

UNIVERSITÄT FREIBURG, SCHWEIZ
ZENTRUM FÜR HOCHSCHULDIDAKTIK

**ÜBERARBEITUNG UND ENTWICKLUNG VON KRITERIEN
ZUR BEWERTUNG VON PRÄSENTATIONEN
IN MASTERSEMINAREN**

Abschlussarbeit zur Erlangung des Diploms
in Hochschuldidaktik und Technologie in der Lehre

Unter der Leitung von Prof. Bernadette Charlier Pasquier

Regina JENSEN

Department für Psychologie | Philosophische Fakultät | Universität Fribourg, Schweiz
2017

Ich erkläre ehrenwörtlich, dass ich meine Abschlussarbeit selbständig und ohne unzulässige fremde Hilfe verfasst habe.

Fribourg, den 17.01.2017

Datum

Regina JENSEN

INHALTSVERZEICHNIS

ZUSAMMENFASSUNG	- 1 -
1. KONTEXT UND ZIELSETZUNG DER VORLIEGENDEN ARBEIT	- 2 -
1.1 FORMALER KONTEXT DER BESCHRIEBENEN EVALUATION	- 2 -
1.2 BEDEUTUNG VON PRÄSENTATIONEN UND DEREN BEWERTUNG IN MASTERSEMINAREN	- 2 -
1.3 PERSÖNLICHE MOTIVATION FÜR DIESE ABSCHLUSSARBEIT	- 3 -
1.4 ZIELSTELLUNG UND AUFBAU DER VORLIEGENDEN ARBEIT	- 4 -
2. DIE BEWERTUNG VON PRÄSENTATIONEN ALS EVALUATION	- 5 -
2.1 ART UND FUNKTION DER EVALUATION	- 5 -
2.2 EVALUATION ALS PROZESS DER URTEILSBILDUNG	- 6 -
2.3 DIE FORMULIERUNG VON VERHALTENSANKERN: METHODE KRITISCHER EREIGNISSE	- 9 -
3. LERNZIELE ALS BASIS FÜR EVALUATIONEN	- 11 -
3.1 LERNZIELE: BEDEUTUNG, DEFINITION UND TAXONOMIE	- 11 -
3.2 LERNZIELE FÜR PRÄSENTATIONEN IN MASTER-SEMINAREN	- 13 -
4. KRITISCHE REFLEXION DER BISHERIGEN BEWERTUNGSSCHEMATA	- 16 -
4.1 VERSION 1 DES BEWERTUNGSSCHEMAS UND DEREN ÜBERARBEITUNG	- 16 -
4.2 VERSION 2 DES BEWERTUNGSSCHEMAS UND IDENTIFIZIERTE PROBLEMFELDER	- 20 -
4.3 ZIELSTELLUNG FÜR DIE ÜBERARBEITUNG DES BEWERTUNGSSCHEMAS	- 22 -
5. METHODISCHES VORGEHEN	- 24 -
5.1 SCHRITT 1: ÜBERARBEITUNG DER BEWERTUNGSKRITERIEN	- 24 -
5.2 SCHRITT 2: ENTWICKLUNG DER VERHALTENSANKER	- 26 -
5.3 SCHRITT 3: ERSTELLUNG DER VERSION 3 DES BEWERTUNGSSCHEMAS	- 28 -
6. ERSTE BEWERTUNG DER VERSION 3 DES BEWERTUNGSSCHEMAS	- 32 -
6.1 VALIDITÄT	- 32 -
6.2 RELIABILITÄT	- 33 -
6.3 RELEVANZ	- 35 -
6.4 QUANTITATIVE VERÄNDERUNG DER BEWERTUNG	- 35 -
6.5 QUALITATIVE EINSCHÄTZUNG DER HANDHABBARKEIT DES BEWERTUNGSSCHEMAS	- 36 -
7. ZUSAMMENFASSUNG	- 38 -
7.1 INHALTLICHES FAZIT	- 38 -
7.2 VORSCHLÄGE FÜR WEITERE VERÄNDERUNGEN	- 38 -
7.3 PERSÖNLICHES FAZIT	- 39 -
8. LITERATURVERZEICHNIS	- 42 -
9. ANHANG	I
ANHANG I: BEWERTUNGSSCHEMA FÜR PRÄSENTATIONEN VERSION 1	II
ANHANG II: BEWERTUNGSSCHEMA FÜR PRÄSENTATIONEN VERSION 2	III
ANHANG III: MATERIALIEN ZUR ERHEBUNG DER VERHALTENSANKER	XII
ANHANG IV: BEWERTUNGSSCHEMA FÜR PRÄSENTATIONEN VERSION 3	XV

ABBILDUNGSVERZEICHNIS

ABBILDUNG 1: STRUKTUR DER TAXONOMIE KOGNITIVER LERNZIELE VON BLOOM NACH KRATHWOHL (2010); ÜBERSETZUNG DURCH AUTORIN	- 12 -
ABBILDUNG 2: AUSSCHNITT AUS DEM BEWERTUNGSSHEMA FÜR PRÄSENTATIONEN V1	- 16 -
ABBILDUNG 3: AUSSCHNITT AUS DEM BEWERTUNGSBOGEN DES BEWERTUNGSSCHEMAS FÜR PRÄSENTATIONEN V2	- 18 -
ABBILDUNG 4: AUSSCHNITT AUS DEN VERHALTENSANKERN DES BEWERTUNGSSCHEMAS FÜR PRÄSENTATIONEN V2	- 19 -
ABBILDUNG 5: AUSSCHNITT AUS DEM BEWERTUNGSBOGEN DES BEWERTUNGSSCHEMAS FÜR PRÄSENTATIONEN V3	- 28 -
ABBILDUNG 6: AUSSCHNITT AUS DEN VERHALTENSANKERN DES BEWERTUNGSSCHEMAS FÜR PRÄSENTATIONEN V3	- 29 -

TABELLENVERZEICHNIS

TABELLE 1: AUS DEN FORMULIERTEN LERNZIELEN ABGELEITETE BEWERTUNGSKRITERIEN ZUM INHALT	- 24 -
TABELLE 2: AUS DEN FORMULIERTEN LERNZIELEN ABGELEITETE BEWERTUNGSKRITERIEN ZUR PRÄSENTATIONSKOMPETENZ	- 25 -
TABELLE 3: ZUORDNUNG MEDIAN - NOTE	- 30 -

ZUSAMMENFASSUNG

Das Vermitteln von Inhalten stellt in vielen Berufen eine alltägliche Herausforderung dar. Um Studierende darauf vorzubereiten, werden in Masterseminaren immer häufiger Präsentationen von Studierenden erwartet, deren Bewertung in die Note des Seminars mit einfließt. Daher werden Evaluationskriterien benötigt, die aus Lernzielen abgeleitet sind und den üblichen Anforderungen bezüglich der Gütekriterien entsprechen. Das Ziel der vorliegenden Arbeit ist es, die bestehenden Evaluationskriterien für Präsentationen dahingehend zu überarbeiten. Dazu wurden zunächst Lernziele formuliert, aus denen Bewertungskriterien abgeleitet wurden. Für diese Kriterien wurden von mehreren Experten mithilfe der Methode kritischer Ereignisse Verhaltensanker definiert. Im Herbstsemester 2016 kamen diese überarbeiteten Evaluationskriterien erstmals zum Einsatz. Die kritische Reflexion dieser ersten Anwendung weist auf Verbesserungen gegenüber der vorherigen Version hin. Gleichzeitig ergibt sich aber auch bei dieser neuen Version noch in manchen Punkten Veränderungsbedarf, insbesondere hinsichtlich der Erhöhung der Interrater-Reliabilität sowie in Bezug auf die Vermittlung der Kriterien an die Studierenden.

1. KONTEXT UND ZIELSETZUNG DER VORLIEGENDEN ARBEIT

1.1 *FORMALER KONTEXT DER BESCHRIEBENEN EVALUATION*

Die Autorin dieser Arbeit ist als Diplomassistentin im Department für Psychologie (Philosophische Fakultät) der Universität Fribourg angestellt. Der überwiegende Anteil der Lehrtätigkeit findet im Masterstudiengang Psychologie statt. Im Ko-Teaching mit einer Professorin werden Masterseminare der Option „Arbeits- und Organisationspsychologie“ gehalten. Im Schnitt werden diese Seminare von ungefähr 10 Studierenden besucht. In diesen Kontext sind die vorliegende Arbeit sowie die gemachten Erfahrungen und vorgenommenen Einschätzungen einzuordnen.

1.2 *Bedeutung von Präsentationen und deren Bewertung in Masterseminaren*

„Abschlüsse der zweiten Stufe werden an Studierende verliehen, die [...] ihre Schlussfolgerungen und das Wissen und die Prinzipien, die ihnen zugrunde liegen, klar und eindeutig kommunizieren können, sowohl an Experten wie auch an Laien.“ (Schweizerische Universitätskonferenz, 2011, S. 11). Diese Anforderung wird im Qualifikationsrahmen für den schweizerischen Hochschulbereich für Studierende auf Masterniveau unter dem Begriff der kommunikativen Fähigkeit formuliert. Hier wird deutlich, dass – wie es auch Race, Brown und Smith (2005) in ihrem Buch schreiben – in verschiedenen potenziellen Berufsfeldern von Masterabsolventen das Halten von guten Präsentationen eine wichtige berufliche Kompetenz darstellt und diese Kompetenz an der Universität entwickelt werden sollte (van Ginkel, Gulikers, Biemans & Mulder, 2016). Für Studierende der philosophischen Fakultät der Universität Fribourg wird nach dem Master-Reglement (Philosophische Fakultät der Universität Freiburg (Schweiz), 2006) zudem am Ende des Masterstudiums eine Masterverteidigung inklusive der Präsentation der eigenen Masterarbeit verlangt.

Um Studierende bestmöglich auf die beruflichen Anforderungen im Allgemeinen, aber auch auf die Anforderungen der Masterverteidigung im Speziellen vorzubereiten, werden immer öfter Präsentationen von Studierenden in Masterseminare eingebunden. Häufig werden diese auch bewertet und zählen mit in die Gesamtnote des Seminars. Zwingend stellt sich damit die Frage nach Bewertungskriterien, anhand derer die Benotung vorgenommen werden kann.

1.3 PERSÖNLICHE MOTIVATION FÜR DIESE ABSCHLUSSARBEIT

Neben dieser eher „objektiven“ Notwendigkeit, Bewertungskriterien der Benotung von Präsentationen zugrunde zu legen, gab es auch verschiedene persönliche Anlässe für das Thema und die Zielstellung der vorliegenden Abschlussarbeit.

Ein Aspekt geht auf das eigene Studium zurück. Grundsätzlich waren Erfahrungen mit den Dozenten während meines Studiums ein starker Motivator für die Did@ctic-Ausbildung. Auch hinsichtlich der Benotung von Präsentationen habe ich verschiedene Erfahrungen gemacht, die mich für diese Arbeit motiviert haben. Während des Bachelor- und Master-Studiums habe ich über 20 Präsentationen in Seminaren gehalten. Bei manchen Dozenten wurde dabei deutlich, dass den Noten ein klares Anforderungsprofil und Benotungsschema zugrunde lag. Diese Dozenten kommunizierten auch im Vorhinein ihre Erwartungen, sodass man sich bei der Erstellung der Präsentationen daran orientieren konnte. Bei anderen Dozenten wiederum hatte man den Eindruck, bei der Note handele es sich um eine globale Einschätzung, die nicht nur stark durch die Sympathie/Antipathie für die Studierenden zusammen hing, sondern auch mit der aktuellen Stimmungslage der Dozenten. Auch wenn ich natürlich nicht sicher sein kann, dass es wirklich so war, empfand ich diese Benotung als höchst frustrierend und mir war klar, dass ich dies einmal anders machen wollte, sollte ich selber in die Situation kommen, Noten für nicht-schriftliche Arbeiten zu vergeben.

Als ich dann in der Situation war, selber Noten zu vergeben, merkte ich, wie schwierig das ist. Nach wie vor empfinde ich das Vergeben von objektiven und fairen Noten – sowohl für schriftliche als auch für mündliche Leistungsnachweise - als schwierigste Aufgabe im Bereich der Lehre. Das Basieren der Noten auf einem fundierten Bewertungsschema sehe ich als Möglichkeit, diese Unsicherheit etwas zu reduzieren. Die bisherigen Versionen der Bewertungsschemata empfand ich dabei aber aus verschiedenen Gründen als unbefriedigend (für eine detaillierte Beschreibung dieser Gründe s. Kapitel 4), was für mich Motivation war, eine grundlegende Überarbeitung der Kriterien anzustreben.

Ein weiterer Anlass für das Thema dieser Abschlussarbeit war, dass ich im Verlauf der Did@ctic-Ausbildung lernte, welche hohe Relevanz die Verknüpfung von Evaluationen mit definierten Lernzielen hat. Gleichzeitig realisierte ich, dass mir die Lernziele von Präsentationen in Seminaren nicht klar waren und somit auch die Evaluation nicht auf Lernzielen basierte. Dies zu ändern war ein weiterer Motivator für diese Abschlussarbeit.

Und nicht zuletzt war mir wichtig, dass das Projekt, das ich im Rahmen der Abschlussarbeit durchführen würde, ein Produkt liefert, welches eine praktische Relevanz für den Arbeitsalltag von Dozierenden hat – nicht nur für mich, sondern vielleicht auch für andere Dozierende. Da zumindest im Bereich der Psychologie in mehreren Masterseminaren Präsentationen durch Studierende verlangt werden, sah ich bei dem gewählten Thema die Möglichkeit, dieses Ziel zu erreichen.

Aus der „objektiven“ Notwendigkeit eines Bewertungsschemas sowie den verschiedenen persönlichen Anlässen ergibt sich die Zielstellung dieser Arbeit.

1.4 ZIELSTELLUNG UND AUFBAU DER VORLIEGENDEN ARBEIT

Das übergreifende Ziel der vorliegenden Arbeit besteht in der Weiterentwicklung des bestehenden Bewertungsschemas für Präsentationen in Masterseminaren. Dabei werden vier Unterziele verfolgt:

- (1) Eine Fundierung der Bewertungskriterien auf formulierten Lernzielen
- (2) Eine Kalibrierung der Verhaltensanker zur Bewertung der einzelnen Kriterien hinsichtlich ihres Anspruchsniveaus
- (3) Eine Verallgemeinerung der Kriterien, um eine breitere Anwendung des Bewertungsschemas zu ermöglichen
- (4) Eine Anpassung der Notenberechnung gemäss der Möglichkeiten einer Ordinalskala

Zunächst wird im Folgenden eine theoretische Einordnung der Bewertung von Präsentationen als Evaluation vorgenommen (Kapitel 2). Anschliessend wird eine Taxonomie von Lernzielen vorgestellt, aus der Lernziele für das Halten von Präsentationen in Masterseminaren abgeleitet werden (Kapitel 3). Nach einer kritischen Reflexion der bisherigen Versionen der Bewertungskriterien wird der Veränderungsbedarf für die neue Version formuliert (Kapitel 4). Nachdem das Vorgehen der Überarbeitung beschrieben wurde (Kapitel 5), werden erste Ergebnisse zur Anwendung des überarbeiteten Bewertungsschemas berichtet (Kapitel 6). Die Arbeit schliesst mit einem Fazit sowie einem Ausblick für weitere mögliche Veränderungen und Anpassungen (Kapitel 7).

2. DIE BEWERTUNG VON PRÄSENTATIONEN ALS EVALUATION

2.1 ART UND FUNKTION DER EVALUATION

Bei einer Lernevaluation (worauf Bezug genommen wird, wenn im Folgenden von Evaluation gesprochen wird) handelt es sich um einen Prozess, durch den die Entwicklung von Kompetenzen bei Studierenden (oder anderen Personengruppen) beurteilt wird (Charlier, 2015). Typischerweise werden zwei **Arten von Evaluation** unterschieden: formative und summative Evaluation (Bloom, Hastings, & Madaus, 1971). Eine formative Evaluation wird während des Lernprozesses durchgeführt und hat das Ziel, den Lernprozess zu verbessern und zu erleichtern („assessment for learning“; Bloxham & Boyd, 2007, S. 22). Dabei wird Feedback als eine zentrale Komponente betrachtet (Bell & Cowie, 2001). Mithilfe einer summativen Evaluation soll dagegen ermittelt werden, in welchem Ausmass die Studierenden die vorab formulierten Lernziele erreicht haben (Yorke, 2003). Diese Art der Evaluation wird auch als „assessment of learning“ bezeichnet (Bloxham & Boyd, 2007, S. 22) und ist meist mit der Vergabe einer Note verbunden (Bloom et al., 1971). Allerdings sind formative und summative Evaluationen nicht immer trennscharf voneinander abzugrenzen. Bei manchen Aufgaben handelt es sich sowohl um eine formative als auch eine summative Evaluation: Studierende erhalten ein Feedback, was einen Lernfortschritt bewirken soll. Gleichzeitig wird die Leistung aber auch benotet und fließt in die Gesamtnote des Moduls ein (Yorke, 2003). Diese wenig trennscharfe Abgrenzung trifft auch auf die in dieser Arbeit beschriebene Evaluation zu.

Stellt man die Evaluation einer Präsentation im Rahmen eines Masterseminars in den Kontext des gesamten Masterstudiums, handelt es sich um eine *formative Evaluation*. Die Studierenden erhalten wiederholt Feedback zu ihrer Präsentation. Dieses Feedback soll helfen, die Studierenden auf die Anforderungen in der Masterverteidigung, aber auch im Berufsleben vorzubereiten. Da Studierende immer wieder Präsentationen halten, handelt es sich, spezifischer ausgedrückt, um eine *formative Zwischen-Evaluation* („évaluation formative d'étape“, Cardinet, 1986). Betrachtet man die Evaluation einer Präsentation dagegen im Rahmen eines einzelnen Seminars, dann ist sie als *summative Evaluation* zu interpretieren. Es wird evaluiert, inwieweit die Studierenden die Anforderungen des Seminars erfüllen und die Benotung fließt in die Kursnote und somit in die Modulnote mit ein.

Eng verknüpft mit der Art ist auch die **Funktion der Evaluation**. Hinsichtlich der summativen Evaluation kommt der Präsentation im Seminar eine *zertifikative Funktion* zu. Ziel ist es zu überprüfen, ob die Studierenden in der Lage sind, ein bestimmtes Inhaltsgebiet des Seminars so

aufzubereiten, dass sie es korrekt, zielgerichtet und zielgruppenspezifisch präsentieren und in den Gesamtkontext des Seminars einbetten können. In Bezug auf die formative Zwischen-Evaluation wird mit der Aufgabe einer Präsentation dagegen eine *diagnostische Funktion* verfolgt. Die Ausgestaltung ist dabei eher individuell als kollektiv. Es wird eher das Verbesserungspotenzial einzelner Studierenden hinsichtlich der Präsentationskompetenz identifiziert, als dass kollektive Schwierigkeiten im Lernprozess definiert werden. Nach der Beschreibung von Charlier (2015b) sollten formative Evaluationen nicht über den Erfolg der Lernsequenz entscheiden. Das Ergebnis dieser Evaluation sollte daher nicht in die Note mit einfließen. Ein zentrales Merkmal von formativer Evaluation besteht zudem darin, dass Fehler erlaubt sind und aus diesen gelernt werden kann. Dieser Grundsatz muss bei den Überlegungen zu passenden Bewertungskriterien berücksichtigt werden.

