen stand die Aussage der Ausbilder. Im nächsten Diagramm kann man die Ergebnisse der Ausbilder sehen. Die Ausbilder waren hauptsächlich der Meinung, dass dies voll zutreffe oder eher zutreffe.



Abbildung 23: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Ausbilder erleichtere die Arbeiten im Laufe der jeweiligen Sequenz zu vergleichen.

Die letzte Frage des zweiten Themenkomplexes des Fragebogens der Ausbilder beschäftigt sich damit, ob die Anwendung der Kriterien die forensische Sicherheit der Beurteilung erhöhte. Alle Ausbilder entschieden sich dafür, dass dies zutreffe. Die Mehrheit entschied sich sogar dafür, dass dies voll zutreffe.



Abbildung 24: Graphische Darstellung der Aussage, dass die Anwendung der Kriterien die forensische Sicherheit der Beurteilung erhöhte.

4.5.3 Einfluss der Kriterien auf die Kommunikation

Der dritte Themenkomplex setzte sich mit dem Einfluss der vordefinierten Bewertungskriterien auf die Kommunikation auseinander.

Die Ausbilder sollten sich zu Beginn ihres zweiten Themenkomplexes zu der Aussage äußern, dass die Anwendung der Kriterien es ihnen erleichtere ein Feedback zu geben. Die Ausbilder waren alle der Meinung, dass die Rückmeldung an den Studierenden durch die Anwendung der Kriterien erleichtert würde. Die Antworten verteilten sich zu gleichen Anteilen auf die Aussagen, dass dies eher zutreffe und dass dies voll zutreffe. Kein Ausbilder empfand, dass es nicht zutreffe.



Abbildung 25: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Ausbilder erleichterte dem Studierenden Feedback zu geben.

Das folgende Diagramm zeigt, die Ergebnisse auf die erste Aussage im Fragebogen der Studierenden, dass die Anwendung der Bewertungskriterien es dem Studierenden erleichterte, die kritischen Rückmeldungen besser hinterfragen zu können. Der Großteil der Studierenden empfand, dass dies eher zutreffe. Kein Studierender gab als Antwort, dass dies gar nicht zutreffe.

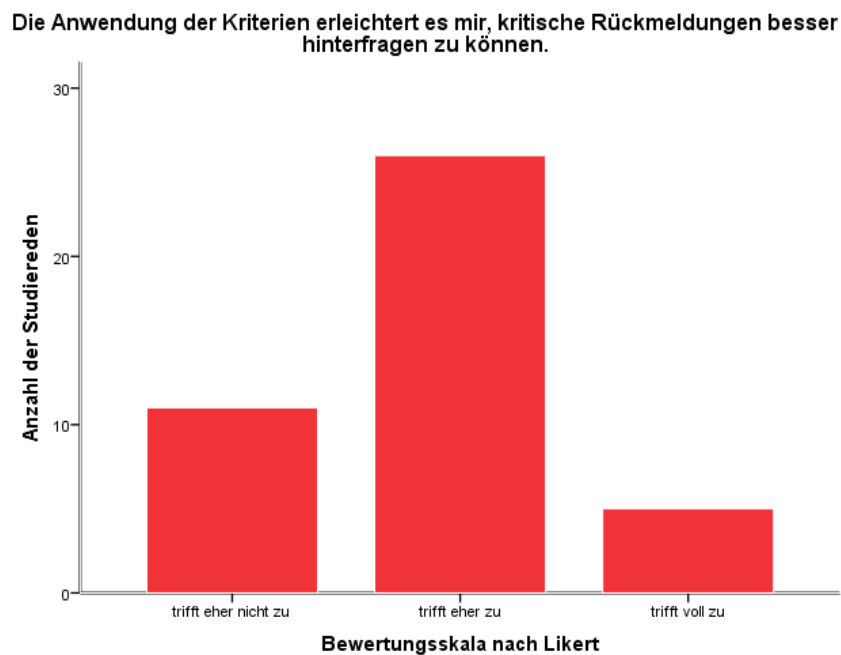


Abbildung 26: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Studierenden erleichtere kritische Rückmeldungen kritisch zu hinterfragen.

Im Folgenden sollten die Ausbilder Stellung beziehen zu der Aussage, dass die Anwendung der Bewertungskriterien, es ihnen erleichterte ihre Bewertung gegen Kritik durch den Studierenden zu verteidigen. Auch in diesem Fall sagten die Ausbilder, dass dies zu treffe und verteilten sich auf die beiden Aussagen, dass dies voll oder eher zutrefte.

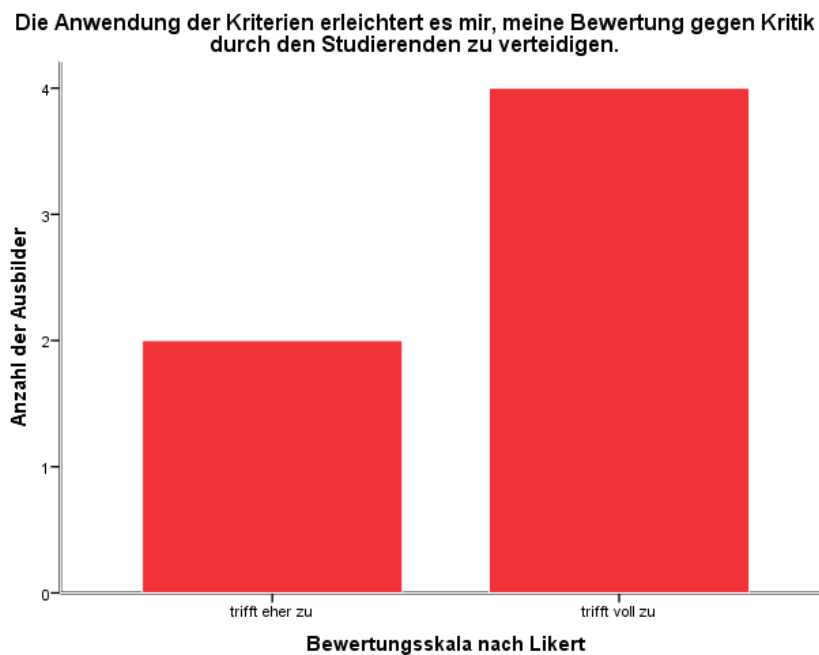


Abbildung 27: Graphische Darstellung der Aussage, dass die Anwendung der Kriterien es dem Ausbilder erleichterte die Bewertung gegen Kritik durch den Studierenden zu verteidigen.

Die letzte Aussage des dritten Themenkomplexes befasste sich mit der Annahme, dass die Anwendung der Bewertungskriterien es dem Ausbilder erleichterte seine Bewertung gegenüber seinem Vorgesetzten zu verteidigen. Auch in diesem Fall wiederholte sich die Verteilung der gegebenen Antworten. Die Ausbilder waren einstimmig der Meinung, dass die Anwendung der Kriterien es erleichterte gegenüber dem Vorgesetzten die eigene Bewertung zu verteidigen. Die Mehrzahl war der Meinung, dass dies sogar voll zutraf.



Abbildung 28: Graphische Darstellung der Aussage, dass die Anwendung der Kriterien es dem Ausbilder erleichterte, seine Bewertung gegenüber dem Vorgesetzten zu verteidigen.

4.5.4 Einfluss der Bewertungskriterien auf den Zeitaufwand

Zum Abschluss der Befragung sollten die Ausbilder sich im vierten Themenkomplex zum Einfluss der vordefinierten Bewertungskriterien auf den Zeitaufwand äußern. Dieser vierte Komplex untergliederte sich in zwei Aussagen.

Die erste Aussage befasste sich damit, dass die Anwendung der Kriterien den Zeitaufwand für die Bewertung im Vergleich zur Bewertung per Augenschein erhöhen würde. Die graphische Darstellung der Ergebnisse zeigt, dass es kein einheitliches Meinungsbild gab. Auf der einen Seite empfanden einige der Zahnärzte, dass dies gar nicht zutreffe, auf der anderen Seite empfanden jedoch doppelt so viele Ausbilder, dass dies eher zutreffe. Dies kann mit der unterschiedlichen klinischen Erfahrung der Ausbilder zusammen hängen.



Abbildung 29: Graphische Darstellung, der Aussage, dass die Anwendung der Kriterien den Zeitaufwand für die Bewertung erhöhe im Vergleich zur Bewertung per Augenschein.

Die zweite Aussage dieses Komplexes, beinhaltet, dass die Anwendung der Kriterien das Erlernen der praktischen Fertigkeiten der Studierenden beschleunigte. Der Großteil der Ausbilder entschied sich für die Antwortmöglichkeit, dass dies eher zutreffe. Dies wird in der nachfolgenden Grafik klar beschrieben.



Abbildung 30: Graphische Darstellung der Aussage, dass die Anwendung der Kriterien das Erlernen der praktischen Fertigkeiten beschleunigen würde.

4.5.5 Vergleich der beiden Bewertungsverfahren

Im Letzten Abschnitt des Fragebogens sollte überprüft werden, ob aus Sicht der Ausbilder manche der praktischen Arbeiten mehr von der Verwendung der Kriterien profitieren als Andere. Das nachfolgende Diagramm zeigt graphisch die Verteilung der Einschätzung der Ausbilder welche der praktischen Arbeiten am meisten von der Verwendung der Bewertungskriterien profitierten. Man kann erkennen, dass das Legen der Kompositfüllung am meisten aus Sicht der Ausbilder von der Verwendung der Bewertungskriterien profitierte. Dem gegenüber steht, dass die Trepanation und Wurzelkanalfüllung am wenigsten profitierte. Zwischen diesen beiden Gruppen lag, die Präparation für die Klasse-II-Kavität für eine Kompositfüllung, die Präparation für die Goldteilkrone sowie für die Goldteilkrone.

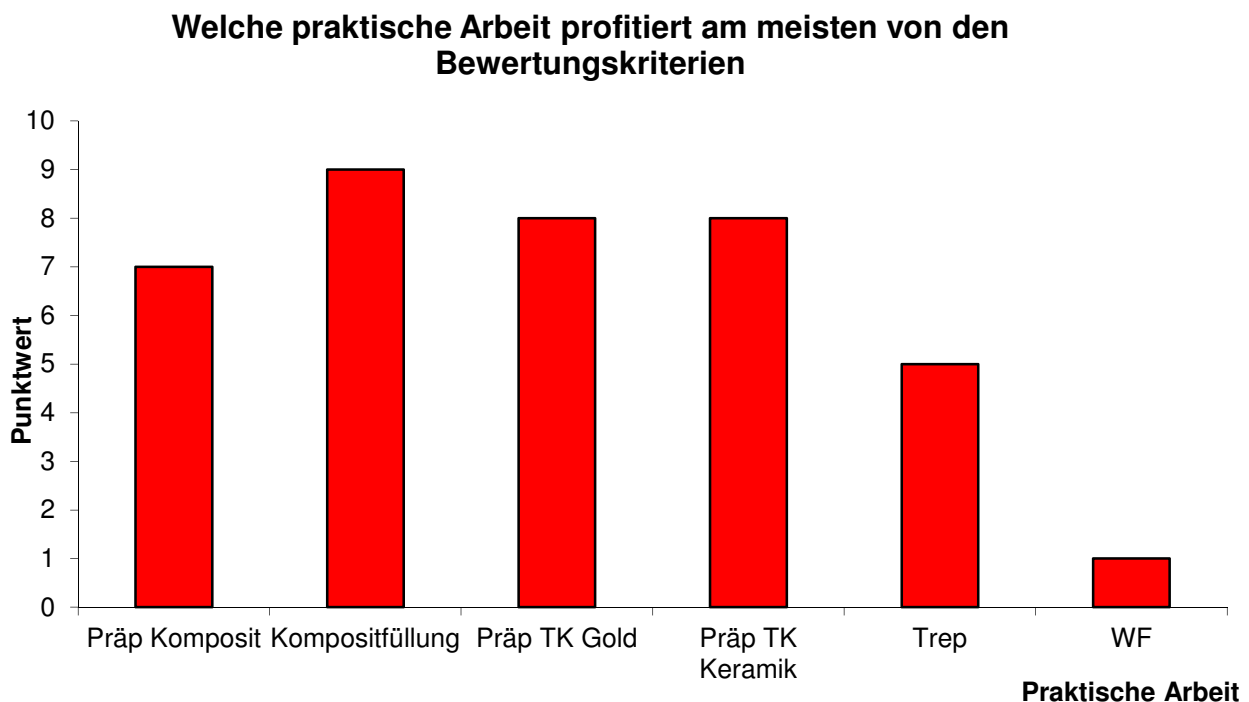


Abbildung 31: Graphische Darstellung der Frage: welche der Arbeiten profitiert am meisten von den Bewertungskriterien? Aus Sicht der Ausbilder.