2.2 EVALUATION ALS PROZESS DER URTEILSBILDUNG

Bei der Benotung einer Präsentation handelt es sich um eine Leistungsbeurteilung durch Beobachtung. Es liegen keine objektiven Leistungsindikatoren vor, die Einschätzung erfolgt vielmehr durch Beobachter und ist somit subjektiv geprägt. Eine Möglichkeit, eine solche Einschätzung vorzunehmen ist die *freie Eindruckschilderung*. Insbesondere für Feedback und Verhaltensentwicklung ist sie hilfreich (Marcus & Schuler, 2006), könnte also für eine formative Evaluation eine gute Basis darstellen. Gleichzeitig sind mit einer freien Eindruckschilderung jedoch auch mehrere kritische Punkte verbunden, die insbesondere für die summative Evaluation problematisch sein können. Diese sollen im Folgenden kurz geschildert werden.

(1) Fehlender Kriteriumsbezug der Beurteilung

Das Charakteristische einer freien Eindruckschilderung liegt darin, dass kein vorgegebenes Schema zur Leistungsbeurteilung verwendet wird. Dies ist mit einer grossen Flexibilität in der Anwendung verbunden (Marcus & Schuler 2006). Gleichzeitig bedeutet es aber auch, dass der Beurteilung keine Kriterien zugrunde liegen. In der Hochschuldidaktik nimmt der Bezug zu definierten Kriterien jedoch eine entscheidende Rolle ein, da hier die Verbindung der Evaluation zu den Lernzielen hergestellt wird (Charlier, 2015b). Eine Fundierung der Bewertung von Präsentationen auf Lernzielen ist daher von grosser Bedeutung.

(2) Verzerrung der Beurteilung durch Urteilstendenzen

Leistungsbeurteilungen durch Beobachtungen sind anfällig für sogenannte Urteilstendenzen. Die wichtigsten Urteilstendenzen sind die Mittelwerts-, Streuungs- und Korrelationstendenzen. Unter *Mittelwertstendenz* versteht man eine Verlagerung der mittleren Beurteilung nach oben oder unten, d.h. die Beurteiler weisen eine Mildetendenz

(Verlagerung nach oben) oder eine Strengetendenz (Verlagerung nach unten) auf. Beispiele für die *Streuungstendenz* sind Streuungseinschränkungen (die Beurteiler nutzen nicht die ganze Skala zur Beurteilung) und Tendenz zur Mitte (die Beurteiler vermeiden die Verwendung der extremen Skalenpunkte). Die meistdiskutierte Korrelationstendenz ist der Halo-Effekt. Mit diesem Effekt wird beschrieben, dass in der Leistungsbeurteilung ein Merkmal des Beurteilten die Beurteilungen anderer Merkmale überstrahlt, dass also das Gesamturteil sehr stark von einem einzelnen (als positiv oder negativ betrachteten) Merkmal geprägt wird (Blickle, 2014). Zu beachten ist, dass diese Tendenzen nicht zwingend Fehler in der Leistungsbeurteilung darstellen. So müssen geringe Urteilsstreuungen keine Verzerrung im Sinne der Streuungstendenz darstellen, sondern können auch tatsächliche Leistungshomogenität abbilden (Marcus & Schuler, 2006). Gleichzeitig können diese Tendenzen jedoch insbesondere bei ungeübten Beobachtern zu einer Verzerrung der Urteile führen.

(3) Fehlende Möglichkeit einer vergleichende Beurteilung

Verschiedene Autoren weisen darauf hin, dass eine freie Eindrucksschilderung für die Leistungsbeurteilung nur dann sinnvoll ist, wenn keine interindividuellen Vergleiche der Leistungsbeurteilung erfolgen sollen (Blickle, 2014; Marcus & Schuler, 2006; Moser & Schuler, 2014). Dies liegt insbesondere darin begründet, dass keine Skalen verwendet werden, sondern der Beurteiler eine individuelle Charakterisierung der zu beurteilenden Person vornimmt (Blickle, 2014). Diese kann in Abhängigkeit der beobachteten Stärken und Schwächen der zu beurteilenden Personen sehr unterschiedlich ausfallen, sodass ein interindividueller Vergleich der Urteile nicht möglich ist. Insbesondere eine summative Evaluation beinhaltet jedoch zwangsläufig durch die Vergabe einer Note einen interindividuellen Vergleich und sollte daher nicht idiosynkratisch in Abhängigkeit der Stärken und Schwächen einer Person geprägt sein.

(4) Geringe Beurteiler-Übereinstimmung

Durch die hohe Flexibilität und Individualisierung der freien Eindrucksschilderung ist eine geringe Beobachter-Übereinstimmung zu konstatieren. Das bedeutet, dass die Interrater-Reliabilität stark eingeschränkt ist (Blickle, 2014). Im Kontext der Hochschuldidaktik ist die Interrater-Reliabilität jedoch ein entscheidendes Gütekriterium für Evaluationsverfahren (Charlier, 2015b).

Eine Möglichkeit, diesen geschilderten Problemen entgegenzuwirken, ist die Verwendung von *Einstufungsverfahren*. Nach Blickle (2014) stellen Einstufungen „kriteriumsbezogene Beurteilungen von Personen auf Ratingskalen ohne Bezug zu anderen Personen dar“ (S. 279). Aus dieser Definition geht hervor, dass der Kriteriumsbezug ein zentrales Merkmal der Einstufungsverfahren ist. Ein weiteres Charakteristikum von Einstufungsverfahren ist die Verwendung von Skalenverankerungen, d.h. die Kennzeichnung einzelner Skalenpunkte mithilfe von Zahlenwerten, Adjektiven oder Verhaltensbeschreibungen. Sprachliche Verankerungen sollen helfen, Urteilstendenzen zu reduzieren und die Verwendung der Skala durch verschiedene Beurteiler zu vereinheitlichen resp. die Interrater-Reliabilität zu erhöhen (Blickle, 2014). Durch die stärkere Standardisierung der Beurteilung werden die Urteile zudem über verschiedene zu beurteilende Personen hinweg vergleichbarer (Marcus & Schuler, 2006). Ein weiterer Vorteil der Einstufungsverfahren ist die Trennung der Beurteilung von Eigenschaften und Verhalten. Diese Vorteile machen die Einstufungsverfahren zu einem geeigneten Instrument für die Beurteilung von Präsentationen in der Hochschuldidaktik.

Es existieren verschiedene Formen von Einstufungsverfahren. Die einfachste Form ist die *grafische Einstufungsskala*. In diesem Fall wird das relevante Merkmal auf einer einfachen Skala beurteilt. Manchmal werden die Skalenpunkte ausschliesslich mit einfachen Bewertungen versehen (z.B. unakzeptable – aussergewöhnlich, Lunenburg, 2012). Sind die Skalenpunkte als Verhaltensweisen formuliert (z.B. „ist im Umgang mit anderen befangen und ängstlich“ – „findet selbst zu schwierigen Menschen leicht Kontakt“ als Skalenpunkte für die Dimension „Kontaktfähigkeit“, Marcus & Schuler, 2006, S. 448), dann werden diese in der Regel ohne grossen Konstruktionsaufwand formuliert.

Bei der *Verhaltensbeobachtungsskala* (Behavior Observation Scale, BOS) werden dagegen zu einem Verhaltenskriterium mehrere Verhaltensweisen gesammelt. Diese werden als Item formuliert und entweder dichotom (Latham & Wexley, 1977) oder auf einer Likert-Skala (z.B. „fast nie“ – „fast immer“, Marcus & Schuler, 2006, S. 448) hinsichtlich ihres Auftretens beurteilt. Die Aggregation der Häufigkeitseinschätzung der verschiedenen einzelnen Verhaltensweisen ergibt eine Beurteilung für ein Verhaltenskriterium. Die einzelnen Verhaltensweisen werden meist mithilfe der *Methode kritischer Ereignisse* gesammelt.

Das nach Marcus & Schuler (2006) am besten erforschte Einstufungsverfahren ist die *verhaltensverankerte Einstufungsskala* (Behaviorally Anchored Rating Scale, BARS). Bei diesem Verfahren werden den Skalenpunkten, anhand derer die Leistung beurteilt werden soll, mithilfe der *Methode kritischer Ereignisse* erhobene beobachtbare Verhaltensweisen zugeordnet. Der Konstruktionsprozess der Verhaltensanker ist meist mehrstufig und komplex und wird in

verschiedenen Varianten umgesetzt (Bernardin & Smith, 1981). Die verhaltensverankerte Einstufungsskala geht über eine reine Häufigkeitseinschätzung von Verhaltensweisen hinaus und ist somit für die Beurteilung von Präsentationen besser geeignet als die Verhaltensbeobachtungsskala. Die in dieser Arbeit überarbeitete Version der Beurteilungskriterien für Präsentationen wird als verhaltensverankerte Einstufungsskala konzipiert. Der Konstruktionsprozess wird dabei jedoch auf die *Methode kritischer Ereignisse* reduziert, die im nächsten Abschnitt beschrieben werden soll.

2.3 DIE FORMULIERUNG VON VERHALTENSANKERN: METHODE KRITISCHER EREIGNISSE

Die *Methode kritischer Ereignisse* (critical incident technique, CIT) geht auf Flanagan (1954) zurück und ist eine Methode zur Sammlung erfolgskritischer, beobachtbarer Verhaltensweisen (Blickle, 2014b; Schuler, 2006). Ihre inhaltlichen Wurzeln hat diese Technik in der Arbeits- und Organisationspsychologie, heute wird sie jedoch in vielen verschiedenen inhaltlichen Bereichen angewendet – unter anderem auch im Bereich der Lehre (Butterfield, Borgen, Amundson & Maglio, 2005). Dieser vielfältige Einsatz liegt sicherlich auch darin begründet, dass Flanagan (1954) selbst schreibt, die Methode bestehe nicht aus einem einzigen, rigiden Set von Regeln, wie sie anzuwenden sei, sondern solle vielmehr als flexibles Set von Prinzipien verstanden werden, dass je nach Situation angepasst und verändert werden könne. Gleichzeitig definiert er fünf Haupt-Bestandteile der Methode, die immer berücksichtigt werden sollten. Diese sollen im Folgenden kurz beschrieben werden.

Der erste Schritt besteht in der genauen Definition des Ziels der Aktivität, für die erfolgskritische Verhaltensweisen gesammelt werden soll. Butterfield et al. (2005) präzisieren diesen ersten Schritt dahingehend, dass es nicht nur darum gehe, das Ziel der Aktivität zu definieren, sondern auch, was von der Person, die die Aktivität ausführt, im Detail erwartet wird. Dieser Punkt ist entscheidend bei der Anwendung der Methode kritischer Ereignisse für die Formulierung von Verhaltensanker zur Beurteilung von Präsentationen und wird in der Formulierung der Lernziele sowie der zu beurteilenden Kriterien aufgegriffen werden (s. Kapitel 3.2 und 5.1). *Der zweite zentrale Bestandteil* der Methode besteht in der Präzisierung der Instruktion für die Personen, mithilfe derer die erfolgskritischen Verhaltensweisen definiert werden sollen (von Flanagan als „Beobachter“ bezeichnet, jedoch sind sie nicht zu verwechseln mit den Beobachtern, die die Beurteilung der Präsentation vornehmen). Insbesondere ist hier auch die Personengruppe zu spezifizieren, von denen die Verhaltensweisen gesammelt werden sollen – in dieser Arbeit sind es die Masterstudierenden. *Die Schritte drei bis fünf* sind die klassischen Schritte einer qualitativen oder empirischen Arbeit und

beinhalten die Datenerhebung, die Datenanalyse sowie die Interpretation und den Bericht der Ergebnisse.

Blickle (2014b) präzisiert, dass es bei der Methode kritischer Ereignisse nicht nur wichtig sei, dass kritische Verhalten so genau wie möglich zu beschreiben, sondern ein Augenmerk auch darauf gelegt werden sollte, was an dem Verhalten effektiv oder ineffektiv war und welche Resultate aus dem Verhalten entstanden sind. Der starke Bezug der mithilfe dieser Methode gesammelten Verhaltensweise zu ihrer Effektivität und erfolgskritischen Bedeutung macht sie so nützlich für die Verwendung als Verhaltensanker einer verhaltensverankerten Einstufungsskala. Die Verhaltensweisen können den Skalenpunkten zugeordnet werden und stellen somit eine Basis für die darauf aufbauende Beurteilung dar. Die genaue Anwendung der Methode kritischer Ereignisse in dem in dieser Arbeit beschriebenen Kontext wird in Kapitel 5.2 beschrieben.

3. LERNZIELE ALS BASIS FÜR EVALUATIONEN

3.1 LERNZIELE: BEDEUTUNG, DEFINITION UND TAXONOMIE

„Wer nicht genau weiss, wohin er will, braucht sich nicht zu wundern, wenn er ganz woanders ankommt.“ Mit diesem Zitat von Robert F. Mager (1965, zitiert nach Wyrsh, 2014) verdeutlicht A. Wyrsh in den Unterlagen zum Basismodul A der Hochschuldidaktik-Ausbildung die Bedeutung von Lernzielen. Aus diesem Zitat gehen zwei wesentliche Merkmale von Lernzielen hervor: Zum einen ist es für die Formulierung von Lernzielen als Dozent notwendig zu wissen, wohin man möchte – in Bezug auf die Hochschuldidaktik ist es also notwendig zu wissen und zu definieren, was den Studierenden vermittelt werden sollen. Zum anderen haben Lernziele die Funktion, den Weg zu weisen – also sowohl den Studierenden als auch den Dozierenden explizit zu verdeutlichen, was am Ende der Einheit gelernt worden sein soll. Beide Aspekte kristallisieren sich in der Definition von Lernzielen von Krathwohl (2010): „an objective defines what students are expected to learn“ (S. 213). Eine entscheidende Rolle spielt diese definierte Lernerwartung auch im Kontext von Evaluationen: Ohne die Definition von Lernzielen können keine Kriterien formuliert werden, anhand derer die Studierenden evaluiert werden sollen. Lernziele stellen also die notwendige Basis für Bewertungskriterien dar. Döring (2008) beschreibt, dass die Einführung von Lernzielen in die Didaktik ihren Ursprung im Behaviorismus hat. Mit diesem Ursprung verbunden ist die Formulierung von Lernzielen in Kategorien von direkt beobachtbarem Verhalten. Dazu passend definiert Wyrsh (2014) Lernziele als „das ausreichend geltende Verhalten“ (S. 36). Insbesondere für Evaluationen, die durch direkte Beobachtung erfolgen, wie dies bei der Evaluation von Präsentationen der Fall ist (s. Kapitel 2.2) stellen Lernziele demnach eine sinnvolle Basis dar. Im folgenden Abschnitt soll deshalb zunächst allgemein eine Taxonomie von Lernzielen vorgestellt werden, aus der im anschliessenden Kapitel 3.2 Lernziele für Präsentationen entwickelt werden sollen. Diese Lernziele bilden die Basis für die Formulierung der überarbeiteten Bewertungskriterien (s. Kapitel 5.1).

Eine weit verbreitete Taxonomie von Lernzielen ist die Taxonomie von Bloom (erstmal veröffentlicht von Bloom, Engelhart, Furst, Hill & Krathwohl, 1956, hier zitiert nach Bloom et al., 1971). Bloom und seine Kollegen definieren drei Bereiche der Taxonomie: kognitive Lernziele, affektive Lernziele und psychomotorische Lernziele. *Kognitive Lernziele* legen den Fokus auf Gelerntes und auf das Bewältigen intellektueller Aufgaben. *Affektive Lernziele* beziehen sich dagegen auf die Entwicklung von Einstellungen, Haltungen und Werten. *Psychomotorische Lernziele* fokussieren motorische Fertigkeiten oder den Umgang mit Gegenständen (Krathwohl, Bloom, & Masia, 1975). Die grösste Verbreitung haben – insbesondere auch im universitären Kontext - kognitive Lernziele (Bloom et al.,

1971; Döring, 2008; Krathwohl et al, 1975). Auf diese beschränken sich auch die folgenden Ausführungen.

Die Taxonomie der kognitiven Lernziele umfasst sechs Bereiche: Wissen, Verständnis, Anwendung, Analyse, Synthese und Bewertung. Jeder dieser Bereiche – mit Ausnahme des Bereichs Anwendung – wird durch verschiedene Unterbereiche definiert. Eine Übersicht über die Bereiche mit den

1 Wissen
1.10 Wissen von Spezifischem
1.11 Wissen von Terminologien
1.12 Wissen spezifischer Fakten
1.20 Wissen von Vorgehensweisen und Mitteln sich mit Spezifischem zu beschäftigen
1.21 Wissen über Konventionen
1.22 Wissen von Trends und Entwicklungen
1.23 Wissen von Klassifikationen und Kategorien
1.24 Wissen von Kriterien
1.25 Methodisches Wissen
1.30 Wissen von Allgemeinem und Abstraktionen eines Gebiets
1.31 Wissen von Prinzipien und Generalisierungen
1.32 Wissen von Theorien und Strukturen
2 Verständnis
2.10 Übertragung
2.20 Interpretation
2.30 Weiterführung
3 Anwendung
4 Analyse
4.10 Analyse von Elementen
4.20 Analyse von Zusammenhängen
4.30 Analyse von Organisations-Prinzipien
5 Synthese
5.10 Erstellen einer eigenen Kommunikation
5.20 Erstellen eines Plans oder eines Sets von Vorgehensweisen
5.30 Ableitung eines Sets abstrakter Beziehungen
6 Bewertung
6.10 Evaluation auf Basis interner Evidenzen
6.20 Bewertung auf Basis externer Kriterien

Abbildung 1: Struktur der Taxonomie kognitiver Lernziele von Bloom nach Krathwohl (2010); Übersetzung durch Autorin

Unterbereichen findet sich in Abbildung 1. Der Aufbau der Taxonomie ist hierarchisch, das bedeutet, dass die Beherrschung einer unteren Kategorie die Voraussetzung für die Beherrschung der darüber liegenden Kategorien darstellt (Krathwohl, 2010).

Krathwohl (2010) beschreibt verschiedene Ziele der Taxonomie von Bloom und Kollegen. Unter anderem erfolgte die Entwicklung der Taxonomie in der Absicht, eine gemeinsame Sprache für Lernziele über verschiedene Personengruppen, Inhalte und Lernstufen hinweg zu entwickeln. Des Weiteren wurde das Ziel verfolgt, eine Kongruenz von Lernzielen in einer Arbeitseinheit, innerhalb eines Curriculums und über verschiedene Kurse hinweg herzustellen. Diese beiden Ziele gehen Hand in Hand mit dem Ziel dieser Arbeit, die Bewertungskriterien für Präsentationen so zu verallgemeinern, dass eine breite Anwendung möglich wird. Auch deshalb eignet sich die Taxonomie von Bloom und Kollegen für die in dieser Arbeit behandelte Ableitung der Lernziele und damit verbunden für die Formulierung der Bewertungskriterien.