4.5.6 Freitext

Am Schluss der Befragung gab es die Möglichkeit die positiven Aspekte und den Verbesserungsbedarf in einem Freitext aufzuführen.

Die Mehrheit der Studierenden und Ausbilder führten als positiven Aspekt auf, dass die Arbeiten objektiv und transparent bewertet werden. Die Studierenden empfanden es zudem als positiv, dass durch die Verwendung der Bewertungskriterien die Arbeitsschritte und Hintergründe der Behandlung klarer werden. Außerdem wurde es als positiv wahrgenommen, dass die Rückfragen klar und deutlich anhand der Bewertungskriterien erläutert werden.

Als Verbesserungsvorschläge wurde angemerkt, dass aus Sicht der Studierenden die klinische Relevanz bei manchen Kriterien mehr hervorgehoben werden sollte. Zudem wünschten sich manche Studierenden weitere Abstufungen und Gewichtungen für einzelne Kriterien, wie beispielsweise die Ausdehnung eines Randspaltes bei einer Kompositfüllung.

Zusammengefasst konnte man anhand des Fragebogens feststellen, dass die Verwendung der Bewertungskriterien sowohl von Studierenden als auch von Ausbildern positiv bewertet wurde. Der erste Themenkomplex setzte sich mit dem Einfluss der Kriterien auf das praktische Arbeiten auseinander. Ausbilder und Studierende vertraten die Meinung, dass das praktische Arbeiten davon profitierte. Alle zu beurteilenden Aussagen wurden positiv eingeschätzt. Im zweiten Themenkomplex sollte die Güte der Bewertung eingeschätzt werden. Das Ergebnis zeigte, dass sowohl die Ausbilder als auch die Studierenden eine Steigerung der Güte der Bewertung durch Anwendung der Bewertungskriterien empfanden. Unter anderem empfanden die Ausbilder, dass die forensische Sicherheit durch die Anwendung der Bewertungskriterien erhöht wurde. Auch in diesem Themenkomplex wurden alle Aussagen positiv zu Gunsten der Anwendung der Bewertungskriterien beantwortet. Der dritte Themenkomplex setzte sich mit dem Einfluss der Bewertungskriterien auf die Kommunikation auseinander. Die Studierenden und Ausbilder kamen anhand des Fragebogens zu dem Schluss, dass die Kommunikation untereinander erleichtert wurde. Für die Ausbilder erleichterte die Anwendung der Bewertungskriterien, dem Studierenden nach der Prüfung ein Feedback zu geben und die Beurteilung gegenüber diesem und auch einem

Vorgesetzten zu verteidigen. Im vierten Themenkomplex sollten die Ausbilder einschätzen, wie sich die Anwendung der Kriterien auf den Zeitaufwand auswirke. Die Ausbilder schätzten den Zeitaufwand höher ein durch die Verwendung der Bewertungskriterien im Vergleich zur Beurteilung per Augenschein. Jedoch verringerte sich der Zeitaufwand im Rahmen des Lernprozesses der jeweiligen praktischen Fertigkeiten und der Erläuterungen der Bewertung in der Prüfung. Im letzten Abschnitt des Fragebogens sollten die Ausbilder evaluieren, welche Prüfung am meisten von der Verwendung der Bewertungskriterien profitierte. Aus Sicht der Ausbilder war dies die Anfertigung der Kompositfüllung. Am wenigsten nutzte dahingegen die Anwendung der Kriterien im Rahmen der Wurzelfüllung.

5 Diskussion

Die zuverlässige Beurteilung von praktischen Arbeiten spielt in der Zahnmedizin eine große Rolle. Dabei sind sowohl die Reproduzierbarkeit als auch die Objektivität nicht nur für Ausbilder sondern auch für Studierende von großer Bedeutung (Gaines *et al.* 1974; Türp 2002; Kellersmann 2008).

5.1 Vorarbeiten

5.1.1 Nominal group technique

Vor der hier beschriebenen Studie wurde eine nominal group technique durchgeführt, um eine inhaltliche Gewichtung der Bewertungskriterien zu erlangen (Delp *et al.* 1977; Horton 1980). Diese Methode ist eine mathematische Zusammenführung einer Gruppenentscheidung. Die nominal group technique besteht aus mehreren aufeinanderfolgenden Arbeitsschritten. In der ersten Phase wurde in der Gruppe erörtert, welche Bewertungskriterien für die jeweiligen Prüfungen wichtig waren. Für diese Kriterien wurde eine Literaturrecherche angefertigt. Die Arbeitsgruppe bestand aus Fr. Prof. Dr. Huth, Dr. Jakob, Dr. Nemecek, ZA Baumann sowie Stud. med. dent. Hr. Kollmuss. Aus den Ergebnissen dieses ersten Arbeitsschrittes wurden Listen mit den Bewertungskriterien angefertigt. Diese Listen wurden den vier an der Studie teilnehmenden Zentren (Prof. Haak, Leipzig, Prof. Hahn, Freiburg, Prof. Huth, München, Prof. Szep, Frankfurt am Main) zugesandt, um noch fehlende Kriterien hinzufügen oder weniger wichtige zu streichen. Diese vier Experten sind an ihren Universitäten für den Phantomkurs der Zahnerhaltung verantwortlich. Aus diesen Ergebnissen entstanden die zu bearbeitenden Kriterien. Diese Kriterien sollten in der zweiten Phase mit jeweils 1, 2 oder 3 Punkten versehen werden. Diese Bewertung wurde von den vier Zentren durchgeführt. Die Ergebnisse wurden für jedes Kriterium pro Prüfung zusammengezählt. Somit wurde aus den Ergebnissen der vier Experten eine Rangliste erstellt, in der die Bewertungskriterien aufgelistet wurden. Im Anschluss wurden die Ergebnisse in der Gruppe diskutiert. In der Literatur werden verschiedene Vorgehensweisen für eine nominal group technique beschrieben. Die am meisten verwendete Methode besteht darin, dass in allen Phasen dieser Analyse immer alle Personen teilnehmen (Delp *et al.* 1977). Im Rahmen dieser Studie haben wir uns bewusst gegen dieses Verfahren entschieden und die NGT-Analyse

modifiziert. An den Diskussionsrunden zu Beginn und am Ende nahmen aus Praktikabilitätsgründen nur Mitarbeiter der eigenen Universität teil. Am Ende der Listen waren die Kriterien zu finden, die mit der geringsten Punktzahl versehen wurden. Dies hatte unterschiedliche Gründe. Zum einen wurde ihnen eine geringere klinische Relevanz als anderen Bewertungskriterien beigemessen. Zum anderen empfanden die Teilnehmer noch immer, dass es geringe inhaltliche Überschneidungen mit anderen Kriterien gab. Auf Grundlage dieser Ergebnisse wurden überprüft, ob der Grund wirklich inhaltliche Überschneidungen oder Ungenauigkeiten in den Formulierungen waren. Um eine praktikable Anzahl an Bewertungskriterien pro Prüfung zu erlangen, entschied man sich die Kriterien mit der geringsten Punktzahl wegfällen zu lassen. Auf diese Weise erhielten wir pro Themengebiet 10-15 Kriterien. Diese Menge an Bewertungskriterien ermöglichte eine praktikable Anwendung. Für die prozentuale Gewichtung wurde die Methode der Matrix-Analyse gewählt. Im Zuge dieses Arbeitsschrittes werden alle zur Verfügung stehenden Kriterien miteinander verglichen und gewichtet. Die Prozentwerte aller Kriterien ergeben zusammen 100%. Für den Fall, dass mehr als 10-15 Kriterien pro Prüfung vorliegen, hätte sich für jedes Kriterium ein zu geringer Prozentwert ergeben, als dass man ihn in der Bewertung hätte heranziehen können. Auch auf Grund dessen hatte man sich entschieden, die Anzahl der Kriterien zu reduzieren. Somit war es möglich im Rahmen der Matrix-Analyse jedem Kriterium einen Punktwert zuzuordnen, der in der späteren Beurteilung eine zutreffende Bewertung des Studierenden ermöglichte (Huth, in Bearbeitung).

Die Literaturrecherche zeigte deutlich, dass es einen großen Bedarf an zuverlässigen Beurteilungen von praktischen studentischen Arbeiten gibt. In keiner der Studien kann man jedoch etwas lesen wie die Kriterien oder Punkte einer Checkliste entstanden sind. Es werden lediglich die verschiedenen Beurteilungsmethoden miteinander verglichen. Türp und Mitarbeiter (Türp 2002) verglichen die Unterschiede der Bewertung prothetischer Arbeiten im Phantomkurs zwischen Assistenz Zahnärzten und Studierenden. Dabei wurde nicht auf die Entstehung der Bewertungsgrundlage eingegangen. Kellersmann (Kellersmann 2008) untersuchte die Reliabilität der Bewertung per Augenschein und anhand eines Bewertungsbogens. Auch in dieser Studie kann man nicht erkennen auf welche Art und Weise die verwendeten Kriterien in den Bögen entstanden waren. Bereits Mitte der sechziger Jahre befasste sich Lilley (Lilley *et al.* 1968) mit der inter- und

intrapersonellen Übereinstimmung von Bewertungen von studentischen Arbeiten. Zwei verschiedene Bewertungsmethoden wurden verglichen, zum einen per Augenschein, zum anderen durch Kriterien. Die Bewertungskriterien sind im Rahmen der Studie von Lilley (Lilley *et al.* 1968) durch die drei teilnehmenden Zahnärzte erörtert und schließlich verwendet worden. Diese Gruppendiskussion kann man nicht als Verfahren interpretieren um allgemeingültige Kriterien zu entwickeln. Fuller (Fuller 1972) und Al Sharaf (Sharaf *et al.* 2007) verglichen die Reproduzierbarkeit einer Bewertung durch Augenschein und durch Kriterien. Die verwendeten Bewertungskriterien werden lediglich erwähnt, jedoch nicht die Methode der Entstehung. Hinkelmann und Long (Hinkelman und Long 1973) stellten den in ihrer Studie verwendeten Bewertungsbogen anhand von Lehrbuchinhalten und Ideen der Assistenzärzte zusammen. In der Gruppendiskussion wurde auf diese Art ein Bewertungsbogen für die studentischen Arbeiten entwickelt. Gaines (Gaines *et al.* 1974) verglich in seiner Arbeit die interindividuelle Reliabilität einer Bewertung per Augenschein und einer Bewertung anhand von Punkten, die bildlich dargestellt wurden. Die Bilder zeigten Idealformen der Arbeiten, auf die nicht weiter eingegangen wurde. Bedi (Bedi *et al.* 1987) untersuchte ob die Verwendung von Bildern in Zusammenhang mit Kriterien die Zuverlässigkeit der Beurteilung erhöhte. Wie auch bei Gaines (Gaines *et al.* 1974) kann man dieser Studie nicht entnehmen wie die Kriterien und die dazugehörigen Bilder erstellt wurden. All diese Studien haben gemeinsam, dass der Entwicklung der Bewertungskriterien kein Verfahren zur Grunde liegt, welches ermöglicht die Kriterien allgemeingültig zu verwenden.

In vorliegender Studie werden erstmals Bewertungskriterien verwendet, die zuvor anhand einer Literaturrecherche sowie NGT und Matrix-Analyse erstellt wurden. Diese werden mit der Bewertung per Augenschein verglichen.

5.1.2 Matrix-Analyse

Im Rahmen der prozentualen Gewichtung wurde die Matrix-Analyse (Horn 1991) angewendet. Diese ermöglicht verschiedene Punkte einer Liste miteinander zu vergleichen und sie gegeneinander zu gewichten. Auf diese Art und Weise kann man einen Wert für die Gewichtung im Vergleich zu den übrigen errechnen. In der Literatur findet sich auch der Begriff Cross-Impact-Analyses (Gordon 1968; Gordon 1969; Enzer 1970; Enzer 1971; Enzer 1972; Sage 1977). Im Zuge dieser Analyse werden, wie auch bei der von uns in dieser Studie verwendeten Art der Matrix-

Analyse, zwei Listen gegenübergestellt und gegeneinander gewichtet. Hauptsächlich findet die Cross-Impact-Analyse ihren Gebrauch in der Risikoabschätzung von beispielsweise Naturkatastrophen oder anderen Vorkommnissen, die einer Einschätzung bedürfen. Das Prinzip besteht darin, im Vergleich jedem der beiden Punkte einen Wert zuzuschreiben. Die Summe der beiden Werte muss immer zehn ergeben.

5.2 Diskussion der Methode

5.2.1 Zeitraum der Bewertung

Zwischen den Beurteilungsdurchgängen jedes Bewerter wurde ein Zeitraum von 2 Wochen festgelegt. Dieser zeitliche Abstand sollte sicherstellen, dass der jeweilige Bewerter sich nicht an die bereits benoteten Prüfungsarbeiten erinnerte und somit nicht unvoreingenommen wäre. In der Literatur sind verschiedene zeitliche Abstände zwischen Bewertungswiederholungen zu finden. Lilley und Mitarbeiter (Lilley *et al.* 1968) wählten wie Hinkelman und Long (Hinkelman und Long 1973) einen Abstand von 1 Monat zwischen 2 Bewertungsdurchgängen. Bedi und Mitarbeiter (Bedi *et al.* 1987) hielten einen Abstand von 1 Woche, Aly Sharaf (Sharaf *et al.* 2007) von 3 Tagen und Vann und Mitarbeiter (Vann *et al.* 1983) von lediglich 1 Tag zwischen den wiederholten Bewertungen. Auf Grundlage dieser Daten wurde der Abstand von 2 Wochen gewählt. Dies hat sich als praktikabel erwiesen.

5.2.2 Fragebogen

Im Rahmen der Studie sollten sowohl die Ausbilder als auch die Studierenden an einer Befragung teilnehmen. Die Fragebögen der beiden Gruppen unterschieden sich in einigen Fragen, um gezielt auf die jeweiligen Gruppen einzugehen. Zur Beantwortung der jeweiligen Fragen sollten die Teilnehmer ihre Antwort mithilfe einer Likert Skala geben (Komorita 1963; Rost 1996). Eine Likert Skala kann mit unterschiedlich vielen Antwortmöglichkeiten ausgestattet werden. In diesem Fall gab es vier Antwortmöglichkeiten. Diese Bestanden aus: trifft voll zu, trifft eher zu, trifft eher nicht zu und trifft gar nicht zu. Eine neutrale Antwortmöglichkeit wurde bewusst weggelassen, um ein eindeutiges Meinungsbild zu erlangen. Lediglich im letzten Abschnitt der Befragung, der Einschätzung der Verbesserung oder Verschlechterung der Beurteilung in Bezug auf die jeweilige Prüfung wurde eine neutrale

Antwortmöglichkeit eingefügt: klare Verbesserung, Verbesserung, Neutral, Verschlechterung und klare Verschlechterung. Dies war nötig, um den aktuellen Eindruck der Ausbilder festhalten zu können.

Alternativ zur verwendeten Methode der Likert Skala gibt es auch die visuelle Analogskala (van Laerhoven *et al.* 2004). Bei dieser Methode macht der Befragte auf einer horizontalen Linie, die durch die Extremwerte begrenzt wird, einen senkrechten Strich. Nachfolgend wird der Strich in einen Wert von eins bis zehn umgerechnet. Auf diese Art erhält man eine Aussage über die Empfindung des Befragten. Diese Art der Befragung wird häufig in der Schmerzmedizin angewandt. Im vorliegenden Fall fand diese Methode der Befragung keine Berücksichtigung, da sich kein differenziertes Bild der Antworten ergeben hätte.

5.2.3 Datenanalyse

Um die Reliabilität der verschiedenen Bewertungssysteme zu errechnen, wurden zum einen der Interklassenkorrelationskoeffizient und zum anderen die mittelwertigen Abweichungen berechnet. Wichtig war es uns zu untersuchen wie groß die Varianz bei der jeweiligen Beurteilungsmethode ist und ob sie verringert werden kann durch die Bewertung durch einen vordefinierten Bewertungsbogen. Daher wurde zu diesem Zweck der Interklassenkorrelationskoeffizient mit absoluter Übereinstimmung, 2fach zufällig mit einem Konfidenzintervall von 95% gewählt. Auf diese Art und Weise konnte sowohl das Ausmaß der Übereinstimmung zwischen den Bewertungsdurchgängen als auch die Bestimmung der Signifikanz vorgenommen werden. Bei der Literaturrecherche wurde klar, dass bei vergleichbaren Studien auch der ICC angewandt wurde (Shrout 1979; Kellersmann 2008).

5.3 Diskussion der Ergebnisse

Zum vorliegenden Thema wurden bereits einige Untersuchungen durchgeführt. Eine Steigerung der Wiederholbarkeit und Transparenz in der Notengebung wurde immer angestrebt. Der Nutzen und die Vorteile der Bewertungskriterien wurden und werden noch immer hinterfragt. Somit bestand die Notwendigkeit Bewertungskriterien zu entwickeln, die auf Literaturrecherche und Expertenmeinung einschlägiger Zentren basieren und die Reliabilität der Beurteilung von klinischen Arbeiten steigert. Sie

sollten zudem anwendbar und verständlich sein sowohl für Ausbilder als auch für Studierenden.

Die Mehrheit der Studien und Untersuchungen zu diesem Thema beschäftigte sich mit vorklinischen Arbeiten. Je nach Untersucher unterscheiden sich die Auswertungsmethoden (Cohen's Kappa, ICC) und die Ergebnisse. Die beschriebenen Ergebnisse lassen sich jedoch in ihrer Tendenz vergleichen.

Ziel dieser Studie war es, die Reliabilität der Bewertung von zwei verschiedenen Bewertungssystemen zu errechnen. Es sollte gezeigt werden, dass die Zuverlässigkeit durch die Bewertung mit einem Bewertungsbogen im Vergleich zu einer Bewertung per Augenschein deutlich höher ist.

Im Rahmen dieser Studie wurde die Reliabilität zwischen Ausbildern und Studierenden verglichen. Sowohl bei Ausbildern als auch bei Studierenden wurde deutlich, dass der Grad der Übereinstimmung im Rahmen der Beurteilung per Augenschein deutlich schlechter war als mittels Bewertungskriterien. Dies konnte durch beide Auswertungsmethoden, Bildung des Intraklassenkorrelationskoeffizienten (ICC) und Berechnung der gemittelten Abweichung der Bewertungsdurchgänge, bewiesen werden. Bei den Ausbildern konnte eine Steigerung des Intraklassenkorrelationskoeffizienten (ICC) von 0,445 bei der Bewertung per Augenschein zu 0,944 bei der Beurteilung mit Bewertungskriterien erreicht werden. Bei den Studierenden konnte eine Steigerung von 0,315 auf 0,949 vollzogen werden bezogen auf alle sechs Module.

Die Ergebnisse der Befragung der Ausbilder und Studierenden zeigte, dass die Verwendung des Kriterienbogens zur Beurteilung der Arbeiten positiv aufgenommen wurde. Die Bewertungskriterien wurden als korrekt, vollständig, klar formuliert und nicht als zu umfangreich empfunden. Die Studierenden hatten das Gefühl, dass sie zudem selbständiger arbeiten konnten und die eigenen Defizite selbst erkannten. Für beide Gruppen erleichterte sich die Kommunikation während des Erlernens und der Erläuterung der Prüfungsnoten. Der Zeitaufwand wurde im Gegensatz dazu als höher empfunden. Aus Sicht der Ausbilder profitierte am wenigsten die Anfertigung einer Wurzelfüllung von den Kriterien. Dies könnte man damit erklären, dass zur Benotung Röntgenbilder vorlagen und Vergleiche der verschiedenen Arbeiten erleichterten. Am meisten profitierte die Anfertigung der Kompositfüllung von der Anwendung der Bewertungskriterien. Die Präparationen für eine Klasse-II-Kavität für

eine Kompositfüllung, für die Goldteilkrone und die Keramikteilkrone lagen sehr dicht beieinander und profitierten aus Sicht der Ausbilder auch von den Bewertungskriterien.

Eine erste Studie, die eine Steigerung der Verlässlichkeit der Benotung durch die Verwendung von Kriterien mit sich brachte, wurde 1968 von Lilley und Mitarbeitern (Lilley *et al.* 1968) veröffentlicht. Zu Beginn lag der ICC bei 0,12- 0,28, nach der Verwendung der Kriterien bei 0,51- 0,63. Dies ist zwar keine signifikante, jedoch eine ernst zu nehmende Verbesserung des Intraklassenkorrelationskoeffizienten, der damals Anlass gab weiter an diesem Thema zu arbeiten. Vergleicht man die Ergebnisse dieser Studie mit den vorliegenden Ergebnissen kann man feststellen, dass beide Untersuchungen eine Verbesserungen der Reliabilität hervorbrachten. Im Gegensatz zu den Ergebnissen von Lilley und Mitarbeitern (Lilley *et al.* 1968) wurde in der vorliegenden Studie eine signifikante Steigerung der Reliabilität erreicht.