3.2 LERNZIELE FÜR PRÄSENTATIONEN IN MASTER-SEMINAREN

Die zentrale Frage ist, wie sich diese von Bloom und Kollegen (1971) formulierten Kategorien von Lernzielen auf Präsentationen in Master-Seminaren anwenden lassen. Dazu soll zunächst noch einmal Bezug genommen werden zum Qualifikationsrahmen für den schweizerischen Hochschulbereich (Schweizerische Universitätskonferenz, 2011). Für Studierende auf Masterniveau werden darin 5 Voraussetzungen definiert, um einen Masterabschluss zu erlangen (sog. „Deskriptoren“): (1) Wissen und Verstehen, (2) Anwendung von Wissen und Verstehen, (3) Urteilen, (4) Kommunikative Fertigkeiten und (5) Selbstlernfähigkeit.

Vergleicht man diese Deskriptoren mit der Taxonomie von Bloom und Kollegen, so wird eine grosse Überschneidung deutlich: Die Kategorien *Wissen*, *Verständnis*, *Anwendung* und *Bewertung* von Bloom und Kollegen finden sich direkt in den Deskriptoren wieder. Aus der Tatsache, dass *Bewertung* die höchste Kategorie in der Taxonomie darstellt und die einzelnen Kategorien, wie in Kapitel 3.1 beschrieben, hierarchisch aufeinander aufbauen, ergibt sich, dass indirekt auch die Kategorien *Analyse* und *Synthese* Voraussetzungen für das Erlangen eines Masterabschlusses sind. Die beiden Deskriptoren (4) Kommunikative Fertigkeit und (5) Selbstlernfähigkeit gehen über die Kategorien der Taxonomie von Bloom und Kollegen hinaus.

In Bezug auf die unterschiedlichen Arten von Evaluationen stellen die Deskriptoren die Basis für eine *summative Evaluation* am Ende des Masterstudiums dar. Gleichzeitig schliessen die Studierenden jedoch auch einzelne Seminare mit einer Note und somit einer summativen Evaluation ab. Da Inhalte

im Masterstudium häufig nur in einem Seminar behandelt werden, ist es wichtig, dass die Studierenden die inhaltsbezogenen Lernziele – in Bezug auf die Deskriptoren also die Aspekte (1)-(4) – für diese spezifischen Inhalte schon am Ende eines Seminars erreichen. Kommunikative Fertigkeiten und Selbstlernfähigkeit dagegen können und sollten über die gesamte Dauer des Masterstudiums entwickelt werden. Gleichzeitig ist hier jedoch zu beachten, dass diese beiden Deskriptoren mit einem geringeren Anspruchsniveau auch in den Deskriptoren für die Bachelor-Stufe zu finden sind. Da ein Bachelor Voraussetzung für die Zulassung zum Master ist, kann im Masterstudium eine gewisse Fertigkeit in diesen beiden Bereichen vorausgesetzt werden.

In Bezug auf Präsentationen, die eine unter anderem eine summative Funktion in Masterseminaren erfüllen (s. Kapitel 2.1) bedeutet dies, dass Lernziele in den 6 von Bloom und Kollegen formulierten Kategorien evaluiert werden können. Hinsichtlich der Aspekte der kommunikativen Fertigkeit und der Selbstlernfähigkeit aus dem Qualifikationsrahmen für den schweizerischen Hochschulbereich (Schweizerische Universitätskonferenz, 2011) sollten die Anforderungen nicht über die zum Ende des Bachelor-Studiums verlangten Anforderungen hinausgehen.

Aufbauend auf diesen Überlegungen sowie der Taxonomie von Bloom und Kollegen (1971) und den Deskriptoren aus dem Qualifikationsrahmen für den schweizerischen Hochschulbereich (Schweizerische Universitätskonferenz, 2011) lassen sich für den inhaltlichen Bereich die folgenden Lernziele formulieren, deren Erreichen anhand der Evaluation von Präsentationen beurteilt werden kann:

1 Wissen

Die Studierenden sind in der Lage, vorgegebene oder selbstgewählte Inhalte so aufzubereiten, dass sie sie anschliessend kennen.

2 Verständnis

Den Studierenden gelingt es, während der Aufbereitung der Inhalte ein vertieftes Verständnis dieser zu erlangen.

3 Anwendung

Die Studierenden können die aufbereiteten Inhalte anwenden. Zur Anwendung der Inhalte gehört im Rahmen der Präsentation auch die zielgruppengerechte Kommunikation der Inhalte.

4 Analyse

Den Studierenden gelingt es, die inhaltlichen Zusammenhänge des aufbereiteten Stoffs zu analysieren.

5 Synthese

Die Studierenden schaffen eine eigene Synthese der erarbeiteten Inhalte.

6 Bewertung

Die Studierenden sind zu einer eigenständigen Bewertung der aufbereiteten Inhalte in der Lage.

In Kapitel 2.1 wurde ausführlich geschildert, dass es sich bei Präsentationen in Masterseminaren nicht ausschliesslich um eine summative Evaluation von Inhalten, sondern auch um eine formative Evaluation der Präsentationskompetenzen handelt. Die Taxonomie von Bloom et al. (1971) bietet für die Formulierung von Lernzielen hinsichtlich der Präsentationskompetenz keine geeignete Basis, da psychomotorische Lernziele zu stark auf muskuläre und motorische Fertigkeiten abzielen (Krathwohl et al., 1975). Der in Kapitel 1.2 zitierte Deskriptor zur kommunikativen Fähigkeit aus dem Qualifikationsrahmen für den schweizerischen Hochschulbereich (Schweizerische Universitätskonferenz, 2011) bietet zwei Anhaltspunkte für die Formulierung von Lernzielen. Zum einen weist „klar und eindeutig kommunizieren“ (S. 11) darauf hin, dass von den Studierenden eine klare Struktur und eine klare Sprache erwartet werden kann. Zum anderen verweist der Aspekt „sowohl an Experten wie auch an Laien“ darauf, dass die Studierenden in der Lage sein sollen, die Präsentation zielgruppengerecht zu gestalten. Aus eigener Erfahrung sowohl im wissenschaftlichen Bereich als auch in der freien Marktwirtschaft, sind für gelungene Präsentationen noch andere Aspekte notwendig. In Zeiten von vielfältigen Möglichkeiten, visuelle Hilfsmittel einzusetzen, gehört dazu die zielführende Auswahl und Gestaltung dieser. Zudem ist es in fast allen Kontexten notwendig, mit den Zuhörern zu interagieren – und sei es nur die Beantwortung von Fragen. Ein letzter Aspekt für gelungene Präsentationen ist die Gestaltung des zeitlichen Rahmens. Fast immer ist eine Obergrenze der Dauer vorgegeben, die es einzuhalten gilt.

Aus diesen fünf Aspekten ergeben sich die folgenden weiteren Lernziele hinsichtlich der Präsentationskompetenz:

7 Klare Kommunikation

Die Studierenden sind in der Lage, die Inhalte mit einer klaren Struktur und einer klaren Sprache zu präsentieren.

8 Zielgruppen-Spezifität

Die Studierenden sind in der Lage, die Präsentation auf die jeweilige Zielgruppe und deren Kenntnisstand anzupassen.

9 Gestaltung visueller Hilfsmittel

Die Studierenden sind in der Lage, die visuellen Hilfsmittel zielgerichtet und angemessen einzusetzen.

10 Interaktion mit den Zuhörern

Die Studierenden sind in der Lage, mit den Zuhörern souverän zu interagieren.

11 Zeitlicher Rahmen

Die Studierenden sind in der Lage, den vorgegebenen zeitlichen Rahmen einzuhalten.

Bevor aus diesen Lernzielen in Kapitel 5.1 Bewertungskriterien abgeleitet werden, sollen die bisherigen Versionen der Bewertungsschemata kritisch reflektiert werden. So können wichtige Aspekte, die bei der Überarbeitung des Bewertungsschemas berücksichtigt werden sollten, herausgearbeitet werden.

4. KRITISCHE REFLEXION DER BISHERIGEN BEWERTUNGSSCHEMATA

4.1 VERSION 1 DES BEWERTUNGSSCHEMAS UND DEREN ÜBERARBEITUNG

Schon längere Zeit werden die Präsentationen von Studierenden in Masterseminaren anhand definierter Bewertungskriterien benotet. Bis zum Herbstsemester 2015 bestand das Bewertungsschemas (im Folgenden als Version 1 (V1) bezeichnet) aus einer Liste von sechs übergeordneten Kategorien. Diese wurden mithilfe einer Aufzählung von Verhaltensweisen konkretisiert. Jede dieser Kategorien wurde auf einer 5-stufigen Skala benotet, die Gesamtnote wurde als Mittelwert der 6 Kategorien gebildet. Zur Verdeutlichung findet sich ein Ausschnitt aus dem Bewertungsschema sich in Abbildung 2, das vollständige Schema der V1 ist in Anhang I dargestellt.

INHALT	0	1	2	3	4
<ul style="list-style-type: none"> - Hinführung/Einführung & Einordnung, Vorwissen der Zielgruppe berücksichtigen - Thema für den jeweiligen Zweck bearbeiten, eigenständige Argumentation - Theoretische Einbettung, Bezüge herstellen - Vor- und Nachteile der methodischen Vorgehens - Verwendung anschaulicher Beispiele, Praxisbezug herstellen - Umgang mit Quellen 					
STRUKTUR	0	1	2	3	4
<ul style="list-style-type: none"> - „roter Faden“, klare Gliederung - Gliederung evtl. visualisieren, besonders wenn es sich um längere Vorträge handelt - Zusammenfassungen (u.a. Erkenntnisgewinn) 					

Abbildung 2: Ausschnitt aus dem Bewertungsschema für Präsentationen V1

In der Validierungsarbeit für das Modul B „Lern- und Lehr-Evaluation“ der Ausbildung für Hochschuldidaktik an der Universität Fribourg (eingereicht im November 2015) wurde dieses Bewertungsschema kritisch reflektiert. Daraus resultierend wurden vier mögliche Verzerrungen des Bewertungsschemas identifiziert. Diese sollen an dieser Stelle nur kurz aufgegriffen werden, eine ausführlichere Darstellung findet sich in der genannten Validierungsarbeit.

(1) Vermischung verschiedener Funktionen

In der Analyse der Funktion der Evaluation wurde herausgearbeitet, dass sowohl eine zertifikative als auch eine diagnostische Funktion verfolgt wird. Das Problem besteht nicht grundsätzlich in der Vermischung dieser Funktionen, sondern darin, dass die Aspekte beider Funktionen äquivalent in die Benotung mit einfließen.

(2) Bruch zwischen Inhalt der Lehre und Inhalt der Evaluation

Aus Anhang I wird deutlich, dass sowohl inhaltliche Aspekte benotet werden, als auch Aspekte der Präsentationskompetenz. Letzteres wird in den Seminaren jedoch nicht als Lehrinhalt aufgegriffen, sodass die Studierenden in Bezug auf etwas evaluiert werden, das nicht Inhalt der Lehre ist.

(3) Verzerrung durch Zeitpunkt der Präsentation im Semester

Studierende halten ihre Präsentation zu unterschiedlichen Zeitpunkten im Semester, wodurch Unterschiede in den Anforderungen und/oder den Möglichkeiten zum Lernen der geforderten Kompetenzen entstehen können.

(4) Berechnung der Note

Wie oben beschrieben, wird die Gesamtnote für die Präsentation durch die Berechnung des Mittelwerts der einzelnen Noten für die 6 Kategorien vorgenommen. Da es sich jedoch um eine Ordinal-Skala handelt, wird hier eine mathematische Operation verwendet, die für dieses Skalenniveau nicht zulässig ist.

Zusätzlich zu den genannten Verzerrungen wurden gewisse Schwächen des Bewertungsschemas V1 hinsichtlich der Validität und der Reliabilität identifiziert. Die **Validität** einer Evaluation beschreibt, ob die Evaluation auch tatsächlich die Aspekte abdeckt, die evaluiert werden sollen (Charlier, 2015b). Die kritischen Punkte in Bezug auf die Validität sind eng mit der unter Punkt (2) beschriebenen Verzerrung verknüpft: In der V1 des Bewertungsschemas wird nicht differenziert, ob Aspekte bewertet werden, die im Seminar gelehrt werden oder ob es sich um Aspekte handelt, die nicht explizit im Seminar aufgegriffen werden. Die **Reliabilität** einer Evaluation beschreibt die Zuverlässigkeit dieser (Charlier, 2015b). Die Reliabilität der V1 des Bewertungsschemas kann in Frage gestellt werden, da die Rolle der Spezifizierungen der übergeordneten Kategorien nicht klar ist. Zum einen sind sie nicht ausschließlich behavioral formuliert, sodass fraglich ist, ob sie beobachtet werden können. Zum anderen wird nicht spezifiziert, ob sie nur eine beispielhafte Aufzählung von Ankerpunkten sind, oder ob all diese Punkte (und nur diese) erfüllt sein müssen, um volle Punktzahl zu erreichen.

Basierend auf dieser Analyse wurde eine grundlegende Überarbeitung des Bewertungsschemas vorgenommen. Das Ziel dieser Überarbeitung war, möglichst viele der oben geschilderten Problembereiche zu adressieren. Im Folgenden werden die zentralen Veränderungen von der Version 1 zur Version 2 (V2) beschrieben. Zur Verdeutlichung werden hier Ausschnitte aus dem Bewertungsbogen (Abbildung 3) sowie den Verhaltensankern (Abbildung 4) dargestellt. Die vollständigen Unterlagen der Version 2 finden sich im Anhang II.

<u>INHALT</u>						
<u>AUFHÄNGER</u>						
EINFÜHRUNG IN DAS THEMA		0	1	2	3	4
<u>PRÄSENTATION WESENTLICHER BESTANDTEILE DES PAPERS</u>						
THEORIE & FRAGESTELLUNG		0	1	2	3	4
METHODE		0	1	2	3	4

Abbildung 3: Ausschnitt aus dem Bewertungsbogen des Bewertungsschemas für Präsentationen V2

Veränderung (1): Trennung der zertifikativen und diagnostischen Funktion

Die Bewertungskriterien in der V2 werden in die Unterkategorien „Inhalt“ und „Präsentationskompetenz“ aufgeteilt. Dabei wird die Note aus den Kriterien der Kategorie „Inhalt“ berechnet. Über die Kriterien der Kategorie „Präsentationskompetenz“ können ausschliesslich einzelne Zusatzpunkte erworben werden. Der Grund für die Beibehaltung einzelner Kriterien hinsichtlich der Präsentationskompetenz liegt in der diagnostischen Funktion der Evaluation: Würde die Präsentationskompetenz in der Evaluation nicht aufgegriffen werden, könnten Lernfelder weniger präzise identifiziert werden – die Identifikation und Rückmeldung von Lernfeldern stellt jedoch eine der zentralen Funktionen der diagnostischen Evaluation dar (Charlier, 2015b). Der Grund für die Vergabe von Extrapunkten für eine gute Präsentation ist ein motivationaler. Heckhausen und Heckhausen (2010) beschreiben Motivation als ein Produkt aus Person und Situation. Auf Seiten der Situation spielen Anreize eine wichtige Rolle, um Motivation entstehen zu lassen. In diesem Sinne ist die Vergabe von Extrapunkten zu interpretieren. Mit dieser Veränderung sollen die oben beschriebenen Punkte (1) und (2) sowie das Problem mit der Validität aufgegriffen werden.

Veränderung (2): „Verknüpfung mit bisherigen Inhalten“ als Bewertungskriterium

Um dem oben beschriebenen Punkt (3) zu begegnen, wurde ein Bewertungskriterium „Verknüpfung mit bisherigen Inhalten“ eingefügt. Je weiter das Semester fortgeschritten ist, desto mehr Inhalte aus den vorherigen Sitzungen können für die spezifische Präsentation relevant sein und in ihr aufgegriffen werden. Somit steigt die Komplexität dieser Anforderung im Verlauf des Semesters. Die Ungleichheit, die entstehen kann, wenn Studierende ihre Präsentationen zu unterschiedlichen Zeiten im Semester halten, kann damit zwar nicht

vollständig vermieden werden. Jedoch kann sie in der entsprechenden Bewertung dieses Punktes etwas abgefedert werden.

Veränderung (3): Formulierung von Verhaltensankern

In Bezug auf die Reliabilität wurde oben verdeutlicht, dass die Rolle der Verhaltenskonkretisierungen im Bewertungsschema V1 nicht klar war. Um diese Unklarheit aufzulösen, wurden für die einzelnen Bewertungskriterien konkrete Verhaltensanker formuliert. Diese haben unter anderem die Vorteile, dass die Anforderungen explizit beschrieben werden und der Fokus auf beobachtbares Verhalten gelegt wird, was wichtige Bedingungen für Evaluationen durch Beobachtung sind (Kapitel 2.2).

	0 (ungenügend)	1 (eher schwach)	2 (mittel)	3 (gut)	4 (exzellent)
AUFHÄNGER					
EINFÜHRUNG IN DAS THEMA	• Es erfolgt keine Einführung in das Thema, die Zuhörer werden nicht „abgeholt“; „Kaltstart“		• Es erfolgt ein „sachlicher“ Einstieg, die Zuhörer werden abgeholt, ohne dass spezifisches Interesse für das Thema geweckt wird		• Der Einstieg in das Thema ist motivierend und weckt Interesse für die Präsentation
PRÄSENTATION WESENTLICHER BESTANDTEILE DES PAPERS					
THEORIE & FRAGESTELLUNG	• Die Fragestellung des Artikels wird nicht klar und nachvollziehbar; die theoretische Grundlage wird nicht in Verbindung mit der Fragestellung gebracht	• Die Fragestellung wird erwähnt/die Hypothesen werden aufgelistet/genannt, ohne dass sie erklärt oder hergeleitet werden	• Die Fragestellung und die dazugehörigen Hypothesen werden hergeleitet, ohne dass eine direkte Verknüpfung zur theoretischen Grundlage gezogen wird	• Die Fragestellung und Hypothesen werden basierend auf der Theorie dargestellt; Theoretische Grundlage wird jedoch nicht selektiert / zielgerichtet ausgewählt	• Die Fragestellung und die dazugehörigen Hypothesen werden präzise auf Basis der Theorie hergeleitet; Die Präsentation der Theorie erfolgt selektiv / zielgerichtet
METHODE	• Weder die Stichprobe noch das Studiendesign und die Durchführung werden nachvollziehbar • Fehlerhafte Darstellung der angewendeten Methode		• Korrekte, aber unselektierte Darstellung der Stichprobe sowie der Durchführung und des Studiendesigns		• Die zum Verständnis notwendigen Informationen zu Stichprobe, Durchführung und Studiendesign werden präzise, selektiv und korrekt dargestellt

Abbildung 4: Ausschnitt aus den Verhaltensankern des Bewertungsschemas für Präsentationen V2

Über diese drei Veränderungen hinaus wurden noch Anpassungen vorgenommen, die nicht in einem direkten Bezug zu den in der Validierungsarbeit herausgearbeiteten Problembereichen stehen. Sie beziehen sich vielmehr darauf, die Transparenz hinsichtlich der Kriterien für die Studierenden zu erhöhen. Um den Studierenden eine angemessene Vorbereitung zu ermöglichen und den Lernerfolg zu fördern, ist die Kommunikation der Anforderungen sowie der Bewertungskriterien von grosser Bedeutung. Dazu dienen die Veränderungen (4) und (5).