Auch Fuller (Fuller 1972) verglich im Jahr 1972 die Beurteilung durch Checkliste mit der damals gängigen Methode per Augenschein und war mit der Steigerung der Reliabilität und des ICC von 0,4 auf 0,6 nicht zufrieden. Man könnte nicht von einer signifikanten Steigerung ausgehen und stand der Verwendung und dem Nutzen der Bewertungskriterien kritisch gegenüber. Diese Studie hat mit der vorliegenden Studie gemeinsam, dass eine Steigerung der Wiederholbarkeit der Bewertung erzielt werden konnte. Die fehlende Signifikanz, die Fuller (Fuller 1972) nicht zufrieden stellte, unterscheidet die Ergebnisse der beiden Studien.

Gaines und Mitarbeiter (Gaines *et al.* 1974) untersuchten ob, eine Verbesserung der Reliabilität zu erreichen wäre, wenn man der gängigen Bewertungsmethode Bilder und Erläuterungen hinzufügte. In den Ergebnissen wurde beschrieben, dass der Intraklassenkorrelationskoeffizient (ICC) von 0,26 im ersten Durchgang auf 0,55 erhöht werden konnte. Die Ergebnisse von Gaines und Mitarbeitern (Gaines *et al.* 1974) und die Eigenen weisen die gleiche Tendenz auf, jedoch fehlt auch Gaines und Mitarbeitern (Gaines *et al.* 1974) die Signifikanz. Zudem kann man auf Grund der abweichenden Methodik beide Studien nur bedingt miteinander vergleichen.

Vier Jahre später arbeiteten Dhuru und Mitarbeiter (Dhuru *et al.* 1978) an einer Beurteilungsmethode um die Verlässlichkeit der Bewertung zu steigern. Auch hier wurde zu Beginn nach subjektiven klinischen Maßstäben bewertet. Im Anschluss musste die Bewertung mit einem 10-Punkte Bogen durchgeführt werden. Eine

Erhöhung des ICC um 0,131 bei den Zahnärzten war zu erkennen. Diese fiel allerdings so gering aus, dass man nicht von einem signifikanten Gewinn an Reliabilität sprechen konnte. Im Gegensatz dazu konnte in der vorliegenden Studie ein deutlicher Anstieg des ICC und die erwünschte Signifikanz erzielt werden. Goepferd und Kerber (Goepferd und Kerber 1980) kamen im Gegensatz zu der aktuell durchgeführten Studie zu einem ähnlichen Ergebnis wie Dhuru und Mitarbeiter (Dhuru *et al.* 1978). Die Reliabilität konnte zwar gesteigert werden, jedoch nicht in dem Maß, dass man von einer statistischen Signifikanz ausgehen konnte und machte somit weitere Studien notwendig.

Vann und Mitarbeiter (Vann *et al.* 1983) jedoch konnten keine Vorteile erkennen durch die Verwendung von Checklisten oder vordefinierten Kriterienlisten. Vann stellte im Rahmen seiner Studie drei verschiedene Beurteilungssysteme einander gegenüber (Augenschein, Checkliste, Checkliste mit zu erfüllenden Unterkriterien pro Punkt) und konnte in Bezug auf die Reliabilität keine signifikanten Unterschiede feststellen. Diese Studie unterscheidet sich deutlich zu der vorliegenden Studie, in der gezeigt werden konnte, dass eine statistisch signifikante Reliabilität vorliegt, wenn bei der Bewertung der Prüfungsarbeiten die vordefinierten Bewertungskriterien verwendet wurden im Gegensatz zur Beurteilung per Augenschein.

Bedi (Bedi *et al.* 1987) verglich Kronenpräparationen der Klasse eins, zwei und drei und errechnete die Wiederholbarkeit der Beurteilung durch Augenschein- und bebilderter Kriterienbewertung. Laut dieser Studie konnte eine Erhöhung der Wiederholbarkeit von $\kappa = 0,52$ auf $\kappa = 0,71$ erzielt werden. In diesem Fall konnte man nicht von einer signifikanten Steigerung der Wiederholbarkeit der Bewertung ausgehen und unterscheidet sich somit von den Ergebnissen der eigenen Studie.

Im Gegensatz zu dieser Studie konnten Aly Sharaf und Mitarbeiter (Sharaf *et al.* 2007) im Jahr 2007 in ihrer Studie keine Verbesserung der Beständigkeit der Benotung durch die Verwendung von Kriterien im Vergleich zur Beurteilung per Augenschein feststellen.

Kellersmann (Kellersmann 2008) befasste sich ein Jahr später mit einer ähnlichen Thematik. Vorklinische Arbeiten wurden auch in dieser Studie sowohl von Zahnärzten als auch von Studierenden bewertet. Diesen Bewertergruppen standen auch zwei Beurteilungsmethoden zur Verfügung, zum einen per Augenschein und zum Anderen durch einen strukturierten Bewertungsbogen. Kellersmann beschrieb

dass sich die Reproduzierbarkeit, berechnet durch den Intraklassenkorrelationskoeffizienten (ICC), der Zahnärzte von 0,59 auf 0,83 steigern ließ. Fasste man Zahnärzte und Studierende zusammen konnte man eine Steigerung von 0,66 auf 0,76 erkennen.

Dieses Ergebnis hat die gleiche Tendenz wie die selbst durchgeführte Studie, steht jedoch im Kontrast zu den zuvor beschriebenen Ergebnissen, die zwar eine Verbesserung des ICC hervorbrachten, jedoch ohne eine statistische Signifikanz. Trotz der bisher veröffentlichten Studien zu diesem Thema bestand somit noch immer das Problem der unzureichenden interpersonellen Reliabilität (Sharaf *et al.* 2007).

Zusammenfassend kann man sagen, dass jede Studie eine Verbesserung der Reliabilität durch die Verwendung der Kriterien zeigte. Keine Studie brachte als Ergebnis hervor, dass die Beurteilung per Augenschein der Bewertung durch Kriterien überlegen sei. Die Gründe für die fehlende Signifikanz sind schwer auszumachen. Die Kriterienlisten und Bilder, die der Bewertung zu Grunde lagen, sind leider nicht immer in den Studien aufgeführt, so dass sich nur mutmaßen lässt, was die Gründe für die fehlende Signifikanz sein könnten. Faktoren, die das Ergebnis der Studie beeinflussen könnten, sind der Ausbildungsgrad und die Erfahrung der Lehrkörper sowie der Wissensstand der teilnehmenden Studierenden. Auch ist die Kollegialität der Studierenden untereinander nicht zu vernachlässigen, die es verhindern könnte gegenseitig schlechte Noten zu vergeben (Kellersmann 2008). Eine weitere Rolle spielt die Art der Kriterienliste. Handelt es sich dabei um eine reine Aufzählung oder sind zusätzlich auch Erläuterungen und Hilfestellungen sowie Bilder verfügbar.

Man kann somit schlussfolgern, dass die Verwendung von allgemeingültigen Bewertungskriterien zur Beurteilung von klinischen Arbeiten in der Ausbildung der Studierenden eine Erhöhung der Zuverlässigkeit und Reproduzierbarkeit der Benotung mit sich bringt.

Die intrapersonelle Reliabilität ist umso besser, je eindeutiger und verständlicher die Kriterien formuliert sind. Anhand der deutlich signifikanten Reliabilitätswerte auf der einen Seite und der subjektiven Empfindung der Ausbilder und Studierenden auf der anderen Seite kann davon ausgegangen werden, dass die Bewertungskriterien sowohl eindeutig als auch verständlich formuliert sind. Man sieht zudem, dass die

ICC-Werte für Ausbilder und Studierende bei der Verwendung der Bewertungskriterien nur minimal differieren. Daraus lässt sich schließen, dass auf Grund der eindeutigen Formulierung der Bewertungskriterien, eine Benotung sogar nicht nur durch Zahnärzte, sondern auch von Studierenden durchgeführt oder zumindest unterstützt werden könnte.

Man kann auf jeden Fall festhalten, dass durch die Verwendung von vordefinierten Bewertungskriterien, die intrapersonelle Reliabilität deutlich gesteigert werden kann im Vergleich zur Beurteilung per Augenschein. Als Ausblick könnte hier sogar darüber nachgedacht werden studentische Tutoren für die Bewertung einzusetzen zur Entlastung des Lehrpersonals. Zudem erreicht man eine höhere forensische Sicherheit der Bewertung mittels Kriterien. Darüber hinaus ist die Kommunikation zwischen Ausbildungspersonal und Studierendem vereinfacht auf Grund der objektiven Beurteilung.

Es wäre von Interesse, wenn durch weitere Studien oder Untersuchungen die erzielten Ergebnisse bestätigt würden. Um den erhöhten Zeitaufwand zu reduzieren, sollte darüber nachgedacht werden eine computerunterstützte Be- und Auswertung der Arbeiten zu entwickeln und zu erproben.

6 Zusammenfassung

Ziel der vorliegenden Arbeit war es erstellte Bewertungskriterien für praktische Studentarbeiten im Phantomkurs der Zahnerhaltung hinsichtlich ihrer Reproduzierbarkeit bei Anwendung durch Ausbilder und Studierende mit der traditionellen Bewertung per Augenschein zu vergleichen.

Die Studie wurde im Phantomkurs der Poliklinik für Zahnerhaltung und Parodontologie der LMU durchgeführt. Die Studiendauer erstreckte sich über zwei Jahre, von September 2011 bis Juli 2013. Pro Themenkomplex (Präparation Klasse-II-Kavität für direkte Kompositfüllung, direkte Kompositfüllung mit Approximalkontakten im Seitenzahnbereich, Präparation einer klassischen Stufenteilkrone für eine indirekte Goldrestauration, Präparation für eine Keramikteilkrone, Trepanation eines bleibenden Molaren, Wurzelkanalfüllung an einem Frontzahn) wurden 20 praktische Arbeiten von zwei Ausbildern und zwei Studierenden unabhängig voneinander im Abstand von zwei Wochen sowohl per Augenschein als auch mit Hilfe der Bewertungskriterien doppelbewertet. Im Anschluss an die praktischen Prüfungen wurde eine Befragung der Ausbilder und Studierenden des Phantomkurses durchgeführt. Im Rahmen der Befragung sollte geklärt werden, wie die Ausbilder und Studierenden die Bewertungskriterien in Bezug auf Transparenz, Forensik, Güte für das Erlernen der praktischen Fertigkeiten und das Feedback einschätzen.

Die statistische Auswertung der Daten erfolgte mit der Software SPSS, Version 21. Zur Bestimmung der intrarater Reliabilität wurde der Intraklassenkorrelationskoeffizient (ICC) mit einem 95%igen Konfidenzintervall gebildet. Die gewonnenen Daten wurden graphisch mit Microsoft Excel dargestellt und deskriptiv beschrieben. Des Weiteren wurden für jeden Prüfer die gemittelten Abweichungen seiner Bewertung für jede der sechs Prüfungen gebildet und graphisch dargestellt, um diese deskriptiv zu vergleichen. Die Auswertung der Fragebögen erfolgte auch deskriptiv.

Die Bestimmung des Intraklassenkorrelationskoeffizienten (ICC) für Ausbilder und Studierende ergab für die Bewertung per Augenschein signifikant geringere Werte als für die Bewertung durch Kriterien. Für die Bewertung per Augenschein konnte für Ausbilder ein Mittelwert von 0,445, für Studierende von 0,316 errechnet werden. Die Mittelwerte lagen im Gegensatz dazu für die Bewertung

durch Kriterien bei 0,944 für Ausbilder und 0,949 für Studierende über alle Themenkomplexe hinweg.