Veränderung (4): Veröffentlichung der Verhaltensanker / Transparenz

Zu Beginn des Semesters wurden den Studierenden das Bewertungsschema sowie die Verhaltensanker der Bewertungsstufe „exzellent“ zur Verfügung gestellt (s. Anhang II.3: Verhaltensanker für Studierende). Die Entscheidung, ausschliesslich die Verhaltensanker für die Ausprägung „exzellent“ aufzugreifen, basiert auf der Theorie des sozialen Lernens von Bandura (1977). Eine der Grundideen dieser Theorie ist, dass Lernen durch Beobachtung von

Verhaltensmodellen geschieht. Auch wenn die Verhaltensmodelle in diesem Fall nur schriftlicher Art sind, ist es sinnvoll, positive Modelle und nicht negative Modelle zur Verfügung zu stellen.

Veränderung (5): Leitfrage und –motto für die Ausarbeitung des Referats

Um die Anforderungen an die Studierenden noch zu verdeutlichen, wurden zusätzlich zu den Verhaltensankern auch eine Leitfrage sowie ein Leitmotto für die Referatsausarbeitung formuliert.

Die V2 des Bewertungsschemas wurde im Frühlingsemester 2016 in einem Masterseminar eingesetzt. Die gemachten Erfahrungen sollen im folgenden Abschnitt reflektiert werden.

4.2 VERSION 2 DES BEWERTUNGSSCHEMAS UND IDENTIFIZIERTE PROBLEMFELDER

Im Vergleich zur V1 haben sich durch die Verwendung der V2 des Bewertungsschemas einige Verbesserungen ergeben. Diese beziehen sich überwiegend auf die Reduktion der oben geschilderten Probleme und sollen deshalb an dieser Stelle nicht noch einmal im Detail aufgegriffen werden.

Parallel zu den Verbesserungen sind jedoch während des ersten Einsatzes der V2 im Frühlingsemester 2016 auch einige Probleme deutlich geworden, die im Folgenden beschrieben werden.

Problem (1): Fehlende Fundierung der Bewertungskriterien

Wie in Kapitel 3 verdeutlicht, kommt der Fundierung von Bewertungskriterien auf Lernzielen eine entscheidende Bedeutung zu. Diese Fundierung ist für die V2 nicht gegeben: Die Kriterien wurden nicht von Lernzielen, sondern aus den inhaltlichen Aspekten der V1 abgeleitet. Eine fehlende Fundierung der Kriterien auf Lernzielen birgt aber die Gefahr, dass wichtige Aspekte bei der Evaluation ausser Acht gelassen werden und Unwichtiges evaluiert wird. Technisch gesprochen, kann es also zu einer Reduktion der Kriteriumsrelevanz kommen, also des Ausmasses, zu dem durch die Evaluation die wesentlichen Aspekte des Leistungskriteriums abgedeckt werden. Gleichzeitig werden die Kriteriumsirrelevanz sowie die Kriteriumsdefizienz erhöht. Unter Kriteriumsirrelevanz versteht man die Aspekte einer Evaluation, die nichts mit dem eigentlichen Leistungskriterium zu tun haben, während die Kriteriumsdefizienz die Aspekte des Leistungskriteriums beschreibt, die durch die Evaluation nicht abgedeckt werden (Marcus & Schuler, 2006). Durch reduzierte Kriteriumsrelevanz ist auch die Validität einer Evaluation gefährdet, da fraglich ist, ob tatsächlich die Aspekte evaluiert werden, die evaluiert werden sollen (Charlier, 2015b).

Problem (2): Anspruchsniveau der Verhaltensanker

Eine Bewertung unter konsequenter Interpretation und Anwendung der Verhaltensanker führte in der ersten Verwendung der V2 zu ungewöhnlich schlechten Noten: Der Median der Noten lag bei 3,5, der Mittelwert bei 3,53. Ohne Anwendung eines korrigierenden Faktors hätten nur 2 von 8 Studierenden die Note 4 erreicht. Auch wenn in dem Seminar insgesamt ein nicht allzu hohes Niveau konstatiert werden konnte, sind diese Noten für Präsentationen aussergewöhnlich schlecht.

Problem (3): Spezifität der V2 des Bewertungsschemas

Wie aus Anhang II deutlich wird, sind die Bewertungskategorien der V2 sehr stark auf die Präsentation eines Papers ausgerichtet. Diese Spezifität der Kriterien kann Vorteile haben, wenn die Anforderung die Präsentation eines einzelnen Papers ist. Aber schon bei der Präsentation von zwei zusammenhängenden Papers sind die Kriterien schwer anzuwenden. Und auch wenn in den meisten Fällen in den Masterseminaren der Arbeits- und Organisationspsychologie Paper die Inhalte von Präsentationen sind, so ist dies bei weitem nicht immer der Fall.

Problem (4): Verrechnung der einzelnen Bewertungskategorien

Für die V2 wurde zwar keine Mittelwertsbildung angewendet, um eine Gesamtnote zu errechnen. Gleichwohl wurden die einzelnen Punkte der Bewertungskriterien addiert, um aus der Gesamtpunktzahl eine Gesamtnote zu berechnen. Dazu wurde eine Formel verwendet, mithilfe derer eine Skala mit beliebig vielen Stufen in eine 6-stufige Skala, wie sie dem Schweizer Notensystem entspricht, überführt werden kann:

$$\left(\left(\frac{\text{maximale Note} - 1}{\text{Maximalpunktzahl}} \right) * \text{erreichte Punktzahl} \right) + 1$$

Mit der 6 als maximale Note sieht die Formel folgendermassen aus:

$$\left(\left(\frac{6 - 1}{\text{Maximalpunktzahl}} \right) * \text{erreichte Punktzahl} \right) + 1$$

Die Entscheidung, diese Variante der Notenberechnung dem Median vorzuziehen, hatte zwei Gründe. Zum einen liegt dem Bewertungsschema eine 5-stufige Skala zugrunde (0-4). Geht man davon aus, dass im Durchschnitt von den Studierenden eine mittlere Leistung erzielt wird, ist zu vermuten, dass sich bei der Nutzung des Medians dieser 5-stufigen Skala und der Umwandlung in die 6-stufige Notenskala eine sehr geringe Differenzierung ergibt. Anders ausgedrückt würden die Studierenden alle sehr ähnliche Noten erhalten. Zum anderen ist bei der Nutzung des Medians die Vergabe von Zusatzpunkten, wie sie unter dem Punkt „Veränderung (1)“

geschildert wird, nicht trivial umzusetzen. Dennoch wird mit der Addition eine arithmetische Funktion angewendet, die für Ordinal-Skalen nicht zulässig ist.

Problem (5): Keine Informationen zur Reliabilität der Evaluation

Auch wenn Verhaltensanker das Ziel haben, die Reliabilität der Evaluation zu erhöhen (s. Kapitel 2.2), kann derzeit keine Aussage darüber gemacht werden, ob die Evaluation reliabel ist.

4.3 ZIELSTELLUNG FÜR DIE ÜBERARBEITUNG DES BEWERTUNGSSCHEMAS

Aus diesen Problemfeldern ergeben sich die in Kapitel 1.3 bereits geschilderten Unterziele dieser Arbeit und damit für die Überarbeitung der V2 des Bewertungsschemas. An dieser Stelle sollen diese Unterziele noch einmal etwas ausführlicher aufgegriffen und in den Kontext der oben geschilderten Probleme der V2 gesetzt werden.

Unterziel (1) Eine Fundierung der Bewertungskriterien auf formulierten Lernzielen

An mehreren Stellen dieser Arbeit wurde bereits auf die Bedeutung von Lernzielen als Basis von Evaluationen und Evaluationskriterien hingewiesen. Um die theoretische Fundierung der Bewertungskriterien der neuen Version 3 (V3) des Bewertungsschemas zu gewährleisten, sollen diese aus den in Abschnitt 3.2 formulierten Lernzielen abgeleitet werden.

Unterziel (2) : Eine Kalibrierung der Verhaltensanker hinsichtlich ihres Anspruchsniveaus

Die V2 des Bewertungsschemas wurde als grafische Einstufungsskala konzipiert, das heisst, die Verhaltensanker wurden ohne allzu grossen Konstruktionsaufwand formuliert (vgl. Kapitel 2.2). Für die V3 des Bewertungsschemas wird eine verhaltensverankerte Einstufungsskala angestrebt, wobei die Verhaltensanker mithilfe der Methode kritischer Ereignisse erhoben werden sollen. Damit wird eine Anpassung der Schwierigkeit auf Basis der Meinung mehrerer Experten angestrebt.

Unterziel (3) Eine Verallgemeinerung der Kriterien

Um die Bewertungskriterien für eine breitere Anwendung zu konzipieren, ist sicherlich schon die Fundierung auf Basis der Lernziele ein entscheidender Schritt. So werden die Kriterien nicht danach strukturiert, was konkret die Basis der Präsentation ist, sondern nach den zu erwerbenden Kompetenzen. Damit sollten die Kriterien für verschiedene Arten von Präsentation anwendbar sein. Zudem wird bei der Erhebung der Verhaltensanker darauf geachtet, Experten aus verschiedenen Bereichen der Psychologie mit einzubeziehen. So fliessen unterschiedliche Perspektiven und Erfahrungen in die formulierten Erwartungen an die Studierenden mit ein.

Unterziel (4) Eine Anpassung der Notenberechnung gemäss der Möglichkeiten einer Ordinalskala

Zur Ermittlung einer Gesamtnote aus mehreren Kriterien, die auf Basis einer Ordinalskala eingeschätzt werden, ist der Median geeignet. Für die V3 soll ein Vorgehen entwickelt werden, mithilfe dessen der Median verwendet werden kann, um die Gesamtnote für die Präsentation zu ermitteln. Dabei sind insbesondere die Überführung des Medians in das Schweizer Notensystem sowie der Einbezug von Extrapunkten für die Kriterien bezüglich der Präsentationskompetenzen als Herausforderungen zu beachten.

5. METHODISCHES VORGEHEN

5.1 SCHRITT 1: ÜBERARBEITUNG DER BEWERTUNGSKRITERIEN

Die Ableitung der Bewertungskriterien aus den Lernzielen erfolgt durch die Operationalisierung der Lernziele. Mit Operationalisierung wird das Verfahren beschrieben „eine nicht direkt beobachtbare Variable [...] für die Beobachtung [...] zugänglich zu machen, indem man eine mit ihr verknüpfte, gut beobachtbare Variable [...] auswählt“ (Gadonne, 2014). Um die Operationalisierung der in Kapitel 3.2 definierten Lernziele für Präsentationen in Masterseminaren vorzunehmen, wurde ein zwei-schrittiges Vorgehen gewählt. Zunächst wurden aus der Taxonomie von Bloom und Kollegen (1971) für jede Überkategorie diejenigen Unterkategorien ausgewählt, die für eine Evaluation von Präsentationen geeignet sind. Dafür wurden auch die detaillierten Beschreibungen der Unterkategorien von Bloom und Kollegen (1971) zur Hilfe genommen (S. 271-273). In einem zweiten Schritt wurden diese Lernziele dann durch die Beantwortung der Frage „Wodurch zeigt sich [...] aufbereiteten Inhalte?“ konkretisiert. Dabei wurde für [...] jeweils einer der Begriffe der Oberkategorien eingesetzt, also z.B. „Wodurch zeigt sich Wissen über die aufbereiteten Inhalte?“. So ergaben sich für jede Oberkategorie 2-3 konkrete Aspekte, die als Bewertungskriterien dienen können. Diese sind in [Tabelle 1](#) ersichtlich. Die numerische Kennzeichnung jedes Kriteriums bezieht sich dabei auf die jeweilige Unterkategorie der Lernziele (s. [Abbildung 1](#)).

Tabelle 1: aus den formulierten Lernzielen abgeleitete Bewertungskriterien zum Inhalt

Lernziel	Bewertungskriterien
1 Wissen Die Studierenden sind in der Lage, vorgegebene oder selbstgewählte Inhalte so aufzubereiten, dass sie sie anschliessend kennen.	Wissen über die aufbereiteten Inhalte zeigt sich durch... 1.11 Kenntnis der für das Thema relevanten Begriffe 1.25 Kenntnis der fachspezifischen Methodik 1.32 Kenntnis der themenspezifischen Theorien und/oder Klassifikationen
2 Verständnis Den Studierenden gelingt es, während der Aufbereitung der Inhalte ein vertieftes Verständnis dieser zu erlangen.	Verständnis der aufbereiteten Inhalte zeigt sich durch... 2.1 Korrekte Wiedergabe / Paraphrasierung der Inhalte in eigenen Worten 2.1 Korrekte Einbettung der Inhalte in die Präsentation 2.2 korrekte Interpretation der Inhalte
3 Anwendung Die Studierenden können die aufbereiteten Inhalte anwenden. Zur Anwendung der Inhalte gehört im Rahmen der Präsentation auch die zielgruppengerechte Kommunikation der Inhalte.	Eine Anwendung der aufbereiteten Inhalte zeigt sich durch... 3 korrekte Verwendung der relevanten Begriffe und Theorien 3 Anpassung der aufbereiteten Inhalte auf den Kontext 3 Anpassung der aufbereiteten Inhalte auf die Zielgruppe

<p>4 Analyse Den Studierenden gelingt es, die inhaltlichen Zusammenhänge des aufbereiteten Stoffs zu analysieren.</p>	<p>Die Fähigkeit zur Analyse zeigt sich durch...</p> <p>4.2 korrekte Analysen der Zusammenhänge innerhalb der aufbereiteten Inhalte</p> <p>4.2 korrekte Analysen der Zusammenhänge zwischen den aufbereiteten Inhalten und den Inhalten des Seminars</p>
<p>5 Synthese Die Studierenden schaffen eine eigene Synthese der erarbeiteten Inhalte.</p>	<p>Die Fähigkeit zur Synthese zeigt sich durch...</p> <p>5.1 die zielgerichtete Auswahl der zentralen Elemente des aufbereiteten Inhalts</p> <p>5.3 die zielgerichtete Zusammenfassung der aufbereiteten Inhalte</p> <p>5.3 die Ableitung (abstrakter) Implikationen und weiterführender Gedanken</p>
<p>6 Bewertung Die Studierenden sind zu einer eigenständigen Bewertung der aufbereiteten Inhalte in der Lage.</p>	<p>Die Fähigkeit zur Bewertung zeigt sich durch...</p> <p>6.1 eine kritische Reflexion der aufbereiteten Inhalte hinsichtlich der Inhalte und der darin enthaltenen Aussagen an sich</p> <p>6.2 eine kritische Reflexion der aufbereiteten Inhalte hinsichtlich fachspezifischer Konventionen, Methodik und Theorien</p>

Um aus den Lernzielen hinsichtlich der Präsentationen Bewertungskriterien abzuleiten, wurde das gleiche Vorgehen gewählt. Die sich daraus ergebenden Kriterien sind in Tabelle 2 ersichtlich.

Tabelle 2: Aus den formulierten Lernzielen abgeleitete Bewertungskriterien zur Präsentationskompetenz

Lernziel	Bewertungskriterien
<p>7 Klare Kommunikation Die Studierenden sind in der Lage, die Inhalte mit einer klaren Struktur und einer klaren Sprache zu präsentieren.</p>	<p>Eine klare Kommunikation zeigt sich durch...</p> <ul style="list-style-type: none"> - eine klare und nachvollziehbare Struktur - freies und gut verständliches Sprechen
<p>8 Zielgruppen-Spezifität Die Studierenden sind in der Lage, die Präsentation auf die jeweilige Zielgruppe und deren Kenntnisstand anzupassen.</p>	<p>Zielgruppenspezifität der Präsentation zeigt sich durch...</p> <ul style="list-style-type: none"> - Berücksichtigung des Vorwissens der Zielgruppe - Anpassung der Geschwindigkeit der Präsentation an die Zielgruppe - Anpassung der Wortwahl an die Zielgruppe
<p>9 Gestaltung visueller Hilfsmittel Die Studierenden sind in der Lage, die visuellen Hilfsmittel zielgerichtet und angemessen einzusetzen.</p>	<p>Eine zielführende Gestaltung der Hilfsmittel zeigt sich durch...</p> <ul style="list-style-type: none"> - ein angemessenes Verhältnis von visualisierten Inhalten und gesprochenem Wort - eine für den Kontext angemessene Gestaltung der Hilfsmittel

<p>10 Interaktion mit den Zuhörern Die Studierenden sind in der Lage, mit den Zuhörern souverän zu interagieren.</p>	<p>Eine souveräne Interaktion mit den Zuhörern zeigt sich durch...</p> <ul style="list-style-type: none"> - Blickkontakt mit allen Zuhörern - souveräne Reaktion auf Fragen von den Zuhörern - Einbindung der Zuhörer
<p>11 Zeitlicher Rahmen Die Studierenden sind in der Lage, den vorgegebenen zeitlichen Rahmen einzuhalten.</p>	<p>Die Einhaltung des zeitlichen Rahmens zeigt sich durch...</p> <ul style="list-style-type: none"> - keine wesentliche Über- oder Unterschreitung der vorgegebenen Zeit

5.2 SCHRITT 2: ENTWICKLUNG DER VERHALTENSANKER

Um diese Kriterien in einer Beobachtungssituation evaluieren zu können, ist es – wie in Kapitel 2 geschildert – hilfreich, wenn nicht notwendig, Verhaltensanker zur Verfügung zu haben. Diese Verhaltensanker wurden mithilfe der in Kapitel 2.3 vorgestellten Methode kritischer Ereignisse (CIT) erhoben. Für die Anwendung der CIT ist die Auswahl der Experten, die die kritischen Ereignisse liefern können, sehr wichtig. Als Experten wurden in dieser Arbeit Personen betrachtet, die regelmässig im Hochschulkontext mit der Bewertung von Präsentationen konfrontiert sind. Bei der Auswahl der Experten wurde darauf geachtet, möglichst alle Funktionsstufen der wissenschaftlichen Angestellten an der Hochschule abzudecken. Da eines der Ziele dieser Arbeit die Verallgemeinerung des Bewertungsschemas für die Anwendung im Department für Psychologie ist, wurden Personen aus diesem Fachgebiet ausgewählt. Zudem wurden zwei Psychologinnen in Deutschland angefragt. Eine ist wissenschaftliche Mitarbeiterin an einer Hochschule, die andere ist sowohl als Lehrbeauftragte an einer Hochschule als auch im Bereich Organisationsentwicklung im öffentlichen Dienst tätig. Der Autorin war es wichtig, auch eine Expertin zu befragen, die in einem Bereich ausserhalb der Hochschule tätig ist, um auch den Aspekt der Vorbereitung der Studierenden auf Anforderungen im Beruf ausserhalb der Universität aufzugreifen.

Konkret wurden vom Department für Psychologie an der Universität Fribourg zwei Professoren, zwei Personen aus dem oberen Mittelbau und zwei Doktoranden sowie als Externe zwei Psychologinnen aus Deutschland angefragt, die als Experten für die Erhebung der kritischen Verhaltensweisen fungierten. Um den Arbeitsaufwand für jede dieser Personen zu reduzieren, wurde die Erhebung der Verhaltensanker in inhaltliche und präsentationskompetenz-bezogene Bewertungskriterien aufgeteilt. Aus jeder Funktionsgruppe wurden je einer Person die inhaltlichen Aspekte und der anderen Person die Aspekte zur Präsentationskompetenzen zugesendet. Um die Methode der kritischen Ereignisse durchzuführen, ist es nach Flanagan (1954) wichtig, fünf Schritte zu befolgen (s. Kapitel 2.3). Der *erste Schritt*, die genaue Definition der Erwartungen an denjenigen, der beobachtet

und evaluiert wird, ist durch die Formulierung der Lernziele und der Bewertungskriterien erfolgt. Der *zweite kritische Schritt* ist die präzise Instruktion der Experten. Diese erfolgte in dieser Arbeit schriftlich per Mail. In dieser Mail wurden kurz der Kontext sowie das Ziel der Erhebung geschildert (für die vollständige Mail s. Anhang III.1). Ohne die Methode kritischer Ereignisse explizit zu nennen, wurde diese kurz beschrieben. Neben dieser Einführung ist die genaue Aufgabenstellung für die Experten von grosser Bedeutung. Diese wurde wie folgt formuliert:

„Stelle dir zu jedem dieser Bewertungskriterien einen Studenten/eine Studentin vor, die während einer Präsentation dieses Kriterium besonders schlecht oder aussergewöhnlich gut erfüllt haben. Welches konkrete Verhalten haben diese Studierenden gezeigt, damit ihr Verhalten besonders ineffektiv oder effektiv für die Erfüllung des Kriteriums war? Beschreibe dieses Verhalten bitte so genau wie möglich und gib auch an, warum dieses Verhalten besonders ineffektiv/effektiv war und welche konkreten Konsequenzen das jeweilige Verhalten hatte“.

Dabei wurde insbesondere darauf geachtet, dass die Instruktion auch – wie von Blicke (2014b) gefordert – Hinweise enthält, die Effizienz des beschriebenen Verhaltens sowie dessen Folgen zu beachten. Abschliessend wurde darauf hingewiesen, dass Fragen zur Formulierung der kritischen Verhaltensweisen jederzeit an die Autorin gestellt werden können. Zudem wurde anhand eines der Bewertungskriterien aus der Version 2 ein Beispiel für das Zielprodukt zur Verfügung gestellt. Dabei wurden die Verhaltensanker gegenüber der Version 2 jedoch etwas umformuliert, um einen stärkeren Fokus auf beobachtbares Verhalten zu legen. Zudem wurde den Experten ein Erhebungsbogen zugesendet, in dem das jeweilige Lernziel mit den dazugehörigen Kriterien aufgeführt wurde, sowie Platz für die Beschreibung des Verhaltens, der Gründe für die Effizienz/Ineffizienz sowie die Konsequenzen vorgesehen war (s. Anhang III.2 und III.3). Den Experten wurde ungefähr drei Wochen Zeit für die Erstellung der Verhaltensanker gegeben. Nach Ablauf dieser drei Wochen lagen die Antworten von allen acht Experten vor.

Der nächste Schritt in der Erstellung der Verhaltensanker bestand darin, aus den von den Experten genannten kritischen Ereignissen eine Synthese zu schaffen. Dieser Schritt erfolgte durch eine qualitative Analyse der eingegangenen Verhaltensanker. Wie durch die Instruktion der Experten beabsichtigt, handelte es sich meist um sehr konkrete Beispiele von Verhaltensweisen. Für die Formulierung der Bewertungskriterien ist es jedoch notwendig, diese Verhaltensweisen auf ein etwas abstrakteres Niveau zu bringen. Dazu wurden zunächst für jedes Bewertungskriterium die Aspekte identifiziert, die von drei bis vier Experten genannt wurden und in einer verallgemeinernden Weise in die Verhaltensanker aufgenommen. Meistens waren damit die wesentlichen Aspekte abgedeckt. In

einem zweiten Schritt wurden die Aspekte analysiert, die von mindestens zwei Experten in den Beispielen genannt wurden. Auch die wurden auf einer abstrakteren Ebene in die Verhaltensanker integriert. Abschliessend wurden die verbleibenden Aspekte, die jeweils nur von einem der Experten genannt wurden, auf ihre Bedeutsamkeit und Möglichkeit zur Verallgemeinerung für verschiedene Anwendungsbereiche überprüft. Wenn beides gegeben war, wurden diese Aspekte auch noch aufgenommen. Abschliessend wurde versucht, die einzelnen Punkte nicht nebeneinander stehen zu lassen, sondern in einem zusammenhängenden Verhaltensanker zu formulieren. Eine gewisse Subjektivität in der Auswertung der kritischen Ereignisse ist nicht auszuschliessen. Jedoch besteht die Objektivierung der Kriterien in dem Schritt, mehrere Experten zur Formulierung der möglichen Verhaltensweisen heranzuziehen. Zudem wurde versucht, der Subjektivität der Auswertung insofern entgegenzuwirken, dass zunächst die Aspekte aufgenommen wurden, die von mehreren Experten genannt wurden.

5.3 SCHRITT 3: ERSTELLUNG DER VERSION 3 DES BEWERTUNGSSCHEMAS

Aus der oben beschriebenen Synthese wurde die finale V3 des Bewertungsschemas für Präsentationen erstellt. Die vollständigen Unterlagen dieser Version finden sich in Anhang IV. Wie aus [Abbildung 5](#) ersichtlich wird, folgt diese dritte Version dem gleichen Grundprinzip der Bewertung wie die zweite Version. Jedes Kriterium wird auf einer 5-stufigen Skala (0 – 4) eingeschätzt, wobei sich in der Bewertungspraxis gezeigt hat, dass auch die Zwischenstufen (0,5; 1,5; 2,5; 3,5) für die differenzierte Bewertung hilfreich sind. Forschung in Bezug auf unterschiedliche Skalenformate hat gezeigt, dass die Reliabilität bei 5- bis maximal 9-stufigen Skalen am höchsten ist (Marcus & Schuler, 2006; Moser & Schuler 2014). Sowohl mit als auch ohne die Zwischenstufen befindet sich die verwendete Skala von 0 bis 4 genau in diesem Bereich, sodass es keinen Grund gab, die Skalierung zu verändern.

INHALT						
1 Wissen						
Die Studierenden sind in der Lage, vorgegebene oder selbstgewählte Inhalte so aufzubereiten, dass sie sie anschliessend kennen.						
Kenntnis der für das Thema relevanten Begriffe		0	1	2	3	4
Kenntnis der fachspezifischen Methodik		0	1	2	3	4
Kenntnis der themenspezifischen		0	1	2	3	4

Abbildung 5: Ausschnitt aus dem Bewertungsbogen des Bewertungsschemas für Präsentationen V3

Auch empfehlen Marcus & Schuler (2006) sowie Moser & Schuler (2014), dass mindestens die Skalenendpunkte mit Verhaltensankern versehen werden sollten, um eine gute Reliabilität der Beurteilung zu erreichen. Wie aus [Abbildung 6](#) ersichtlich wird, ist auch dieses Kriterium bei der dritten Version weiterhin erfüllt. Im Vergleich zu der Version 2 sind die Beschreibung der Lernziele sowie einer detaillierten Beschreibung der Bewertungskriterien, abgeleitet aus den Lernzielen, hinzugekommen. So wird für die Studierenden besser ersichtlich, welche Basis die jeweiligen Kriterien haben, als dies bei Version 2 der Fall war.

Lernziel	Bewertungskriterien	besonders schlechte Erfüllung des Kriteriums	aussergewöhnlich gute Erfüllung des Kriteriums
1 Wissen Die Studierenden sind in der Lage, vorgegebene oder selbstgewählte Inhalte so aufzubereiten, dass sie sie anschliessend kennen.	Wissen über die aufbereiteten Inhalte zeigt sich durch...		
	Kenntnis der für das Thema relevanten Begriffe	Die Studierenden verwenden zentrale Begriffe gar nicht oder falsch, vertauscht sie oder setzen sie in den falschen Kontext und führen sie nicht korrekt ein.	Die Studierenden präsentieren und verwenden die zentralen Begriffe korrekt präsentiert und erklären sie präzise.
	Kenntnis der fachspezifischen Methodik	Die Studierenden beschreiben und erklären die verwendete Methodik nicht korrekt (z.B. <i>wissen die Studierenden nicht, welche Methode angewendet wurde</i>) und verwechseln methodische Termini (z.B. <i>Experiment vs Feldforschung</i>).	Die Studierenden beschreiben die verwendete Methodik einfach, korrekt und prägnant; Sie verknüpfen sie korrekt mit den inhaltlichen Aspekten.
	Kenntnis der themenspezifischen Theorien und/oder Klassifikationen	Die Studierenden erklären die theoretischen Grundkonzepte nicht oder falsch, sodass der Kern der Theorie/Konzepte nicht deutlich wird.	Die Studierenden erläutern die relevanten theoretischen Grundkonzepte korrekt und prägnant.

Abbildung 6: Ausschnitt aus den Verhaltensankern des Bewertungsschemas für Präsentationen V3

Im Vergleich zu den in Kapitel 3.2 formulierten Lernzielen und in Kapitel 5.1 abgeleiteten Bewertungskriterien wird in der finalen Version 3 (s. Anhang IV) eine Veränderung deutlich: Das Kriterium „8 Zielgruppen-Spezifität“ wurde nicht mehr aufgegriffen. Der Grund hierfür ist darin zu sehen, dass es bei den von den Experten formulierten Verhaltensbeispielen sehr grosse Überschneidungen zwischen diesem Kriterium und dem Kriterium 3.3 „Anpassung der aufbereiteten Inhalte auf die Zielgruppe“ gab. Diese Überschneidung ist insofern besonders beachtenswert, da die kritischen Ereignisse für die Kriterien 1-6 (Inhalt) von anderen Experten formuliert wurden als die Kriterien 7-10 (Präsentationskompetenz). Die Überschneidung ist also nicht darauf zurückzuführen, dass bestimmte Beispiele von den Experten bei einem früheren Kriterium formuliert und bei Kriterium 8 wegen der erhöhten Zugänglichkeit durch die vorausgehende Formulierung wieder abgerufen wurde. Es ist nicht zielführend, die Bewertung einzelner Kriterien doppelt in die Benotung einfließen zu lassen. Da die Anpassung an die Zielgruppe im Bereich der Anwendung (Kriterium 3) ein bedeutendes Element ist, wurde das Kriterium 8 aus den Präsentationskriterien herausgenommen.

Neben der Veränderung der Kategorien der Bewertungskriterien sowie der Neuformulierung der Verhaltensanker wurden zwei weitere Veränderungen von der Version 2 zur Version 3 vorgenommen, die im Folgenden kurz geschildert werden sollen.

(1) Verwendung des Medians zur Festlegung der Note

In den Kapiteln 4.2 und 4.3 wurde dargestellt, dass bei der Version 2 eine arithmetische Formel verwendet wurde, um aus der Bepunktung der einzelnen Kategorien eine Gesamtnote zu berechnen. Da dem Bewertungsschema eine Ordinal-Skala zugrunde liegt, ist die Berechnung eines Summenwertes nicht zulässig. Daher wird im Herbstsemester der Median aus den Punkten der einzelnen Kategorien zur Bildung der Gesamtnote verwendet. Im vorliegenden Bewertungsschema ist die Verwendung des Medians mit zwei Herausforderungen verbunden: (a) Die Umwandlung des Medians in eine Note des Schweizerischen Notensystems (1-6) und (b) der Einbezug der Extrapunkte für die Kriterien hinsichtlich der Präsentationskompetenzen. Die Herausforderung (a) entsteht, da die Kriterien des Bewertungsschemas auf einer 9-stufigen Skala bewertet werden, während für die Notenvergabe im Schweizerischen Notensystem eine 11-stufige Skala (1 – 1.5 – 2 – 2.5 – 3 – 3.5 – 4 – 4.5 – 5 – 5.5 – 6) zur Verfügung steht. Um diese Umwandlung vorzunehmen, ist es zunächst notwendig zu definieren, bei welcher Erfüllung der Kriterien die Präsentation der Studierenden als „genügend/bestanden“ betrachtet werden kann (Note ≥ 4). Für das vorliegende Bewertungsschema wurde dies für den Bereich 1-4 der Skala als gegeben angesehen. Für diesen Bereich wurde eine möglichst gleichmässige Zuordnung der möglichen Median-Werte zu den Noten 4-6 vorgenommen. Gleiches wurde auch für den Bereich 0-0.75 und die Noten 1-3 vorgenommen. Da der Median aber nur bestimmte Werte annehmen kann, war es nicht möglich, allen möglichen Notenstufen einen Wert zuzuordnen. Aus den oben geschilderten Überlegungen ergibt sich die folgende Zuordnung (Tabelle 3):

Tabelle 3: Zuordnung Median - Note

Median	Note	
0	1	ungenügend
0.25 - 0.5	2	
0.75	3	
1.0 - 1.5	4	genügend
1.75 - 2.25	4.5	
2.5 - 3.0	5	
3.25 - 3.5	5.5	
3.75 - 4	6	

Die Herausforderung (b) entsteht, da es nun nicht mehr möglich ist – wie bei der Version 2 – die Extrapunkte für die Präsentationskompetenzen auf den Summenwert aufzuaddieren. Um diese Herausforderung zu lösen, wurde das folgende Vorgehen entwickelt: Für jeden Studierenden wird bestimmt, für wie viele Kategorien im Bereich „Präsentationskompetenzen“ der Studierende Extrapunkte bekommt (z.B. 3 Extrapunkte). Extrapunkte werden vergeben, wenn der Median für eine der Kategorien 7-10 ≥ 3 ist. Also sind maximal 4 Extrapunkte möglich. Der

Punktliste für den Bereich „Inhalt“, die zur Bestimmung des Medians erstellt wurde, wird nun so oft die „4“ hinzugefügt, wie Extrapunkte vergeben wurden (in dem beschriebenen Beispiel wird 3 Mal). Je öfter die Maximalpunktzahl hinzugefügt wird, desto mehr wird der Median nach oben verschoben. Gleichzeitig verändert sich der Median aber nur, wenn die dem ursprünglichen Median naheliegenden Punkte höher sind als der ursprüngliche Median. Dies wird dem Gedanken der „Extrapunkte“ gerecht. Bei einer Liste von 16 Zahlen und einem Hinzufügen von maximal 4 Zahlen können die hinzugefügten Werte (also die „4“) nicht zum Median werden. Somit wird zudem verhindert, dass die Studierenden eine Note bekommen, die besser ist als die über den Bereich „Inhalt“ maximal erreichte Punktzahl. Auch dies entspricht dem Gedanken der „Extrapunkte“. Zusammengefasst wird durch dieses Vorgehen erreicht, dass über eine gute Präsentation der Median zwar verbessert werden kann, jedoch nicht über die ursprünglich maximal erreichte Punktzahl hinaus.

(2) Veröffentlichung der Kriterien / Transparenz

Wie in Kapitel 4.1 unter Veränderung (4) beschrieben, wurden den Studierenden die Bewertungskriterien der V2 ausschliesslich mit dem Verhaltensanker der Stufe „exzellent“ zur Verfügung gestellt. Hintergrund dieser Entscheidung war die soziale Lerntheorie von Bandura (1977), in welcher auf die Bedeutung von positiven Verhaltensbeispielen hingewiesen wird. Im Bereich der Personalentwicklung wird die soziale Lerntheorie von Bandura im Rahmen der Verhaltensmodellierung (Behavior Modeling Training) angewendet (Demmerle, Schmidt, Hess, Solga & Ryschka, 2011). Zwar ist diese Technik in der Umsetzung sehr umfangreich, dennoch können die Erkenntnisse dieses Ansatzes zum Teil auf die Veröffentlichung der Bewertungskriterien übertragen werden. Ein zentraler Bestandteil der Verhaltensmodellierung ist die Verwendung von Lernzielen, welche möglichst als Handlungsanweisungen formuliert sein sollten (Demmerle et al., 2011). Daher wurden bei der Version 3 gemeinsam mit den Verhaltensankern auch die Lernziele an die Studierenden kommuniziert. Des Weiteren zeigt Forschung im Bereich der Verhaltensmodellierung, dass eine Anwendung des Gelernten am besten dann gelingt, wenn sowohl positive als auch negative Verhaltensmodelle gezeigt wurden (Taylor, Russ-Eft & Chan, 2005). Daher wurden für die V3 neben den Lernzielen und den positiven Verhaltensankern auch die negativen Verhaltensanker veröffentlicht (s. Abb. 6 und Anhang IV.2).

Basierend auf diesen Veränderungen wurde die Version 3 des Bewertungsschemas in den ersten Wochen des Herbstsemesters 2016 getestet.

6. ERSTE BEWERTUNG DER VERSION 3 DES BEWERTUNGSSCHEMAS

6.1 VALIDITÄT

Die *Validität* – im Deutschen auch häufig als *Gültigkeit* bezeichnet – eines Verfahrens gibt an, ob dieses Verfahren tatsächlich das misst, was es zu messen vorgibt und was es messen soll (Harting, Frey & Jude, 2012). In Bezug auf Evaluationsverfahren ist Validität dann gegeben, wenn auch tatsächlich das evaluiert wird, was evaluiert werden soll (Charlier, 2015b). Durch die Fundierung der V3 des Bewertungsschemas auf Lernzielen wird der Fokus der Evaluation auf die Aspekte gelegt, die von den Studierenden gelernt werden sollen. Somit führt die Verknüpfung mit Lernzielen zu einer Erhöhung der Validität. Auch die gezielte Trennung der diagnostischen und der zertifikativen Evaluation im Rahmen der Bewertung von Präsentationen führt zu einer Erhöhung der Validität. Die Note basiert primär auf der Bewertung von Kompetenzen, die in einem spezifischen Seminar erworben werden sollen (d.h. den inhaltlichen Aspekten). Kompetenzen, die erst im Verlauf des Masterstudiums zu erwerben sind (d.h. Präsentationskompetenzen), gehen dagegen nur über Zusatzpunkte in die Bewertung mit ein. Somit werden Studierende im Vergleich zur V1 nicht mehr in Bezug auf Kompetenzen bewertet, was sie noch nicht vollständig erworben haben können.