Die Ergebnisse zeigten sowohl für Ausbilder als auch Studierende signifikant, dass im Gegensatz zur Beurteilung per Augenschein die Bewertung durch Kriterien ein hohes Maß an intrapersoneller Übereinstimmung hervorbrachte. Dies galt für alle sechs Themenkomplexe. Zudem wurde deutlich, dass die Abweichungen der Benotungen zwischen Ausbildern und Studierenden unter der Verwendung der Bewertungskriterien im Vergleich zur Augenscheinbewertung signifikant geringer waren.

Die Auswertung der Fragebögen zeigte, dass Ausbilder und Studierende die Verwendung von Bewertungskriterien mit einem Zugewinn an Transparenz, Forensik, Güte für das Lehren und Erlernen der praktischen Fertigkeiten und mit einer Erleichterung des Feedbacks verbanden. Aus Sicht der Ausbilder profitierte die Anfertigung einer Kompositfüllung im Seitenzahnggebiet am meisten von der Verwendung der Bewertungskriterien. Im Gegensatz dazu war der Nutzen am niedrigsten bei der Anfertigung der Wurzelfüllung.

Basierend auf den vorliegenden Daten kann man somit davon ausgehen, dass die Sicherheit und Wiederholbarkeit der Bewertung durch die Anwendung von vordefinierten Bewertungskriterien gesteigert werden kann. Durch die Anwendung dieser Kriterien wird die Beurteilung der Prüfungsarbeiten für die Studierenden transparenter. Zudem wird die Kommunikation zwischen Ausbildern und Studierenden erleichtert.

7 Literaturverzeichnis

- ÄAppO (2002) Approbationsordnung für Ärzte. B. f. Justiz. §14 (7): 21.
- About I., Murray P. E., Franquin J. C., Remusat M., Smith A. J. (2001). The effect of cavity restoration variables on odontoblast cell numbers and dental repair. *J Dent* 29(2): 109-117.
- Ahmad I. A. (2009). Rubber dam usage for endodontic treatment: a review. *Int Endod J* 42(11): 963-972.
- André Hutsky, Roland Frankenberger (2011). "Keramisch denken" - Präparationsregeln für Inlays und Teilkronen aus Keramik. *Quintessenz* 62(1): 19-22.
- Ayad M. F., Johnston W. M., Rosenstiel S. F. (2009). Influence of dental rotary instruments on the roughness and wettability of human dentin surfaces. *J Prosthet Dent* 102(2): 81-88.
- Banks R. G. (1990). Conservative posterior ceramic restorations: a literature review. *J Prosthet Dent* 63(6): 619-626.
- Basrani B. (2011). Irrigation in endodontic treatment. *Alpha Omegan* 104(1-2): 18-25.
- Bedi R., Lo E., King N. M., Chan T. (1987). The effect of pictorial criteria upon the reliability of assessments of cavity preparations. *J Dent* 15(5): 222-224.
- Benninghoff, Drenckhahn (2008). *Makroskopische Anatomie, Histologie, Embryologie, Zellbiologie. Bd.1. 17. durchgesehene Auflage.* München, Jena, 607 f. Urban und Fischer
- Berger S. B., Paliolol A. R., Cavalli V., Giannini M. (2011). Surface roughness and staining susceptibility of composite resins after finishing and polishing. *J Esthet Restor Dent* 23(1): 34-43.
- Buckley S., Coleman J., Davison I., Khan K. S., Zamora J., Malick S., Morley D., Pollard D., Ashcroft T., Popovic C., Sayers J. (2009). The educational effects of portfolios on undergraduate student learning: a Best Evidence Medical Education (BEME) systematic review. *BEME Guide No. 11. Med Teach* 31(4): 282-298.
- Burke F. J., Cheung S. W., Mjor I. A., Wilson N. H. (1999). Reasons for the placement and replacement of restorations in vocational training practices. *Prim Dent Care* 6(1): 17-20.
- Carney O., McIntosh J., Worth A. (1996). The use of the Nominal Group Technique in research with community nurses. *J Adv Nurs* 23(5): 1024-1029.
- Cavalcanti B. N., Otani C., Rode S. M. (2002). High-speed cavity preparation techniques with different water flows. *J Prosthet Dent* 87(2): 158-161.
- Cavalcanti B. N., Serairdarian P. I., Rode S. M. (2005). Water flow in high-speed handpieces. *Quintessence Int* 36(5): 361-364.

- Chuang S. F., Su K. C., Wang C. H., Chang C. H. (2011). Morphological analysis of proximal contacts in class II direct restorations with 3D image reconstruction. *J Dent* 39(6): 448-456.
- Clark DJ (2005). Shaping and restoring ovoid canal systems. *Endodontic Therapy* 5, 9-13.
- Cohen S. C. (1989). Endodontics and litigation: an American perspective. *Int Dent J* 39(1): 13-16.
- Cohen S., Schwartz S. (1987). Endodontic complications and the law. *J Endod* 13(4): 191-197.
- Davis R. D., Marshall J. G., Baumgartner J. C. (2002). Effect of early coronal flaring on working length change in curved canals using rotary nickel-titanium versus stainless steel instruments. *J Endod* 28(6): 438-442.
- Delp P., Thesen A., Motiwalla J., Seshardi N. (1977). Nominal Group Technique. Systems tools for project planning.
- Demarco F. F., Correa M. B., Cenci M. S., Moraes R. R., Opdam N. J. (2012). Longevity of posterior composite restorations: not only a matter of materials. *Dent Mater* 28(1): 87-101.
- Dhuru V. B., Rypel T. S., Johnston W. M. (1978). Criterion-oriented grading system for preclinical operative dentistry laboratory course. *J Dent Educ* 42(9): 528-531.
- El-Badrawy W. A., Leung B. W., El-Mowafy O., Rubo J. H., Rubo M. H. (2003). Evaluation of proximal contacts of posterior composite restorations with 4 placement techniques. *J Can Dent Assoc* 69(3): 162-167.
- Enzer Selwyn (1970). A case study using forecasting as a decision-making aid. *Futures* 2(4): 341-362.
- Enzer Selwyn (1971). Delphi and cross-impact techniques: An effective combination for systematic futures analysis. *Futures* 3(1): 48-61.
- Enzer Selwyn (1972). Cross-impact techniques in technology assessment. *Futures* 4(1): 30-51.
- European Society of Endodontology (2006). Quality guidelines for endodontic treatment: Consensus report of the European society of Endodontology. *International Endodontic Journal* 2006(39): 921-930.
- Fahl N. Jr. (1997). Optimizing the esthetics of Class IV restorations with composite resins. *J Can Dent Assoc* 63(2): 108-111, 114-105.
- Federlin M., Manner T., Hiller K. A., Schmidt S., Schmalz G. (2006). Two-year clinical performance of cast gold vs ceramic partial crowns. *Clin Oral Investig* 10(2): 126-133.
- Federlin M., Schmidt S., Hiller K. A., Thonemann B., Schmalz G. (2004). Partial ceramic crowns: influence of preparation design and luting material on internal adaptation. *Oper Dent* 29(5): 560-570.
- Feil P. H. (1982). An analysis of the reliability of a laboratory evaluation system. *J Dent Educ* 46(8): 489-494.

- Frankenberger R., Mörig G., Blunck U., Hajto J., Pröbster L., Ahlers MO. (2008). Präparationsregeln für Keramikinlays und -teilronen unter besonderer Berücksichtigung der CAD/CAM-Technologie. Bayerisches Zahnärzteblatt(3): 45-51.
- Frankenberger R., Roth S., Kramer N., Pelka M., Petschelt A. (2003). Effect of preparation mode on Class II resin composite repair. J Oral Rehabil 30(6): 559-564.
- Fuller J. L. (1972). The effects of training and criterion models on interjudge reliability. J Dent Educ 36(4): 19-22.
- Gaines W. G., Bruggers H., Rasmussen R. H. (1974). Reliability of ratings in preclinical fixed prosthodontics: effect of objective scaling. J Dent Educ 38(12): 672-675.
- Gallagher M., Hares T., Spencer J., Bradshaw C., Webb I. (1993). The nominal group technique: a research tool for general practice? Fam Pract 10(1): 76-81.
- Goepferd S. J., Kerber P. E. (1980). A comparison of two methods for evaluating primary class II cavity preparations. J Dent Educ 44(9): 537-542.
- Goldstein R. E. (1989). Finishing of composites and laminates. Dent Clin North Am 33(2): 305-318, 210-309.
- Goodman A., Reader A., Beck M., Melfi R., Meyers W. (1985). An in vitro comparison of the efficacy of the step-back technique versus a step-back/ultrasonic technique in human mandibular molars. J Endod 11(6): 249-256.
- Gordon T. J. (1969). Cross-impact matrices: An illustration of their use for policy analysis. Futures 1(6): 527-531.
- Gordon T. J., Hayward, H. (1968). Initial experiments with the cross impact matrix method of forecasting. Futures 1(2): 100-116.
- Hickel R., Manhart J. (2001). Longevity of restorations in posterior teeth and reasons for failure. J Adhes Dent 3(1): 45-64.
- Himel V. T., Levitan M. E. (2003). Use of nickel titanium instruments for cleaning and shaping root canal systems. J Tenn Dent Assoc 83(2): 29-33.
- Hinkelman K. W., Long N. K. (1973). Method for decreasing subjective evaluation in preclinical restorative dentistry. J Dent Educ 37(9): 13-18.
- Horn R. A., Johnson, C. A. (1991). Topics in Matrix Analysis. Cambridge, New York, Melbourne. Cambridge University Press
- Horton J. N. (1980). Nominal group technique. A method of decision-making by committee. Anaesthesia 35(8): 811-814.
- Hsieh Y. D., Gau C. H., Kung Wu S. F., Shen E. C., Hsu P. W., Fu E. (2007). Dynamic recording of irrigating fluid distribution in root canals using thermal image analysis. Int Endod J 40(1): 11-17.