Allerdings zeigt die Anwendung im Herbstsemester 2016 auch die Grenzen der V3 hinsichtlich der Validität auf. Diese Grenzen hängen eng mit der Schwierigkeit zusammen, kognitive Lernziele anhand von beobachtbarem Verhalten zu operationalisieren und sollen beispielhaft an dem Kriterium des Wissens verdeutlicht werden. Das Lernziel für dieses Kriterium ist wie folgt formuliert: *Die Studierenden sind in der Lage, vorgegebene oder selbstgewählt Inhalte so aufzubereiten, dass sie sie anschliessend kennen.* Darauf aufbauend heisst das erste Bewertungskriterium: *Kenntnis der für das Thema relevanten Begriffe.* Der Verhaltensanker für die Stufe „exzellent“ wurde aus den Experten-Sammlungen wie folgt abgeleitet: *Die Studierenden präsentieren und verwenden die zentralen Begriffe korrekt und erklären sie präzise.* Die Frage, die sich stellt, ist jedoch, wie gut im Kontext einer Präsentation von der korrekten Verwendung und präzisen Erklärung von Begriffen tatsächlich auf das Wissen der Studierenden geschlossen werden kann. Präsentationen können im Detail vorbereitet werden. Technische Umsetzungen wie z.B. PowerPoint bieten die Möglichkeit, mit Notizen zu arbeiten, in denen die präzisen Definitionen von Begriffen geschrieben stehen können. In diesem konkreten Fall ist bei der Erfüllung des Verhaltensankers also nicht unbedingt gegeben, dass auch tatsächlich die Kenntnis der Begriffe vorhanden ist. Auf der anderen Seite kann aber sicherlich bei einer falschen Verwendung im Sinne des Verhaltensankers der geringen Ausprägung (*Die Studierenden verwenden zentrale Begriffe gar nicht oder falsch, vertauschen sie oder setzen sie in*

den falschen Kontext und führen sie nicht korrekt ein) davon ausgegangen werden, dass das Wissen nicht vorhanden ist oder im gegebenen Kontext nicht abgerufen werden kann. Somit ist es nicht so, dass mit dem Bewertungsschema nicht das gemessen werden kann, was gemessen werden soll. Vielmehr muss man sich bei der Anwendung bewusst machen, dass der Kontext einer Präsentation ggf. nicht ausreicht, um das Erreichen der Lernziele vollständig zu evaluieren. In den Masterseminaren der Option Arbeits- und Organisationspsychologie wird die Präsentation daher häufig durch eine Klausur am Ende des Semesters ergänzt.

6.2 RELIABILITÄT

Die *Reliabilität* beschreibt die *Genauigkeit* oder auch die Messfehlerfreiheit einer Messung oder eines Verfahrens (Schermelleh-Engel & Werner, 2012). Es existieren verschiedene Möglichkeiten, die Reliabilität zu bestimmen. Eine davon ist die *Test-Retest Reliabilität*, bei der die Genauigkeit einer Evaluation über mehrere Situationen hinweg geprüft wird. Diese ist in der Lernevaluation insofern schwer umzusetzen, da sie auf der Voraussetzung beruht, dass sich die „wahren Werte“ der evaluierten Personen zwischen den Evaluationen nicht verändern (Schermelleh-Engel & Werner, 2012). In der Lehre zwei Situationen zu schaffen, die auf den gleichen Voraussetzungen beruhen und in Bezug auf welche sich die Studierenden in der Zeit zwischen den beiden Evaluationen nicht verändert haben, ist jedoch kaum möglich. Ein Student wird eine Präsentation ein zweites Mal nicht in exakt der Weise halten, sondern idealerweise aus den Erfahrungen der ersten Präsentation gelernt haben. Daher bietet es sich an, die Reliabilität von Evaluationen über mehrere Evaluatoren hinweg zu bestimmen (*Interrater-Reliabilität*). Grundsätzlich trägt die Basierung der Evaluation auf klar formulierten Verhaltensankern bereits zu einer Erhöhung der Reliabilität bei. Die Verhaltensanker schaffen über mehrere Evaluatoren hinweg ein gemeinsames Verständnis für die Anforderungen, die an die Studierenden gestellt werden. Somit wird die Subjektivität der Bewertung und damit auch der Messfehler in der Bewertung reduziert.

Da in dem Seminar, in dem die V3 des Bewertungsschemas getestet wurde, immer zwei Dozenten anwesend waren, kann die Interrater-Reliabilität für die V3 aber auch numerisch überprüft werden. Zu beachten sind dabei jedoch zwei Dinge. Erstens wurden zum jetzigen Zeitpunkt des Semesters (November 2016) erst 4 Präsentationen gehalten (drei mit je zwei Präsentierenden und eine mit einer Präsentierenden). Zum anderen wurde die Beurteilung der Präsentationskompetenzen bei zwei Präsentierenden zwischen den Evaluatoren aufgeteilt, um die Komplexität etwas zu reduzieren. Somit kann die Interrater-Reliabilität nur für die 16 inhaltsbezogenen Kriterien berechnet werden.

Es gibt verschiedene Ansätze, die Interrater-Reliabilität zu bestimmen (McHugh, 2012). Eine davon ist die *prozentuale Übereinstimmung*. Hier wird bestimmt, für wie viele einzelne Einschätzungen die Werte verschiedener Rater exakt übereinstimmt. Im hier vorliegenden Fall bedeutet das, dass die Prozentzahl der Fälle bestimmt wird, für die die Bewertung der 16 Kriterien in den vier Präsentationen zwischen den beiden Evaluatoren exakt gleich ist. Die prozentuale Übereinstimmung liegt bei 28.6%. Diese Zahl ist recht ernüchternd. Allerdings muss man sich vor Augen halten, dass die Einschätzung auf einer 9-stufigen Skala erfolgt (s. Kapitel 5.3). Eine exakte Übereinstimmung ist bei einer so feinen Differenzierung sehr schwer zu erreichen. Dies wird deutlich, wenn man eine Abweichung von 0.5 Skaleneinheiten noch als Übereinstimmung betrachtet. In diesem Fall steigt die prozentuale Übereinstimmung auf 69.8%. Auch dieser Wert ist sicher noch nicht zufriedenstellend. Grundsätzlich wird an dem Mass der prozentualen Übereinstimmung jedoch die Kritik geübt, dass eine Übereinstimmung mit einer gewissen Wahrscheinlichkeit auch rein zufällig auftreten kann und dies in diesem Mass nicht berücksichtigt wird. In der Berechnung von Cohen's Kappa als Mass der Interrater-Reliabilität wird dagegen für diese potenziell zufällig auftretende Übereinstimmung korrigiert. Bei diesem Mass bedeutet die 0, dass die Übereinstimmung zwischen zwei Einschätzungen nicht über das hinausgeht, was rein zufällig erwartet werden kann. Eine 1 bedeutet dagegen über den Zufall hinaus perfekte Übereinstimmung (McHugh, 2012). Berechnet man Cohen's Kappa für die vorliegenden vier Präsentationen, erhält man einen Wert von $\kappa=.198$. Wongpakaran, Wongpakaran, Wedding und Gwet (2013) nennen in ihrem Artikel Benchmark-Werte zur Einschätzung von Cohen's Kappa. Nach den von ihnen genannten Werten ist ein Wert von $\kappa=.198$ als gering einzuschätzen. Aber auch hier ist wieder zu beachten, dass nur exakte Übereinstimmungen berücksichtigt werden. Wertet man auch für die Berechnung von Cohen's Kappa eine Abweichung von 0.5 noch als Übereinstimmung, wird ein Wert von $\kappa=.621$ erreicht. Dieser ist nach Wongpakaran et al. (2013) als gut einzuschätzen.

Bislang wurde die Interrater-Reliabilität auf der Ebene der einzelnen Kriterien betrachtet. Da explizit einzelne Kriterien bewertet werden, ist diese Betrachtungsweise sicherlich wichtig. Genauso wichtig ist aber auch, ob im Endresultat die verschiedenen Evaluatoren basierend auf der Einschätzung der einzelnen Kriterien zu der gleichen Note kommen. Wenn dies gegeben ist, sind Abweichungen in den einzelnen Kriterien ggf. nicht so hoch einzuordnen. Auf Basis der bisherigen vier Präsentationen wird auf Ebene der Noten für die inhaltsbezogenen Aspekte eine prozentuale Übereinstimmung von 75% erreicht, d.h. für drei Präsentationen ergaben die Einschätzungen der beiden Dozenten die gleiche Note. Bei der vierten Präsentation lag eine Abweichung von 0.5 vor (Note 4 vs. Note 4.5). Die Berechnung von Cohen's Kappa ergibt für die Noten einen Wert von $\kappa=.636$ und ist somit auch als gut einzuschätzen.

Zusammengefasst kann gesagt werden, dass die Übereinstimmung zwischen den beiden Evaluatoren noch nicht zufriedenstellend ist. Ein Grund hierfür ist sicherlich darin zu finden, dass die beiden Dozenten unterschiedliche Vorgehensweisen gewählt haben, wie stark die Verhaltensanker bei der Benotung als Basis gedient haben. Dieses Vorgehen müsste bei der weiteren Anwendung der V3 des Bewertungsschemas vereinheitlicht werden. Als zufriedenstellend kann dagegen die Übereinstimmung in den Gesamtnoten betrachtet werden.

6.3 RELEVANZ

Mit der *Relevanz* einer Evaluation wird beschrieben, inwieweit die Aktivität, die evaluiert wird, in den Augen der Studierenden bedeutsam ist (Charlier, 2015b). Am Anfang der Arbeit (vgl. Kapitel 1.2) wurde verdeutlicht, dass das Halten von Präsentationen – oder allgemeiner formuliert, das zielgerichtete Präsentieren bestimmter Inhalte, abgestimmt auf eine bestimmte Zielgruppe – für den Abschluss des Masterstudiums, aber mit grosser Wahrscheinlichkeit auch im späteren Berufsleben für die Studierenden relevant ist. Damit ist die Relevanz der in dieser Arbeit beschriebenen Evaluation vermutlich stärker gegeben als für manche schriftliche Evaluation. Die Frage ist jedoch, inwieweit den Studierenden diese Relevanz auch bewusst ist. Es besteht sicherlich Verbesserungsbedarf dahingehend, dass die Relevanz von Präsentationen sowie die Transfermöglichkeiten der erworbenen Kompetenzen den Studierenden vorgängig vermittelt werden.

6.4 QUANTITATIVE VERÄNDERUNG DER BEWERTUNG

Im Kapitel 4.2 wurde beschrieben, dass bei der Verwendung der V2 des Bewertungsschemas relativ schlechte Noten für die Präsentationen vergeben wurden. Als eine Erklärung dafür wurde das Anspruchsniveau der Verhaltensanker gesehen. Durch den Einbezug mehrerer Experten in die Formulierung der Verhaltensanker der V3 sollte diesem Problem begegnet werden. Im ersten Test der V3 stellt sich daher nun die Frage, ob sich das Niveau der Noten durch die Überarbeitung verändert hat. Natürlich ist der Vergleich der Noten in zwei verschiedenen Seminaren mit Vorsicht zu interpretieren. Es wurden nicht ausschliesslich die gleichen Studierenden bewertet. Zudem ist der Kontext in jedem Seminar ein anderer und, wie oben beschrieben, haben sich die Studierenden in der Zwischenzeit sicher auch weiterentwickelt. Dennoch kann der Vergleich der Noten einen ersten Aufschluss über die Erreichung des Ziels geben. Zu beachten ist, dass auch hier zum jetzigen Zeitpunkt des Semesters (November 2016) nur von 7 Studierenden die Noten in die Betrachtung einfließen.

Um einen sinnvollen Vergleich ziehen zu können, wurden für das laufende Seminar die Noten nach der in Kapitel 4.2 beschriebenen Formel berechnet. Hier zeigt sich ein ähnlich ernüchterndes Bild wie im Frühlingsemester 2016: von 7 Studierenden erreichten ohne Anwendung eines Korrekturfaktors nur 2 Studierende eine Note besser als 4 – im Frühlingsemester waren es 2 von 8. Dieses Ergebnis legt die Vermutung nahe, dass die relativ schlechten Noten nicht ausschliesslich auf ein unpassendes Anspruchsniveau der Verhaltensanker zurück zu führen ist, sondern dass ein anderer Grund relevant ist. Aufschlussreich ist hier die Betrachtung der Noten, die auf Basis des Medians ermittelt wurden: Hier erreichen alle 7 Studierenden eine Note von 4 oder besser. Der grosse Unterschied in den Noten je nach Art der Ermittlung ist leicht zu erklären: Der in Kapitel 4.2 dargestellte Formel liegt die Annahme zugrunde, dass Studierende eine Prüfung dann bestehen, wenn sie mehr als die Hälfte der möglichen Maximal-Punktzahl erreichen. Bei dem der V3 zugrundeliegenden Bewertungsschema bedeutet das, dass die Note 4 erreicht wird, wenn im Schnitt für jede Kategorie eine Bewertung von 2 erreicht wird (ohne Berücksichtigung der Zusatzpunkte). Für die Ermittlung der Note auf Basis des Medians wird dagegen eine andere Annahme getroffen: Wie aus [Tabelle 3](#) und der Schilderung in Kapitel 5.3 hervorgeht, wird die Bewertung der einzelnen Kriterien mit einer 1 als genügend für eine Note von 4 oder besser betrachtet. Schlussfolgernd ergibt sich daraus, dass also weniger das Anspruchsniveau der Verhaltensanker, als vielmehr die Art der Notenermittlung verantwortlich war für die verhältnismässig schlechten Resultate bei der Anwendung der V2. Eine Kalibrierung der Verhaltensanker basierend auf der Meinung mehrerer Experten war sicherlich dennoch sehr nützlich. Entscheidend ist in Zukunft aber sicherlich die für eine Ordinalskala zulässige Verwendung des Medians zur Notenermittlung anstelle einer Summenberechnung.

6.5 QUALITATIVE EINSCHÄTZUNG DER HANDHABBARKEIT DES BEWERTUNGSSCHEMAS

Für die Bewertung eines Bewertungsschemas sind nicht nur quantitative Kriterien relevant, sondern auch, wie gut sich das Schema in der alltäglichen Verwendung tatsächlich einsetzen lässt. Daher sollen im Folgenden einzelne qualitative Beobachtungen in der Verwendung der V3 des Bewertungsschemas geschildert werden.

Zu bemerken ist an dieser Stelle sicherlich zunächst die verhältnismässig hohe Komplexität des Schemas, die die Anwendung nicht immer ganz einfach macht. Dies ist auf verschiedene Aspekte zurück zu führen. Im Vergleich zur V2 hat sich verändert, dass die Kriterien nicht mehr dem vermeintlichen Ablauf der Präsentation folgen, sondern übergreifend über die gesamte Präsentation zu bewerten sind: Auch am Ende einer Präsentation können z.B. noch relevante Begriffe richtig oder falsch verwendet werden, was für das Kriterium „Wissen“ zu beachten wäre. Für die Evaluatoren

steigt damit die Anforderung, während der gesamten Präsentation alle Kriterien im Blick zu haben. In der Anwendungspraxis hat sich gezeigt, dass dies so explizit wiederum kaum möglich ist. Vielmehr ist es praktikabler, während des Vortrags Beobachtungen zu notieren und diese erst anschliessend den Kriterien zuzuordnen und zu einer Bewertung der einzelnen Kriterien zu kommen. Auf theoretischer Ebene ist dies sicherlich ein Vorteil des Bewertungsschemas, da so der Grundsatz der Trennung von Beobachtung, Bewertung und Benotung berücksichtigt wird (Charlier, 2015b). Für die Evaluatoren bedeutet dies jedoch, nach der Präsentation noch einmal Zeit investieren zu müssen, die Bewertung und Benotung vorzunehmen – im Semesteralltag nicht immer ganz einfach.

Auch hat sich im Vergleich zur V2 die Anzahl der zu bewertenden Kriterien erhöht: Insgesamt muss für 24 Kriterien eine Einschätzung vorgenommen werden. Hinzu kommt, dass die Verhaltensanker relativ ausführlich gehalten sind. Dies hat den Vorteil, dass sie sehr konkret sind. Allerdings bringt es auch den Nachteil mit sich, dass es mehrere Anwendungen braucht, bis man wirklich mit den Verhaltensankern vertraut ist. Bei der Bewertung und Benotung ist es unerlässlich, die Verhaltensanker durchzulesen. Auch dies führt jedoch dazu, dass von den Evaluatoren ein gewisser Zeitaufwand gefordert wird, um zu einer Benotung der Kriterien zu kommen.

Grundsätzlich muss man sich bei der Anwendung der V3 des Bewertungsschemas also bewusst machen, dass es nicht möglich ist, die Seminarsitzung mit einem fertigen Evaluationsbogen zu verlassen, sondern das im Anschluss noch eine intensive Auseinandersetzung mit der Bewertung der Präsentation notwendig ist. Um Erinnerungsverzerrungen zu vermeiden, sollte dies zudem möglichst rasch nach dem Seminar erfolgen. Bedenkt man jedoch, dass jede Evaluation Zeit und Aufwand erfordert (im Vergleich sind hier z.B. Klausurkorrekturen zu nennen), so sind diese Zeitinvestition und der Aufwand jedoch gerechtfertigt und auch nicht als unverhältnismässig hoch einzuschätzen.

7. ZUSAMMENFASSUNG

7.1 INHALTLICHES FAZIT

Das Ziel der vorliegenden Arbeit bestand darin, das bestehende Bewertungsschema für Präsentationen in Masterseminaren dahingehend zu überarbeiten, dass (1) die Bewertungskriterien auf Lernzielen basieren, (2) die Verhaltensanker für die Bewertungskriterien hinsichtlich ihres Anspruchsniveaus angepasst werden, (3) die Bewertungskriterien so verallgemeinert werden, dass sie in verschiedenen Masterseminaren des psychologischen Departments und gegebenenfalls auch darüber hinaus eingesetzt werden können und (4) die Notenberechnung auf die zugrundeliegende Ordinalskala angepasst wird. Um dieses Ziel zu erreichen wurden in der überarbeiteten Version des Bewertungsschemas (Version 3) mehrere Veränderungen vorgenommen. Die erste Veränderung besteht darin, dass konkrete Lernziele für das Halten von Präsentationen in Masterseminaren formuliert wurden. Daraus wurden anschliessend Bewertungskriterien formuliert. Die Verhaltensanker zu diesen Bewertungskriterien wurden anschliessend mithilfe der Methode kritischer Ereignisse erhoben. Dazu wurden insgesamt 8 Expertinnen und Experten gebeten, konkrete Verhaltensbeispiele für die Extrempole der Ausprägung der Bewertungskriterien zu formulieren. Damit wurde zum einen erreicht, dass das Anspruchsniveau von mehr als einer Person definiert wurde. Zudem wurden Personen aus verschiedenen Bereichen der Psychologie befragt, sodass sich eine Verallgemeinerung der Kriterien ergibt. Zuletzt wurde eine Variante der Notenberechnung entwickelt, der der Median der verschiedenen Bewertungskriterien zugrunde liegt.

Grundsätzlich wurde somit versucht, alle identifizierten Probleme der 2. Version des Bewertungsschemas sowie die definierten Ziele für die 3. Version des Bewertungsschemas zu adressieren. Darüber hinaus zeigt die Betrachtung der verschiedenen Gütekriterien (s. Kapitel 6) ein zufriedenstellendes Ergebnis.

Gleichzeitig hat die erste Anwendung des überarbeiteten Bewertungsschemas aber auch aufgezeigt, an welchen Stellen noch Veränderungsbedarf besteht. Dieser soll abschliessend im folgenden Kapitel beschrieben werden.

7.2 VORSCHLÄGE FÜR WEITERE VERÄNDERUNGEN

Die Werte der Interrater-Reliabilität können zwar als zufriedenstellend betrachtet werden, sind jedoch nicht durchgängig als gut einzuschätzen. Hier besteht daher noch Verbesserungspotenzial. Zum einen kann dies sicherlich dadurch erreicht werden, dass bei einer Bewertung einer Präsentation

durch zwei Evaluatoren ein gemeinsames Vorgehen abgesprochen wird. So ist es sicherlich wichtig, die Verhaltensanker bei der Bewertung der Präsentation stärker als Basis zu nutzen. Gleichzeitig sollte aber auch darüber nachgedacht werden, ob es gegebenenfalls helfen könnte, die Anzahl der Kriterien zu reduzieren. Somit würde die Komplexität der Bewertung reduziert werden und die Evaluatoren wären schneller mit den Kriterien und den dazugehörigen Verhaltensankern vertraut. Wichtig ist hierbei jedoch, die Operationalisierung der Lernziele beizubehalten, die Reduktion also nur so weit vorzunehmen, wie es mit einer Umsetzung der Lernziele in den Kriterien vereinbar wäre.

Bereits jetzt werden den Studierenden die Kriterien am Anfang des Semesters zur Verfügung gestellt, sodass sie als Art schriftliche Verhaltensmodelle dienen und die Vorbereitung der Präsentation leiten können. Jedoch ist zu überlegen, ob nicht eine gemeinsame Diskussion des Schemas einer reinen zur Verfügung Stellung vorzuziehen wäre. So könnte bei den Studierenden ein besseres Verständnis der Kriterien erreicht werden. Zudem könnten Fragen hinsichtlich der Erwartungen geklärt werden. Auch wäre es möglich, die Relevanz der erworbenen Kompetenzen verdeutlichen und somit die wahrgenommene Relevanz der Evaluation zu erhöhen. Daher wird empfohlen, die Kriterien in der ersten Sitzung eines Semesters mit den Studierenden kurz zu diskutieren.

7.3 PERSÖNLICHES FAZIT

Am Ende dieser Arbeit möchte ich hinsichtlich zweier Aspekte ein persönliches Fazit ziehen: (1) hinsichtlich der Auswirkung und des Nutzens dieser Arbeit für die eigene Lehre und (2) hinsichtlich der im Rahmen dieser Arbeit erworbenen Kompetenzen.

(1) Fazit hinsichtlich der Auswirkung und des Nutzens der Arbeit für die eigene Lehre

Wie in Kapitel 1.3 dargestellt, war es für mich ein Anliegen, eine Abschlussarbeit über ein Thema zu schreiben, das eine praktische Relevanz meinen Arbeitsalltag als Dozentin und ggf. auch für andere Dozenten hat. Damit ein Projekt eine praktische Relevanz hat, ist die Anwendung des im Rahmen des Projekts entwickelten Produkts wichtig. Da wir in jedem Semester mindestens ein Seminar im Masterstudiengang halten und in jedem dieser Seminare Präsentationen von Studierenden gehalten werden, wird das Bewertungsschema jedes Semester von neuem angewendet. Neben der Anwendung ist für die praktische Relevanz aber auch entscheidend, dass das Projekt einen Einfluss auf die eigene Lehre hat. Dies ist bei mir in drei Aspekten der Fall.

Zum einen hilft mir das neue Bewertungsschema, die Unsicherheit bei der Notenvergabe zu reduzieren, die ich in Kapitel 1.3 beschrieben habe. Wichtig ist für mich dabei unter anderem

auch der Bezug zu konkretem Verhalten. Auch wenn ich in der Anwendungspraxis durch die Betrachtung der Interrater-Reliabilität (s. Kapitel 6.2) festgestellt habe, dass auch bei dem vermeintlich konkret beobachtbaren Verhalten noch gewisser Interpretationsspielraum bleibt, so bieten die Verhaltensanker doch eine Basis, die jeweilige Beobachtung und Beurteilung zu diskutieren und die Entscheidung zu begründen. Wichtig ist für mich aber auch das Bewusstsein, dass die Verhaltensanker durch die Technik der kritischen Ereignisse zusammen mit anderen Dozierenden entwickelt wurden und somit nicht allein meine subjektive Sichtweise widerspiegeln. Das gibt mir mehr Vertrauen in die Anwendung des Bewertungsschemas.

Zum anderen hilft mir das Bewertungsschema sowie die Fundierung auf den Lernzielen, in der Kommunikation mit den Studierenden unsere Erwartungen an eine gute Präsentation besser beschreiben und begründen zu können. Die bisherigen Bewertungskriterien wirkten auf mich selber zwar nicht willkürlich und dennoch konnte ich nicht fundiert begründen, warum diese Kriterien wichtig sind. Das ist nun durch die Lernziele anders. Darüber hinaus bieten die Verhaltensanker für die Studierende konkrete Anhaltspunkte in der Gestaltung der Präsentation. Als Nebenaspekt zeigte sich in der ersten Anwendung des Bewertungsschemas, dass dies einen Einfluss auf die Qualität der Präsentation haben kann. Bei einer Präsentation wirkte es so, als haben die Studierenden versucht, die Bewertungskriterien nacheinander „abzuhaken“. Auch wenn die starke Orientierung an den Kriterien noch spürbar war und vielleicht noch etwas unnatürlich wirkte, war die Präsentation ausgesprochen gut. Da gut vorstellbar ist, dass die Erfüllung der Kriterien mit jedem Mal natürlicher erfolgt, kann das neue Bewertungsschema also auch einen direkten Einfluss auf die Leistung und den Lernerfolg der Studierenden haben, was ein sehr positiver Nebeneffekt ist.

Als dritter Aspekt ist die Anwendung des Medians zur Notenbildung zu nennen. Auch wenn mir auf theoretischer Ebene bewusst war, dass das Bilden des Mittelwerts über mehrere Kriterien bei einer Ordinalskala unzulässig ist, wendete ich dieses Verfahren bislang bei der Notenbildung an. Diese Arbeit hat dazu geführt, dass ich diese Praxis detaillierter reflektiert und zumindest bei der Evaluation von Präsentationen verändert habe.

(2) Fazit hinsichtlich der im Rahmen der Arbeit erworbenen Kompetenzen

Der erste Aspekt hinsichtlich des Erwerbs von Kompetenzen bezieht sich direkt auf die Anwendung der Ordinalskala. Es war für mich nicht allein die Umstellung auf die Verwendung des Medians, sondern der direkte Vergleich der Noten, die durch die Verwendung von Mittelwert und Median entstanden. Durch die Unterschiede, die sich in den

Noten ergaben, wurde mir erst richtig bewusst, dass es sich nicht nur um ein abstraktes, theoretisches Thema handelt, dem theoretische Überlegungen zu verschiedenen Skalenniveaus zugrunde liegen, sondern dass dieses Thema auch direkte Konsequenzen für die resultierenden Noten hat. Diese Erkenntnis ist für mich eine wichtige Basis, für die Verwendung des Medians zu argumentieren.

Ein zweiter Aspekt bezieht sich auf die Formulierung von Lernzielen. Wie ich im Lernportfolio verdeutlicht habe, war das Verstehen der Bedeutung von Lernzielen sowie deren Verknüpfung mit den gewählten Lehrmethoden und der Gestaltung der Lernevaluation ein wesentlicher Lernerfolg der gesamten Didactic-Ausbildung. Gleichzeitig habe ich im Lernportfolio aber auch verdeutlicht, dass ich das Formulieren von Lernzielen nach wie vor als grosse Herausforderung empfinde. Aus meiner Sicht ist hier Übung ein wesentlicher Faktor, diese Herausforderung Stück für Stück zu meistern. Die vorliegende Abschlussarbeit hat mir nicht nur die Gelegenheit gegeben, für einen Bereich der Lehre Lernziele zu formulieren, sondern auch, mich noch einmal intensiv mit der Theorie zu den Lernzielen, den unterschiedlichen Taxonomien und Kategorien auseinander zu setzen. So konnte ich die praktische Anwendung mit der Theorie verknüpfen. Ich hoffe, dass das für die Zukunft eine gute Basis für die Formulierung von Lernzielen darstellt.

Ein letzter Aspekt, den ich hier nennen möchte, bezieht sich auf die Frage, welche Aspekte für eine Evaluation wichtig sind. Hier hat mir insbesondere der Prozess der Reflexion über die Stärken und Schwächen der verschiedenen Versionen des Bewertungsschemas geholfen, diesen Punkt in Bezug auf die Evaluation von Präsentationen zu vertiefen. Die Reflexion von Stärken und Schwächen setzt eine klare Zielfokussierung voraus, was wiederum zur Folge hat, dass dieses Ziel – oder anders ausgedrückt, dass die Lernziele – klar formuliert werden.

8. LITERATURVERZEICHNIS

- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bell, B. & Cowie, B. (2001). The Characteristics of Formative Assessment in Science Education. *Science Education*, 85, 536–553.
- Bernardin, H. J. & Smith, P. C. (1981). A Clarification of Some Issues Regarding the Development and Use of Behaviorally Anchored Rating Scales (BARS). *Journal of Applied Psychology*, 66(4), 458-463.
- Blickle, G. (2014). Leistungsbeurteilung. In F.W. Nerdinger, G. Blickle & N. Schaper (Hrsg). *Arbeits- und Organisationspsychologie* (3. Aufl., S. 271-289). Berlin: Springer-Verlag.
- Blickle, G. (2014b). Anforderungsanalyse. In F.W. Nerdinger, G. Blickle & N. Schaper (Hrsg). *Arbeits- und Organisationspsychologie* (3. Aufl., S. 207-221). Berlin: Springer-Verlag.
- Bloom, B.S., Hastings, J.T., & Madaus, G.F. (1971). *Handbook on Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill Book Company.
- Bloxham, S. & Boyd, P. (2007). *Developing Effective Assessment in Higher Education: a practical guide*. Maidenhead: Open University Press.
- Butterfield, L. D., Borgen, W. A., Amundson, N. E., Maglio, A.-S. T. (2005). Fifty years of the critical incident technique: 1954–2004 and beyond. *Qualitative Research*, 5(4), 475-496.
- Cardinet, J. (1986). *Pour apprécier le travail des élèves*. Bruxelles : De Boeck-Wesmael.
- Charlier, B. (2015). *Glossaire zu den Unterlagen des Module B. Evaluation des apprentissages. Formation Did@cTIC: note de cours*. Unveröffentlichte Unterrichtsmaterialien, Universität Freiburg.
- Charlier, B. (2015b). *Module B. Evaluation des apprentissages. Formation Did@cTIC: note de cours*. Unveröffentlichte Unterrichtsmaterialien, Universität Freiburg.
- Demmerle, C., Schmidt, J. M., Hess, M., Solga, M. & Ryschka, J. (2011). Verhaltensmodellierung. In J. Ryschka, M. Solga & A. Mattenklott (Hrsg). *Praxishandbuch Personalentwicklung* (S.289-294). Wiesbaden: Gabler

- Döring, K. W. (2008). *Handbuch Lehren und Trainieren in der Weiterbildung*. Weinheim: Beltz Verlag.
- Flanagan, J.C. (1954). The Critical Incident Technique. *Psychological Bulletin* 51(4), 327–358.
- Gadenne, V. (2014). Operationalisierung. In M. A. Wirtz (Hrsg.), *Dorsch - Lexikon der Psychologie* (17. Aufl.). Abgerufen unter: <https://portal.hogrefe.com/dorsch/operationalisierung/>
- Hartig, J., Frey, A. & Jude, N. (2012). Validität. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion*, (S. 143–171). Berlin: Springer.
- Heckhausen, J. & Heckhausen, H. (2010). Motivation und Handeln: Einführung und Überblick. In: J. Heckhausen & H. Heckhausen (Hrsg.), *Motivation und Handeln* (S.1-9). Berlin: Springer-Verlag.
- Krathwohl, D. R. (2010). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, 41(4), 212-218.
- Krathwohl, D. R., Bloom, B. S. & Masia, B. B. (1975). *Taxonomie von Lernzielen im affektiven Bereich*. Weinheim: Beltz Verlag.
- Latham, G. P. & Wexley, K. N. (1977). Behavioral Observation Scales for Performance Appraisal Purposes. *Personnel Psychology*, 30, 255-268.
- Lunenburg, F. C. (2012). Performance Appraisal: Methods and Rating Errors. *International Journal of Scholarly Academic Intellectual Diversity*, 14(1), 1-9.
- Marcus, B. & Schuler, H. (2006). Leistungsbeurteilung. In H. Schuler (Hrsg.). *Lehrbuch der Personalpsychologie* (2. Aufl., S. 433-469). Göttingen: Hogrefe.
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 22(3), 276-282.
- Moser, K. & Schuler, H. (2014). Leistungsbeurteilung. In H. Schuler & K. Moser (Hrsg.). *Lehrbuch Organisationspsychologie* (5. Aufl., S. 177-216). Bern: Verlag Hans Huber.
- Philosophische Fakultät der Universität Freiburg (Schweiz) (2006). *Reglement zur Erlangung des Masters an der Philosophischen Fakultät der Universität Freiburg (Schweiz)*. Universität Freiburg. Abgerufen unter: <http://www3.unifr.ch/lettres/deformation/cours-et-examens/reglements-et-directives/>
- Race, P., Brown, S., & Smith, B. (2005). *500 Tips on Assessment* (2. Aufl.). London: Routledge.

- Schermelleh-Engel, K. & Werner, C. S. (2012). Methoden der Reliabilitätsbestimmung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion*, (S. 119–141). Berlin: Springer.
- Schuler, H. (2006). Arbeits- und Anforderungsanalyse. In H. Schuler (Hrsg.). *Lehrbuch der Personalpsychologie* (2. Aufl., S. 45-68). Göttingen: Hogrefe.
- Schweizerische Universitätskonferenz (2011). *Qualifikationsrahmen für den schweizerischen Hochschulbereich – nqf.ch-HS*. Abgerufen unter:
<http://www.swissuniversities.ch/de/hochschulraum/qualifikationsrahmen/>
- Taylor, P. J., Russ-Eft, D. F. & Chan, D. W. L. (2005). A meta-analytic review of behavior modeling training. *Journal of Applied Psychology*, 90, 692-709.
- Van Ginkel, S., Gulikers, J., Biemans, H., & Mulder, M. (2016). Fostering oral presentation performance: does the quality of feedback differ when provided by the teacher, peers or peers guided by tutor? *Assessment & Evaluation in Higher Education*, DOI: 10.1080/02602938.2016.1212984
- Wongpakaran, N., Wongpakaran, T., Wedding, D., & Gwet, K. L. (2013). A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC Medical Research Methodology*, 13(61), 1-7. doi:10.1186/1471-2288-13-61.
- Wyrsh, A. (2014). *Basismodul A – Lehren und Lernen*. Unveröffentlichte Unterrichtsmaterialien, Universität Fribourg.
- Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education*, 45, 477-501.

9. ANHANG

ANHANG I: BEWERTUNGSSCHEMA FÜR PRÄSENTATIONEN VERSION 1	II
ANHANG II: BEWERTUNGSSCHEMA FÜR PRÄSENTATIONEN VERSION 2	III
ANHANG II.1: BEWERTUNGSBOGEN	III
ANHANG II.2: VERHALTENSANKER	IV
ANHANG II.3: VERHALTENSANKER FÜR STUDIERENDE	X
ANHANG III: MATERIALIEN ZUR ERHEBUNG DER VERHALTENSANKER	XII
ANHANG III.1: MAIL ZUR ERKLÄRUNG DES VORGEHENS	XII
ANHANG III.2: ERHEBUNGSBOGEN FÜR DEN BEREICH «INHALT»	XIII
ANHANG III.3: ERHEBUNGSBOGEN FÜR DEN BEREICH «PRÄSENTATIONSKOMPETENZ»	XIV
ANHANG IV: BEWERTUNGSSCHEMA FÜR PRÄSENTATIONEN VERSION 3	XV
ANHANG II.1: BEWERTUNGSBOGEN	XV
ANHANG II.2: VERHALTENSANKER	XVIII

ANHANG I: BEWERTUNGSSCHEMA FÜR PRÄSENTATIONEN VERSION 1

Referenten:

Datum:

Thema:

INHALT

0	1	2	3	4
---	---	---	---	---

- Hinführung/Einführung & Einordnung, Vorwissen der Zielgruppe berücksichtigen
- Thema für den jeweiligen Zweck bearbeiten, eigenständige Argumentation
- Theoretische Einbettung, Bezüge herstellen
- Vor- und Nachteile der methodischen Vorgehens
- Verwendung anschaulicher Beispiele, Praxisbezug herstellen
- Umgang mit Quellen

STRUKTUR

0	1	2	3	4
---	---	---	---	---

- „roter Faden“, klare Gliederung
- Gliederung evtl. visualisieren, besonders wenn es sich um längere Vorträge handelt
- Zusammenfassungen (u.a. Erkenntnisgewinn)

FORMULIERUNG

0	1	2	3	4
---	---	---	---	---

- verständliche Sprache (deutlich, laut, Tempo, Modulation)
- Konzepte erläutern
- freier Vortrag

MEDIEN/HILFSMITTEL

0	1	2	3	4
---	---	---	---	---

- klare, übersichtliche & interessante Gestaltung (Konsistenz & Abwechslung)
- adäquater Einsatz, Verhältnis Wort & Visualisierung

EINBEZIEHEN DER ZUHÖRENDEN

0	1	2	3	4
---	---	---	---	---

- Blickkontakt, Fragen stellen
- Beantwortung von Fragen
- Anregung zur Weiterbeschäftigung

ZEITMANAGEMENT

0	1	2	3	4
---	---	---	---	---

- Zeitlichen Rahmen nicht überschreiten (ca. 10 Minuten pro Person)

DISKUSSIONSLEITUNG

0	1	2	3	4
---	---	---	---	---

- Strukturierung der Diskussion / Zielführung
- Leitfragen
- Moderation (keine Vorgabe der Inhalte)
- Einbezug aller Teilnehmer

ANHANG II: BEWERTUNGSSCHEMA FÜR PRÄSENTATIONEN VERSION 2**II.1: Bewertungsbogen****Referenten:** _____ **Datum:** _____**Thema:** _____ **Gesamtpunktzahl:** _____**INHALT**

<u>AUFHÄNGER</u>						
EINFÜHRUNG IN DAS THEMA		0	1	2	3	4
<u>PRÄSENTATION WESENTLICHER BESTANDTEILE DES PAPERS</u>						
THEORIE & FRAGESTELLUNG		0	1	2	3	4
METHODE		0	1	2	3	4
ERGEBNISSE		0	1	2	3	4
DISKUSSION		0	1	2	3	4
<u>FAZIT UND ZUSAMMENFASSUNG</u>						
ERKENNTNIS- GEWINN		0	1	2	3	4
PRAXIS- RELEVANZ		0	1	2	3	4
<u>EINBETTUNG IN DEN GESAMTKONTEXT</u>						
VERKNÜPFUNG GRUNDLAGEN- TEXT		0	1	2	3	4
VERNKNÜPFUNG BISH. INHALTE		0	1	2	3	4

<u>ZIELGRUPPENSPEZIFISCHE AUFBEREITUNG</u>						
BERÜCKSICHTIGUNG VORWISSEN		0	1	2	3	4
VERANSCHAULICHUNG / BEISPIELE		0	1	2	3	4
<u>STRUKTUR</u>						
ROTER FADEN		0	1	2	3	4
KONSISTENZ DER VERWENDETEN BEGRIFFE		0	1	2	3	4
ZEITL. RAHMEN		0	1	2	3	4

PUNKTZAHL INHALT: _____

PRÄSENTATIONSKOMPETENZ

<u>MEDIEN / HILFSMITTEL</u>						
GESTALTUNG / (AUSWAHL)		0	1	2	3	4
VERHÄLTNIS VISUALISIERUNG / WORT		0	1	2	3	4
<u>VORTRAG</u>						
SPRACHE UND HALTUNG		0	1	2	3	4
FREIES SPRECHEN		0	1	2	3	4
<u>EINBEZUG DER ZUHÖRER</u>						
BLICKKONTAKT		0	1	2	3	4
EINBEZUG DURCH FRAGEN		0	1	2	3	4
BEANTWORTUNG VON FRAGEN		0	1	2	3	4
<u>GESCHWINDIGKEIT</u>						
GESCHWINDIG- KEIT DER PRÄSENTATION		0	1	2	3	4

EXTRAPUNKTE PRÄSENTATION: _____

II.2: Verhaltensanker

INHALT

Leitfrage für die Referatsausarbeitung : was sollen die Mit-Studierenden an zentralen Erkenntnissen und Ergebnissen mitnehmen?