- Hugo B. (1999). [Oscillating procedures in the preparation technic (I)]. Schweiz Monatsschr Zahnmed 109(2): 140-160.
- Huth K.C. Erstellung von Bewertungskriterien für praktische Arbeiten im Phantomkurs der Zahnerhaltung und Parodontologie der Zahnklinik der Universität München. Projektarbeit, Master of Medical Education. In Bearbeitung.
- Huth K.C., Baumann M., Hickel R., Paschos E. Assessment of practical tasks in the Phantom course of Conservative Dentistry by predefined criteria: a comparison between self assessment by students and assessment by instructors. In Bearbeitung
- Javaheri M., Maleki-Kambakhsh S., Etemad-Moghadam Sh (2010). Efficacy of two caries detector dyes in the diagnosis of dental caries. J Dent (Tehran) 7(2): 71-76.
- Jenkins S. M., Dummer P. M., Gilmour A. S., Edmunds D. H., Hicks R., Ash P. (1998). Evaluating undergraduate preclinical operative skill; use of a glance and grade marking system. J Dent 26(8): 679-684.
- Johnson B. R. (2009). Endodontic access. Gen Dent 57(6): 570-577; quiz 578-579, 595, 679.
- Jünger J., Schellberg D., Nikendei C. (2006). Student's self-assessment of clinical competence and objective clinical performance in OSCE evaluation. GMS Zeitschrift für Medizinische Ausbildung 23(3).
- Kellersmann C. T. (2008). Reliabilität der Beurteilung vorklinischer Phantomarbeiten bei Einsatz eines strukturierten Bewertungsbogens. Dissertation. Münster.
- Klaiber B., Hugo B. (1998). Metallinlay und Teilkrone. Praxis der Zahnheilkunde, Bd. 2 (Kariologie und Füllungstherapie). München, Heidemann, D. 4. Auflage: 255-278.
- Komorita S. S. (1963). Attitude Content, Intensity, and the Neutral Point on a Likert Scale. J Soc Psychol 61: 327-334.
- Kramer N., Reinelt C., Richter G., Petschelt A., Frankenberger R. (2009). Nanohybrid vs. fine hybrid composite in Class II cavities: clinical results and margin analysis after four years. Dent Mater 25(6): 750-759.
- Krasner P., Rankow H. J. (2004). Anatomy of the pulp-chamber floor. J Endod 30(1): 5-16.
- Lanning S. K., Waldrop T. C., Gunsolley J. C., Maynard J. G. (2003). Surgical crown lengthening: evaluation of the biological width. J Periodontol 74(4): 468-474.
- Leo Mente (2012). Risiko-Management Perforation; Risk-management perforation. Deutsche Zahnärztliche Zeitschrift 67(11): 695-707.
- Lesage B. (2011). Finishing and polishing criteria for minimally invasive composite restorations. Gen Dent 59(6): 422-428.
- Li K. Z., Gao Y., Zhang R., Hu T., Guo B. (2011). The effect of a manual instrumentation technique on five types of premolar root canal geometry

- assessed by microcomputed tomography and three-dimensional reconstruction. *BMC Med Imaging* 11: 14.
- Lilley J. D., ten Bruggen Cate H. J., Holloway P. J., Holt J. K., Start K. B. (1968). Reliability of practical tests in operative dentistry. *Br Dent J* 125(5): 194-197.
- Lin Hui-Ching (2011). Use of rubber dams during root canal treatment in Taiwan. *Journal of the Formosan Medical Association* 110(6): 397-400.
- Loomans B. A., Opdam N. J., Roeters F. J., Bronkhorst E. M., Burgersdijk R. C. (2006). Comparison of proximal contacts of Class II resin composite restorations in vitro. *Oper Dent* 31(6): 688-693.
- Lussi A. (1995). [Damage to neighboring teeth during the preparation of proximal cavities. An in-vivo study]. *Schweiz Monatsschr Zahnmed* 105(10): 1259-1264.
- Lussi A., Gygax M. (1998). Iatrogenic damage to adjacent teeth during classical approximal box preparation. *J Dent* 26(5-6): 435-441.
- Macluskey M., Durham J., Balmer C., Bell A., Cowpe J., Dawson L., Freeman C., Hanson C., McDonagh A., Jones J., Millsopp L., Oliver R. (2011). Dental student suturing skills: a multicentre trial of a checklist-based assessment. *Eur J Dent Educ* 15(4): 244-249.
- Margeas R. C. (2010). Keys to success in creating esthetic class IV restorations. *J Esthet Restor Dent* 22(1): 66-71.
- Mizrahi B. (2011). All-ceramic silica/glass-based crowns--clinical protocols. *Br Dent J* 211(6): 257-262.
- Mjor I. A., Toffenetti F. (2000). Secondary caries: a literature review with case reports. *Quintessence Int* 31(3): 165-179.
- Motamedi M. H. (2006). Surgical management of iatrogenic root perforation following endodontic therapy. *N Y State Dent J* 72(5): 40-41.
- Mount G. J., Ngo H. (2000). Minimal intervention: advanced lesions. *Quintessence Int* 31(9): 621-629.
- Murray P. E., Smith A. J., Windsor L. J., Mjor I. A. (2003). Remaining dentine thickness and human pulp responses. *Int Endod J* 36(1): 33-43.
- Naito T. (2005). Better success rate for root canal therapy when treatment includes obturation short of the apex. *Evid Based Dent* 6(2): 45.
- Natkin E., Guild R. E. (1967). Evaluation of preclinical laboratory performance: a systematic study. *J Dent Educ* 31(2): 152-161.
- Norman G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Adv Health Sci Educ Theory Pract* 15(5): 625-632.
- Opdam N. J., Loomans B. A., Roeters F. J., Bronkhorst E. M. (2004). Five-year clinical performance of posterior resin composite restorations placed by dental students. *J Dent* 32(5): 379-383.
- Osborne J. W., Summitt J. B. (1998). Extension for prevention: is it relevant today? *Am J Dent* 11(4): 189-196.

- Ottl P., Lauer H. C. (1998). Temperature response in the pulpal chamber during ultrahigh-speed tooth preparation with diamond burs of different grit. *J Prosthet Dent* 80(1): 12-19.
- Ozturk B., Usumez A., Ozturk A. N., Ozer F. (2004). In vitro assessment of temperature change in the pulp chamber during cavity preparation. *J Prosthet Dent* 91(5): 436-440.
- Parashos P., Messer H. H. (2006). Rotary NiTi instrument fracture and its consequences. *J Endod* 32(11): 1031-1043.
- Parkinson C. F. (1976). Excessive crown contours facilitate endemic plaque niches. *J Prosthet Dent* 35(4): 424-429.
- Patel S., Rhodes J. (2007). A practical guide to endodontic access cavity preparation in molar teeth. *Br Dent J* 203(3): 133-140.
- Qvist V., Johannessen L., Bruun M. (1992). Progression of approximal caries in relation to iatrogenic preparation damage. *J Dent Res* 71(7): 1370-1373.
- Raghu R., Srinivasan R. (2011). Optimizing tooth form with direct posterior composite restorations. *J Conserv Dent* 14(4): 330-336.
- Rodrigues Junior S. A., Pin L. F., Machado G., Della Bona A., Demarco F. F. (2010). Influence of different restorative techniques on marginal seal of class II composite restorations. *J Appl Oral Sci* 18(1): 37-43.
- Rost J (1996). *Lehrbuch Testtheorie, Testkonstruktion*. Bern u.a.
- Sage A. P. (1977). *Methodology for Large-scale Systems*. New York. McGraw-Hill
- Schafer E., Tepel J., Hoppe W. (1995). Properties of endodontic hand instruments used in rotary motion. Part 2. Instrumentation of curved canals. *J Endod* 21(10): 493-497.
- Schilder H. (2006). Filling root canals in three dimensions. 1967. *J Endod* 32(4): 281-290.
- Sharaf A. A., AbdelAziz A. M., El Meligy O. A. (2007). Intra- and inter-examiner variability in evaluating preclinical pediatric dentistry operative procedures. *J Dent Educ* 71(4): 540-544.
- Shrout P. E. & Fleiss, J. L. (1979). Intraclass correlation: Uses in assessing rater reliability. *Psychological Bulletin* 86: 420-428.
- St-Georges A. J., Sturdevant J. R., Swift E. J., Jr., Thompson J. Y. (2003). Fracture resistance of prepared teeth restored with bonded inlay restorations. *J Prosthet Dent* 89(6): 551-557.
- Stakenas J. A., Merrick N. P. (1982). Application of SPSS to management of computer-assisted instruction usage statistics. *Bull Med Libr Assoc* 70(2): 231-233.
- Stoll R., Sieweke M., Pieper K., Stachniss V., Schulte A. (1999). Longevity of cast gold inlays and partial crowns--a retrospective study at a dental school clinic. *Clin Oral Investig* 3(2): 100-104.

-
- Summitt J. B., Robbins, W. J., Hilton, T. J. & Schwartz, R. S. (2006). Fundamentals of operative Dentistry A Contemporary Approach (3 ed.), 265-266. Quintessence Publishing
- Tarnow D. P., Magner A. W., Fletcher P. (1992). The effect of the distance from the contact point to the crest of bone on the presence or absence of the interproximal dental papilla. *J Periodontol* 63(12): 995-996.
- Taschieri S., Bettach R., Lolato A., Moneghini L., Fabbro M. D. (2011). Endodontic surgery failure: SEM analysis of root-end filling. *J Oral Sci* 53(3): 393-396.
- Terry D. A. (2005). Restoring the incisal edge. *N Y State Dent J* 71(5): 30-35.
- Türp JC Gerds TH, Schneider U (2002). Variability of rates when assessing the the preclinical laboratory performance of dental students. *Dtsch Zahnärztl Z* 57(9): 526-531.
- Usha G., Thippeswamy H., Nagesh L. (2012). Comparative assessment of validity and reliability of the Oral Impacts on Daily Performance (OIDP) frequency scale: a cross-sectional survey among adolescents in Davanagere city, Karnataka, India. *Int J Dent Hyg*.
- van Laerhoven H., van der Zaag-Loonen H. J., Derkx B. H. (2004). A comparison of Likert scale and visual analogue scales as response options in children's questionnaires. *Acta Paediatr* 93(6): 830-835.
- Vann W. F., Machen J. B., Hounshell P. B. (1983). Effects of criteria and checklists on reliability in preclinical evaluation. *J Dent Educ* 47(10): 671-675.
- Wikipedia (2013). Intra-Klassen-Korrelation.
- Wilson E. G., Mandradjieff M., Brindock T. (1990). Controversies in posterior composite resin restorations. *Dent Clin North Am* 34(1): 27-44.
- Wirtz M. & Caspar, F. (2002). Beurteilerübereinstimmung und Beurteilerreliabilität. Göttingen: Hogrefe.
- Wu M. K., Wesselink P. R., Walton R. E. (2000). Apical terminus location of root canal treatment procedures. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 89(1): 99-103.
- Yared G. M., Dagher F. B., Machtou P. (1997). Influence of the removal of coronal gutta-percha on the seal of root canal obturations. *J Endod* 23(3): 146-148.
- Young G. R., Parashos P., Messer H. H. (2007). The principles of techniques for cleaning root canals. *Aust Dent J* 52(1 Suppl): S52-63.

8 Abbildungsverzeichnis

Abbildung 1: Graphische Darstellung der gemittelten Abweichung in Noten für 12 Arbeiten pro Prüfungsart der Überprüfung der Verständlichkeit pro Prüfung. MW = Mittelwert	52
Abbildung 2: Graphische Darstellung der gemittelte Abweichung der Bewertung der praktischen Arbeiten durch beide Methoden pro Prüfung und Bewerter. MW = Mittelwert	55
Abbildung 3: Graphische Darstellung der gemittelte Abweichung der Bewertung der praktischen Arbeiten durch beide Methoden pro Prüfung und Studierender. MW = Mittelwert	57
Abbildung 4: Graphische Darstellung der mittelwertigen Notenabweichungen aller Bewerter pro Prüfung und Beurteilungssystem für 20 Arbeiten pro Prüfungsart. MW = Mittelwert	59
Abbildung 5: Graphische Darstellung der Aussage, dass die Bewertungskriterien korrekt seien aus Sicht der Studierenden.	60
Abbildung 6: Graphische Darstellung der Aussage, dass die Bewertungskriterien korrekt seien aus Sicht der Ausbilder	61
Abbildung 7: Graphische Darstellung der Aussage, dass die Bewertungskriterien vollständig seien aus Sicht der Studierenden	62
Abbildung 8: Graphische Darstellung der Aussage, dass die Bewertungskriterien vollständig seien aus Sicht der Ausbilder.	63
Abbildung 9: Graphische Darstellung der Aussage, dass die Bewertungskriterien zu umfangreich seien aus Sicht der Studierenden.....	64
Abbildung 10: Graphische Darstellung der Aussage, dass die Bewertungskriterien zu umfangreich seien aus Sicht der Ausbilder.....	65
Abbildung 11: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien klar für den Studierenden seien.	66
Abbildung 12: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien klar für den Ausbilder seien.	67
Abbildung 13: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Studierenden erleichtere das praktische Arbeiten zu strukturieren.	68

Abbildung 14: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Ausbilder erleichtere das praktische Arbeiten strukturiert zu lehren. 69

Abbildung 15: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien die Vorbereitung des Studierenden erleichtere..... 70

Abbildung 16: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Studierenden erleichtere die Arbeitsschritte selbstständig durchzuführen..... 71

Abbildung 17: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Studierenden erleichtern eigene Defizite zu erkennen und zu üben. 72

Abbildung 18: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Ausbilder erleichtere, die Unterschiede der Präparationsformen Gold, Keramik und Komposit aufzuzeigen. 73

Abbildung 19: Graphische Darstellung der Aussage der Ausbilder, dass durch die Anwendung der Kriterien das Wesentliche der praktischen Arbeiten besser gelernt würde als ohne Kriterien..... 74

Abbildung 20: Graphische Darstellung der Aussage, dass der Ausbilder durch die Kenntnis der Kriterien die Lernziele des Kurses besser vermitteln könnte. 75

Abbildung 21: Graphische Darstellung der Aussage, dass die Bewertungskriterien es dem Studierenden erleichterten Stärken und Schwächen bei den einzelnen Arbeitsschritten isoliert zu erkennen. 76

Abbildung 22: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es den Studierenden erleichtere die Arbeiten im Laufe der jeweiligen Sequenz zu vergleichen..... 77

Abbildung 23: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Ausbilder erleichtere die Arbeiten im Laufe der jeweiligen Sequenz zu vergleichen. 78

Abbildung 24: Graphische Darstellung der Aussage, dass die Anwendung der Kriterien die forensische Sicherheit der Beurteilung erhöhte..... 79

Abbildung 25: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Ausbilder erleichterte dem Studierenden Feedback zu geben..... 80

Abbildung 26: Graphische Darstellung der Aussage, dass die Anwendung der Bewertungskriterien es dem Studierenden erleichtere kritische Rückmeldungen kritisch zu hinterfragen. 81

Abbildung 27: Graphische Darstellung der Aussage, dass die Anwendung der Kriterien es dem Ausbilder erleichterte die Bewertung gegen Kritik durch den Studierenden zu verteidigen..... 82

Abbildung 28: Graphische Darstellung der Aussage, dass die Anwendung der Kriterien es dem Ausbilder erleichterte, seine Bewertung gegenüber dem Vorgesetzten zu verteidigen..... 83

Abbildung 29: Graphische Darstellung, der Aussage, dass die Anwendung der Kriterien den Zeitaufwand für die Bewertung erhöhe im Vergleich zur Bewertung per Augenschein..... 84

Abbildung 30: Graphische Darstellung der Aussage, dass die Anwendung der Kriterien das Erlernen der praktischen Fertigkeiten beschleunigen würde. 85

Abbildung 31: Graphische Darstellung der Frage: welche der Arbeiten profitiert am meisten von den Bewertungskriterien? Aus Sicht der Ausbilder. 86

9 Tabellenverzeichnis

- Tabelle 1: NGT- Analyse für Prüfung Präparation Klasse II für Komposit. In der linken und mittleren Spalte sind die Bewertungskriterien aufgelistet. In der rechten Spalte ist pro Bewertungskriterium die Summe der von den 4 Zentren vergebenen Punkte der NGT-Analyse aufgelistet. ... 33
- Tabelle 2: NGT- Analyse für Prüfung Legen einer Kompositfüllung Klasse II. In der linken und mittleren Spalte sind die Bewertungskriterien aufgelistet. In der rechten Spalte ist pro Bewertungskriterium die Summe der von den 4 Zentren vergebenen Punkte der NGT-Analyse aufgelistet. 33
- Tabelle 3: NGT- Analyse für Prüfung Präparation Gold TK. In der linken und mittleren Spalte sind die Bewertungskriterien aufgelistet. In der rechten Spalte ist pro Bewertungskriterium die Summe der von den 4 Zentren vergebenen Punkte der NGT-Analyse aufgelistet. ... 34
- Tabelle 4: NGT- Analyse für Prüfung Präparation Keramik TK. In der linken und mittleren Spalte sind die Bewertungskriterien aufgelistet. In der rechten Spalte ist pro Bewertungskriterium die Summe der von den 4 Zentren vergebenen Punkte der NGT-Analyse aufgelistet. 35
- Tabelle 5: NGT- Analyse für Prüfung Trepanation. In der linken und mittleren Spalte sind die Bewertungskriterien aufgelistet. In der rechten Spalte ist pro Bewertungskriterium die Summe der von den 4 Zentren vergebenen Punkte der NGT-Analyse aufgelistet..... 35
- Tabelle 6: NGT- Analyse für Prüfung Wurzelfüllung. In der linken und mittleren Spalte sind die Bewertungskriterien aufgelistet. In der rechten Spalte ist pro Bewertungskriterium die Summe der von den 4 Zentren vergebenen Punkte der NGT-Analyse aufgelistet..... 36
- Tabelle 7: Matrix-Analyse für Prüfung 1. In der linken Spalte ist das Kriterium, in der mittleren Spalte die gesamte Summe der Bewertungspunkte aller 4 Zentren und rechts der Gesamtprozentwert zu sehen..... 37
- Tabelle 8: Matrix-Analyse für Prüfung 2. In der linken Spalte ist das Kriterium, in der mittleren Spalte die gesamte Summe der Bewertungspunkte aller 4 Zentren und rechts der Gesamtprozentwert zu sehen..... 38
- Tabelle 9: Matrix-Analyse für Prüfung 3. In der linken Spalte ist das Kriterium, in der mittleren Spalte die gesamte Summe der Bewertungspunkte aller 4 Zentren und rechts der Gesamtprozentwert zu sehen..... 39

Tabelle 10: Matrix-Analyse für Prüfung 4. In der linken Spalte ist das Kriterium, in der mittleren Spalte die gesamte Summe der Bewertungspunkte aller 4 Zentren und rechts der Gesamtprozentwert zu sehen. 40

Tabelle 11: Matrix-Analyse für Prüfung 5. In der linken Spalte ist das Kriterium, in der mittleren Spalte die gesamte Summe der Bewertungspunkte aller 4 Zentren und rechts der Gesamtprozentwert zu sehen. 41

Tabelle 12: Matrix-Analyse für Prüfung 6. In der linken Spalte ist das Kriterium, in der mittleren Spalte die gesamte Summe der Bewertungspunkte aller 4 Zentren und rechts der Gesamtprozentwert zu sehen. 41

Tabelle 13: Überprüfung der Verständlichkeit. 1. Spalte: Angabe der Konstellation. 2.-7. Spalte: Prüfungsart. In den drei Zeilen werden die ICC-Werte mit den Konvidenzintervallen für den jeweiligen Vergleich angegeben..... 52

Tabelle 14: Bestimmung der Intraklassenkorrelationskoeffizienten zum Vergleich der Reliabilität der Bewertung zwischen Augenschein und Kriterien der sechs Prüfungen durch die Ausbilder. 1 Spalte: Bewertungsart. 2.-7. Spalte Prüfungsart. In den vier Zeilen sind die ICC Werte mit den Konvidenzintervallen zu sehen..... 54

Tabelle 15: Bestimmung der Intraklassenkorrelationskoeffizienten zum Vergleich der Reliabilität der Bewertung zwischen Augenschein und Kriterien der sechs Prüfungen durch die Studierenden. 1 Spalte: Bewertungsart. 2.-7. Spalte Prüfungsart. In den vier Zeilen sind die ICC Werte mit den Konvidenzintervallen zu sehen..... 56

Tabelle 16: Bestimmung der Intraklassenkorrelationskoeffizienten zur Gegenüberstellung der Reliabilität der Bewertung per Augenschein der sechs Prüfungen durch die Ausbilder und Studierenden. 1. Spalte: Bewerter. 2.-7. Spalte Prüfungsart. In den vier Zeilen sind die ICC Werte mit den Konvidenzintervallen zu sehen..... 58

Tabelle 17: Bestimmung der Intraklassenkorrelationskoeffizienten zur Gegenüberstellung der Reliabilität der Bewertung per Kriterien der sechs Prüfungen durch die Ausbilder und Studierenden. 1. Spalte: Bewerter. 2.-7. Spalte Prüfungsart. In den vier Zeilen sind die ICC Werte mit den Konvidenzintervallen zu sehen..... 58

10 Abkürzungen

Abb.	Abbildung
cm	Zentimeter
et al.	et alii (und andere)
evtl.	eventuell
Kap.	Kapitel
mm	Millimeter
mod	mesial-okklusal-distal
MW	Mittelwert
n	Anzahl
o.ä.	oder ähnliches
o.g.	oben genannt
Präp.	Präparation
Tab.	Tabelle
TK	Teilkrone
Trep.	Trepanation
u.a.	unter anderem
v.a.	vor allem
WF	Wurzelfüllung
%	Prozent
z.B.	zum Beispiel

11 Anhang

11.1 Aufgabenstellung für Prüfung 1

Bewertungskriterien: Präparation einer auf die Karieslokalisation bezogenen Klasse-II-Kavität für direkte Kompositfüllungen

Karieslokalisation: OK SZ, Klasse II, mod, Karies an beiden Approximalkontakten sowie okklusal

11.2 Aufgabenstellung für Prüfung 2

Bewertungskriterien: Anatomisch korrekte und ästhetisch hochwertige direkte Kompositfüllung mit Approximalkontakten im Seitenzahnbereich

Auf folgende anatomische Parameter ist insbesondere zu achten:

- Anzahl, Lage und Form der Höcker
- Zentralfissur
- Gestaltung der Randleiste mit Grube in Richtung Zahnmitte und Einziehung koronal des Approximalkontaktes
- ein deutlicher Widerstand an den Approximalkontakten muss mit Zahnseide zu spüren sein

Defektlokalisation: Zahn 36, Klasse II, mod mit Karies im Bereich der Approximalkontakte sowie okklusal

11.3 Aufgabenstellung für Prüfung 3

Bewertungskriterien: Präparation einer klassischen Stufenteilkrone für eine indirekte Goldrestauration

Auf folgende Parameter ist insbesondere zu achten:

- Stufe 1 mm breit palatinal
- Ansträgung der Stufe 30-45° zur Einschubrichtung mit einer Breite von 0,5-1 mm
- Federrand approximal 0,5-1 mm breit
- Kavitätenmindesttiefe 0,5-1 mm ausgehend von der tiefsten Fissur

-
- Approximale Kastentiefe 1,5 mm ausgehend vom Isthmus
 - Planer Kastenboden liegt 90° zur Einschubrichtung
 - Okklusaler Kasten nicht breiter als $\frac{1}{3}$ des bucco-palatinalen Höckerspitzenabstandes
 - Divergenzwinkel der senkrechten Wände je 3-6°
 - Sondierung der Ränder im Bereich des Approximalkontaktes muss möglich sein
 - Minimale oro-vestibuläre Ausdehnung der approximalen Kästen
 - Kontaktpunkte dürfen nicht im Randbereich der Restauration liegen (Außenschliff Kaukantenschutz, Approximalbereich)

Defektlokalisierung: Zahn 16, Klasse II, modp unter Einbeziehung aller Höcker

11.4 Aufgabenstellung für Prüfung 4

Bewertungskriterien: Präparation für eine Keramikteilkrone

Auf folgende anatomische Parameter ist insbesondere zu achten:

- Mindestschichtstärke der späteren Restauration 1,5 mm in allen Dimensionen
- Mindestbreite Isthmus 2 mm
- Öffnungswinkel 8°-10°
- 90° Winkel zwischen Keramik und Zahnschmelze
- Sondierung der Ränder im Bereich des Approximalkontaktes muss möglich sein
- Kontaktpunkte dürfen nicht im Randbereich der Restauration liegen

Defektlokalisierung: Zahn 26, modbp, Einbezug aller Höcker, Karies an beiden Approximalkontakten

11.5 Aufgabenstellung für Prüfung 5

Bewertungskriterien: Trepanation und Zugangskavität

Defektlokalisierung: Molar

11.6 Aufgabenstellung für Prüfung 6

Bewertungskriterien: Wurzelkanalfüllung

Aufgabenstellung: Wurzelkanalfüllung an einem Frontzahn

Aufbereitungsmethode: Step-back-Technik
Obturationsmethode: Laterale Kondensation (Guttapercha, Ah-Plus)
Dokumentation: Datum:
Zahn:
Kanal:
Meßlänge:
Feilenart:
Referenzpunkt:
Arbeitslänge:
ISO- apikale Master-Feile:
ISO- letzte Feile:

Anfertigung apikaler Röntgenbilder und Beschriftung:

- Diagnoseaufnahme
- Meßaufnahme
- Kontrollaufnahme

11.7 Fragebogen Studierende

Studententitel: Evaluation objektiver Bewertungskriterien für praktische Studentenarbeiten

In den bisherigen praktischen Kursen, die Sie während Ihres vorklinischen Studienabschnitts besucht haben, wurden Ihre Arbeiten per Augenschein beurteilt. In diesem Kurs wird die Beurteilung an Hand von Bewertungskriterien durchgeführt. Diese Bewertungskriterien werden Ihnen zu Beginn jeder Themensequenz ausgehändigt, damit Sie Ihr Ziel beim Arbeiten kennen. Der Mehrwert für die Beurteilung der studentischen Arbeiten sowohl für Ausbilder als auch für Studierende durch die Anwendung von Bewertungskriterien soll durch diesen Fragebogen erfasst werden.

Fragebogen zur Anwendung der Bewertungskriterien durch Studenten

Bitte kreuzen Sie die zutreffende Aussage an.

Fragenkomplex 1:

Einfluss der Kriterien auf praktisches Arbeiten:

Die Bewertungskriterien sind korrekt.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Bewertungskriterien sind vollständig.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Bewertungskriterien sind zu umfangreich.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Anwendung der Bewertungskriterien ist klar für mich.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Anwendung der Kriterien erleichtert es mir, das praktische Arbeiten zu strukturieren.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Anwendung der Kriterien erleichtert mir die eigene Vorbereitung.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Anwendung der Kriterien erleichtert es mir, die aufeinanderfolgenden Arbeitsschritte selbstständig durchzuführen.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Anwendung der Kriterien erleichtert es mir, eigene Defizite zu erkennen und gezielt zu üben.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Fragenkomplex 2:

Einfluss der Kriterien auf die Güte der Bewertung:

Die Anwendung von Kriterien erleichtert es mir, Stärken und Schwächen bei den einzelnen Arbeitsschritten isoliert zu erkennen.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Anwendung der Kriterien erleichtert es mir, meine Arbeiten im Laufe der jeweiligen Sequenz zu vergleichen.

trifft gar nicht zu

trifft eher nicht zu

trifft eher zu

trifft voll zu

--	--	--	--

Likert Skala

Fragenkomplex 3:

Einfluss der Kriterien auf die Kommunikation:

Die Anwendung der Kriterien erleichtert es mir, kritische Rückmeldungen besser hinterfragen zu können.

trifft gar nicht zu

trifft eher nicht zu

trifft eher zu

trifft voll zu

--	--	--	--

Likert Skala

Freitext:

1. An der Einführung der vordefinierten Bewertungskriterien für die sechs praktischen Arbeiten finde ich besonders positiv...
2. An der Einführung der vordefinierten Bewertungskriterien für die sechs praktischen Arbeiten sollte Folgendes noch verbessert oder verändert werden...

11.8 Fragebogen Ausbilder

Studententitel: Evaluation objektiver Bewertungskriterien für praktische Studentenarbeiten

In den bisherigen praktischen Kursen, die Sie während Ihres vorklinischen Studienabschnitts besucht haben, wurden Ihre Arbeiten per Augenschein beurteilt. In diesem Kurs wird die Beurteilung an Hand von Bewertungskriterien durchgeführt. Diese Bewertungskriterien werden Ihnen zu Beginn jeder Themensequenz ausgehändigt, damit Sie Ihr Ziel beim Arbeiten kennen. Der Mehrwert für die Beurteilung der studentischen Arbeiten sowohl für Ausbilder als auch für Studierende durch die Anwendung von Bewertungskriterien soll durch diesen Fragebogen erfasst werden.

Fragebogen zur Anwendung der Bewertungskriterien durch die Ausbilder

Bitte kreuzen Sie die zutreffende Aussage an.

Fragenkomplex 1: Einfluss der vordefinierten Bewertungskriterien auf das praktische Arbeiten:

Die Bewertungskriterien sind korrekt.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Bewertungskriterien sind vollständig.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Bewertungskriterien sind zu umfangreich.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Anwendung der Bewertungskriterien ist klar für mich.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Anwendung der Kriterien erleichtert es mir, das praktische Arbeiten strukturiert zu lehren.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Anwendung der Kriterien erleichtert es mir, die Unterschiede der Präparationsformen für Gold, Keramik und Komposit aufzuzeigen.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Durch die Anwendung der Kriterien wird das Wesentliche der praktischen Arbeiten besser gelernt als ohne Kriterien.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Durch Kenntnis der Kriterien kann ich die Lernziele des Kurses besser vermitteln.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Fragenkomplex 2: Einfluss der vordefinierten Bewertungskriterien auf die Güte der Bewertung:

Die Anwendung der Kriterien erleichtert es mir, die Arbeiten eines Studierenden pro Themenkomplex in seiner Entwicklung zu beurteilen.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Anwendung der Kriterien erhöht die forensische Sicherheit meiner Beurteilung, weil die Gründe dokumentiert sind.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Fragenkomplex 3:

Einfluss der vordefinierten Beurteilungskriterien auf die Kommunikation:

Die Anwendung der Kriterien erleichtert es mir, dem Studierenden Feedback zu geben.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Anwendung der Kriterien erleichtert es mir, meine Bewertung gegen Kritik durch den Studierenden zu verteidigen.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Anwendung der Kriterien erleichtert es mir, meine Bewertung gegenüber meinem Vorgesetzten zu verteidigen.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Fragenkomplex 4:

Einfluss der vordefinierten Bewertungskriterien auf den Zeitaufwand:

Die Anwendung der Kriterien erhöht den Zeitaufwand für die Bewertung im Vergleich zur früheren Methode der Bewertung durch Augenschein.

trifft gar nicht zu trifft eher nicht zu trifft eher zu trifft voll zu

--	--	--	--

Likert Skala

Die Anwendung der Kriterien beschleunigt das Erlernen der praktischen Fertigkeiten für die Studierenden.

trifft gar nicht zu

trifft eher nicht zu

trifft eher zu

trifft voll zu

--	--	--	--

Likert Skala

Fragenkomplex 5:

Vergleich der beiden Bewertungsverfahren

Welche der sechs praktischen Arbeiten profitiert in der Qualität am meisten von der Benennung der Bewertungskriterien gegenüber den Studierenden im Vergleich zur Bewertung per Augenschein?

Bitte kreuzen Sie Ihre Beurteilung an.

	Klare Verbesserung	Verbesserung	Neutral	Verschlechterung	Klare Verschlechterung
Präparation für direkte Kompositfüllungen SZ, Klasse II, mod	++	+	O	-	--
Direkte Kompositfüllungen Zahn 26 Klasse II, mod	++	+	O	-	--
Präparation für Teilkrone Gold Zahn 16	++	+	O	-	--
Präparation für Teilkrone Keramik Zahn 26	++	+	O	-	--
Trepanation an einem Molaren (Echtzahn)	++	+	O	-	--
Wurzelkanalfüllung an einem Frontzahn (Echtzahn)	++	+	O	-	--

Freitext:

3. An der Einführung der vordefinierten Bewertungskriterien für die sechs praktischen Arbeiten finde ich besonders positiv...
4. An der Einführung der vordefinierten Bewertungskriterien für die sechs praktischen Arbeiten sollte Folgendes noch verbessert oder verändert werden...

12 Danksagung

Mit dem Abschluss der Dissertation ist es an der Zeit den Menschen zu danken, die mich begleitet und gefördert haben. Mein besonderer Dank gilt dabei zunächst Frau Prof. Dr. Karin Christine Huth für die Möglichkeit, unter ihrer Obhut promovieren zu dürfen. Ich danke, dass sie die wissenschaftliche Betreuung meines Dissertationsprojektes übernommen und dessen Fortgang durch kritische wie anregende Diskussionen mitgestaltet hat.

Mein Dank gilt auch Herrn Prof. Dr. Reinhard Hickel, dass ich die vorliegende Dissertation in seiner Klinik anfertigen durfte.

Bedanken möchte ich mich zudem bei meinen Freunden, meinen Kollegen und Kolleginnen während meiner Zeit an der Ludwig-Maximilians-Universität für die vielfältigen Möglichkeiten der beruflichen Weiterentwicklung. Besonderer Dank gilt Dr. Franz Michael Jakob, der mich während der gesamten Zeit tatkräftig unterstützt hat. Zudem danke ich Dr. Hans-Georg Kirchner, Dr. Michael Nemecek, Herrn M. Kollmuß und den Studierenden, die bei der Umsetzung des Forschungsziels mitgeholfen haben. Danken möchte ich außerdem Dr. Moritz Berger für die kreativen Inspirationen bei der graphischen Umsetzung der Dissertation.

Ich danke meiner lieben Freundin, die mich stets unterstützt und mich in meiner Arbeit bestärkt und angetrieben hat. Hätte sie mich nicht auf diese Art gefördert und motiviert, wäre diese Arbeit in dieser Form nicht möglich gewesen.

Nicht zuletzt meinen Eltern bin ich für die unermüdliche sowie moralische wie tatkräftige Unterstützung nicht nur während meines Studiums sondern auch bereits während meiner Schulzeit zu tiefstem Dank verpflichtet. Sie haben in jeglicher Hinsicht die Grundsteine für meinen Weg gelegt.

13 Lebenslauf

Persönliche Daten

Name Moritz Peter Baumann

Eidesstattliche Versicherung

Baumann, Moritz Peter

Name, Vorname

Ich erkläre hiermit an Eides statt,

dass ich die vorliegende Dissertation mit dem Thema

Evaluation von Bewertungskriterien für praktische Studentarbeiten im Vergleich zu einer Bewertung per Augenschein

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

Bad Homburg, 20.10.2014

Ort, Datum

Unterschrift Doktorandin/Doktorand