Leitmotto für die Referatsausarbeitung: *Eigenständige* Bearbeitung des vorgegebenen Papers

	0 (ungenügend)	1 (eher schwach)	2 (mittel)	3 (gut)	4 (exzellent)
AUFHÄNGER					
EINFÜHRUNG IN DAS THEMA	<ul style="list-style-type: none"> • Es erfolgt keine Einführung in das Thema, die Zuhörer werden nicht „abgeholt“; „Kaltstart“ 		<ul style="list-style-type: none"> • Es erfolgt ein „sachlicher“ Einstieg, die Zuhörer werden abgeholt, ohne dass spezifisches Interesse für das Thema geweckt wird 		<ul style="list-style-type: none"> • Der Einstieg in das Thema ist motivierend und weckt Interesse für die Präsentation
PRÄSENTATION WESENTLICHER BESTANDTEILE DES PAPERS					
THEORIE & FRAGESTELLUNG	<ul style="list-style-type: none"> • Die Fragestellung des Artikels wird nicht klar und nachvollziehbar; die theoretische Grundlage wird nicht in Verbindung mit der Fragestellung gebracht 	<ul style="list-style-type: none"> • Die Fragestellung wird erwähnt/die Hypothesen werden aufgelistet/genannt, ohne dass sie erklärt oder hergeleitet werden 	<ul style="list-style-type: none"> • Die Fragestellung und die dazugehörigen Hypothesen werden hergeleitet, ohne dass eine direkte Verknüpfung zur theoretischen Grundlage gezogen wird 	<ul style="list-style-type: none"> • Die Fragestellung und Hypothesen werden basierend auf der Theorie dargestellt; Theoretische Grundlage wird jedoch nicht selektiert / zielgerichtet ausgewählt 	<ul style="list-style-type: none"> • Die Fragestellung und die dazugehörigen Hypothesen werden präzise auf Basis der Theorie hergeleitet; Die Präsentation der Theorie erfolgt selektiv / zielgerichtet
METHODE	<ul style="list-style-type: none"> • Weder die Stichprobe noch das Studiendesign und die Durchführung werden nachvollziehbar • Fehlerhafte Darstellung der angewendeten Methode 		<ul style="list-style-type: none"> • Korrekte, aber unselektierte Darstellung der Stichprobe sowie der Durchführung und des Studiendesigns 		<ul style="list-style-type: none"> • Die zum Verständnis notwendigen Informationen zu Stichprobe, Durchführung und Studiendesign werden präzise, selektiv und korrekt dargestellt
ERGEBNISSE	<ul style="list-style-type: none"> • Die Ergebnisse werden nicht klar oder fehlerhaft dargestellt 	<ul style="list-style-type: none"> • Die Ergebnisdarstellung ist klar, enthält jedoch Fehler 	<ul style="list-style-type: none"> • Die Ergebnisse werden korrekt und klar dargestellt, jedoch erfolgt keine Schwerpunktsetzung 	<ul style="list-style-type: none"> • Es werden ausgewählte Ergebnisse präsentiert, die Auswahl ist jedoch nicht zielführend oder begründet 	<ul style="list-style-type: none"> • Klare Darstellung ausgewählter Ergebnisse, die im Zusammenhang mit den abgeleiteten Erkenntnissen stehen
DISKUSSION	<ul style="list-style-type: none"> • Die Vor- und Nachteile des methodischen Vorgehens werden nicht kritisch reflektiert 		<ul style="list-style-type: none"> • Es werden ausschliesslich Vor- und Nachteile des methodischen Vorgehens aus dem Paper übernommen 		<ul style="list-style-type: none"> • Die Vor- und Nachteile des methodischen Vorgehens werden präzise beschrieben und basieren auf eigenständigen Überlegungen

FAZIT UND ZUSAMMENFASSUNG					
ERKENNTNISGEWINN	<ul style="list-style-type: none"> • Es erfolgt keine Zusammenfassung der relevanten Aspekte des Papers 		<ul style="list-style-type: none"> • Es werden einzelne Punkte zusammengefasst, ohne dass diese zweckgerichtet ausgewählt sind 		<ul style="list-style-type: none"> • Es werden die für die eigene Argumentation relevanten Punkte kurz und präzise dargestellt
PRAXISRELEVANZ	<ul style="list-style-type: none"> • Die Praxisrelevanz der diskutierten Inhalte wird nicht verdeutlicht 		<ul style="list-style-type: none"> • Zur Verdeutlichung der Praxisrelevanz wird ausschliesslich auf die im Paper genannten Punkte zurückgegriffen 		<ul style="list-style-type: none"> • Die Verdeutlichung der Praxisrelevanz erfolgt auf Basis einer eigenen Schwerpunktsetzung und Argumentation
EINBETTUNG IN DEN GESAMTKONTEXT					
VERKNÜPFUNG GRUNDLAGENTEXT	<ul style="list-style-type: none"> • Die Präsentation des Papers steht für sich allein, es wird keine Verknüpfung zum Grundlagentext hergestellt 		<ul style="list-style-type: none"> • Es werden einzelne Querverweise gezogen, ohne dass diese sinnvoll in den Kontext der Präsentation eingebunden sind; Die Studierenden profitieren von diesen Querverweisen nicht 		<ul style="list-style-type: none"> • Der Bezug zum Grundlagentext wird klar und präzise herausgearbeitet und sinnvoll in die Präsentation integriert; Die dargestellten Inhalte werden so klarer
VERKNÜPFUNG BISH. INHALTE	<ul style="list-style-type: none"> • Es werden keine inhaltlichen Verknüpfungen zu vorher diskutierten Themen hergestellt 		<ul style="list-style-type: none"> • Es erfolgen einzelne Querverweise zu bereits diskutierten Themen, ohne dass diese sinnvoll in die Präsentation eingebettet sind / ohne dass die Präsentation inhaltlich davon profitiert 		<ul style="list-style-type: none"> • Das Thema wird sinnvoll in den Gesamt-Seminar-Kontext eingebettet; Die Bezüge zu vorher diskutierten Themen erfolgen an sinnvollen Stellen
ZIELGRUPPENSPEZIFISCHE AUFBEREITUNG					
BERÜCKSICHTIGUNG VORWISSEN	<ul style="list-style-type: none"> • Die Präsentation wird weder inhaltlich noch methodisch auf das Vorwissen der Zielgruppe angepasst (d.h. zu viel oder zu wenig Erklärung) 		<ul style="list-style-type: none"> • Einzelne Aspekte werden definiert/erklärt, allerdings bleiben andere Begriffe offen oder es werden einzelne Begriffe erklärt, die als bekannt vorausgesetzt werden können 		<ul style="list-style-type: none"> • Sowohl inhaltlich als auch methodisch wird das Vorwissen der Zielgruppe berücksichtigt; Unbekannte Begriffe und Methoden werden definiert/erklärt, bekannte als gegeben vorausgesetzt
VERANSCHAULICHUNG / BEISPIELE	<ul style="list-style-type: none"> • Es erfolgt keine Veranschaulichung der dargestellten Inhalte anhand von Beispielen 		<ul style="list-style-type: none"> • Es werden Beispiele zur Verdeutlichung herangezogen, die im Text genannt werden 		<ul style="list-style-type: none"> • An sinnvollen Stellen werden selbstgewählte Beispiele zur Verdeutlichung der Inhalte verwendet

STRUKTUR					
ROTER FADEN	<ul style="list-style-type: none"> Die Inhalte der Präsentation stehen lose nebeneinander, es ergibt sich kein roter Faden 		<ul style="list-style-type: none"> Die Inhalte der Präsentation werden an einzelnen Stellen sinnvoll miteinander verknüpft, an anderen Stellen stehen sie lose nebeneinander 		<ul style="list-style-type: none"> Die Inhalte der Präsentation folgen einem klaren roten Faden, sie sind sinnvoll miteinander verknüpft
KONSISTENZ DER VERWENDETEN BEGRIFFE	<ul style="list-style-type: none"> Die Bezeichnungen für zentrale Konzepte sind nicht kohärent, es werden verschiedene Begriffe verwendet. 		<ul style="list-style-type: none"> Die Begriffswahl für die zentralen Konzepte ist konsistent, es gibt jedoch Uneinheitlichkeiten in der Verwendung anderer Begriffe 		<ul style="list-style-type: none"> Die eingeführten Begriffe werden konsistent und korrekt verwendet
ZEITLICHER RAHMEN	<ul style="list-style-type: none"> Die Dauer der Präsentation entspricht nicht der Vorgabe (+/- 10 min Abweichung nach oben oder unten) 		<ul style="list-style-type: none"> Es gibt leichte Abweichungen von der vorgegebenen Dauer (+/- 5 min) 		<ul style="list-style-type: none"> Die Dauer der Präsentation entspricht der Vorgabe

Anmerkung : wenn zwei einzelne Punkte genannt werden, handelt es sich um eine « Oder »-Verknüpfung

PRÄSENTATIONSKOMPETENZ

Für jede Oberkategorie, in der im Schnitt 3 Punkte erreicht werden, wird für die Gesamt-Punktzahl ein Extrapunkt vergeben (→ 0-4 Extrapunkte möglich).

	0 (ungenügend)	1 (eher schwach)	2 (mittel)	3 (gut)	4 (exzellent)
MEDIEN / HILFSMITTEL					
GESTALTUNG / (AUSWAHL)	<ul style="list-style-type: none"> Die Gestaltung der Visualisierung ist unübersichtlich, nicht gut lesbar oder unangemessen (z.B. „überladen“, zu verspielt...) 		<ul style="list-style-type: none"> Die Gestaltung ist „sachlich“ passend (gut lesbar, übersichtlich, angemessen), jedoch wenig originell; Beschränkung auf PPT 		<ul style="list-style-type: none"> Die Gestaltung der Visualisierung ist angemessen und originell; es werden andere Hilfsmittel als PPT passend eingesetzt
VERHÄLTNIS VISUALISIERUNG / WORT	<ul style="list-style-type: none"> Alle Aspekte des mündlichen Vortrags sind auch auf den Visualisierungen zu finden; Vortrag geht nicht über die Visualisierung hinaus Mündlicher Vortrag passt nicht zu dem, was auf den Visualisierungen steht 		<ul style="list-style-type: none"> Übereinstimmung von Visualisierung und mündl. Vortrag und grundsätzlich gutes Verhältnis, jedoch immer wieder „Konkurrenzverhältnis“, d.h. Zuhörer muss entweder lesen oder zuhören 		<ul style="list-style-type: none"> Die zentralen Punkte des Vortrags sind auf den Folien zu finden, der Vortrag greift diese kohärent auf, geht jedoch über das hinaus, was visualisiert wurde

<u>VORTRAG</u>					
SPRACHE UND HALTUNG	<ul style="list-style-type: none"> • Es wird zu schnell/langsam gesprochen, die Sprache ist wenig verständlich, zu laut/leise und ohne Modulation; Die Körperhaltung ist unangemessen 		<ul style="list-style-type: none"> • Insgesamt ergeben Sprache und Haltung ein gutes Gesamtbild, jedoch sind einzelne Punkte verbesserungswürdig 		<ul style="list-style-type: none"> • Die Geschwindigkeit und Lautstärke der Sprache sind angemessen, das Gesprochene ist gut zu verstehen; Sprachmodulation wird angemessen eingesetzt die Körperhaltung ist passend
FREIES SPRECHEN	<ul style="list-style-type: none"> • Der Vortrag wird abgelesen 		<ul style="list-style-type: none"> • Insgesamt wird der Vortrag frei gehalten, jedoch sind Hilfsmittel notwendig, damit der Faden nicht verloren wird 		<ul style="list-style-type: none"> • Der Vortrag wird völlig frei gehalten
<u>EINBEZUG DER ZUHÖRER</u>					
BLICKKONTAKT	<ul style="list-style-type: none"> • Während des Vortrags besteht kein Blickkontakt zu den Zuhörern, es wird ausschliesslich auf Hilfsmittel, Computer oder Leinwand geschaut 		<ul style="list-style-type: none"> • Der Blickkontakt zu den Zuhörern wird immer wieder gesucht, bricht jedoch auch immer wieder ab; es werden nicht <u>alle</u> Zuhörer angeschaut 		<ul style="list-style-type: none"> • Während des Vortrags besteht die meiste Zeit ein direkter Blickkontakt zu <u>allen</u> Zuhörern (d.h. es wird nicht nur eine Seite des Raums angeschaut o.ä.)
EINBEZUG DURCH FRAGEN	<ul style="list-style-type: none"> • Die Zuhörer werde nicht durch Fragen eingebunden 		<ul style="list-style-type: none"> • Es werden zwar immer wieder Fragen gestellt, jedoch ohne, dass die Antworten der Zuhörer wirklich abgewartet und aufgegriffen werden 		<ul style="list-style-type: none"> • An passenden Stellen werden die Zuhörer durch Fragen aktiviert; es wird genügend Zeit gegeben, um Antworten zu finden; diese werden angemessen und wertschätzend aufgegriffen und weiterverarbeitet
BEANTWORTUNG VON FRAGEN	<ul style="list-style-type: none"> • Zuhörer werden nicht ermutigt, Fragen zu stellen; aufkommende Fragen werden unsicher beantwortet 		<ul style="list-style-type: none"> • Mit aufkommenden Fragen wird grundsätzlich souverän umgegangen; jedoch wird den Zuhörern nicht explizit die Möglichkeit gegeben, Fragen zu stellen; Fragen werden eher vermieden 		<ul style="list-style-type: none"> • Die Zuhörer werden ermutigt, Fragen zu stellen; mit gestellten Fragen wird souverän umgegangen
<u>GESCHWINDIGKEIT</u>					
GESCHWINDIGKEIT DER PRÄSENTATION	<ul style="list-style-type: none"> • Die Geschwindigkeit der Präsentation ist durchgängig zu schnell oder zu langsam 		<ul style="list-style-type: none"> • Grundsätzlich ist die Geschwindigkeit i.O.; jedoch wird über wichtige Punkte immer wieder zu schnell hinweggegangen, an unwichtigen Punkte zu lange verharret 		<ul style="list-style-type: none"> • Die Geschwindigkeit der Präsentation ist passend

II.3: Verhaltensanker für Studierende

INHALT

Leitfrage für die Referatsausarbeitung : was sollen die Mit-Studierenden an zentralen Erkenntnissen und Ergebnissen mitnehmen?

Leitmotto für die Referatsausarbeitung: *Eigenständige* Bearbeitung des vorgegebenen Papers

	++ (exzellent)
AUFHÄNGER	
EINFÜHRUNG IN DAS THEMA	♦ Der Einstieg in das Thema ist motivierend und weckt Interesse für die Präsentation
PRÄSENTATION WESENTLICHER BESTANDTEILE DES PAPERS	
THEORIE & FRAGESTELLUNG	♦ Die Fragestellung und die dazugehörigen Hypothesen werden präzise auf Basis der Theorie hergeleitet; Die Präsentation der Theorie erfolgt selektiv / zielgerichtet
METHODE	♦ Die zum Verständnis notwendigen Informationen zu Stichprobe, Durchführung und Studiendesign werden präzise, selektiv und korrekt dargestellt
ERGEBNISSE	♦ Klare Darstellung ausgewählter Ergebnisse, die im Zusammenhang mit den abgeleiteten Erkenntnissen stehen
DISKUSSION	♦ Die Vor- und Nachteile des methodischen Vorgehens werden präzise beschrieben und basieren auf eigenständigen Überlegungen
FAZIT UND ZUSAMMENFASSUNG	
ERKENNTNISGEWINN	♦ Es werden die für die eigene Argumentation relevanten Punkte kurz und präzise dargestellt
PRAXISRELEVANZ	♦ Die Verdeutlichung der Praxisrelevanz erfolgt auf Basis einer eigenen Schwerpunktsetzung und Argumentation
EINBETTUNG IN DEN GESAMTKONTEXT	
VERKNÜPFUNG GRUNDLAGENTEXT	♦ Der Bezug zum Grundlagentext wird klar und präzise heraus-gearbeitet und sinnvoll in die Präsentation integriert; Die dargestellten Inhalte werden so klarer
VERKNÜPFUNG BISH. INHALTE	♦ Das Thema wird sinnvoll in den Gesamt-Seminar-Kontext eingebettet; Die Bezüge zu vorher diskutierten Themen erfolgen an sinnvollen Stellen
ZIELGRUPPENSPEZIFISCHE AUFBEREITUNG	
BERÜCKSICHTIGUNG VORWISSEN	♦ Sowohl inhaltlich als auch methodisch wird das Vorwissen der Zielgruppe berücksichtigt; Unbekannte Begriffe und Methoden werden definiert/erklärt, bekannte als gegeben vorausgesetzt
VERANSCHAULICHUNG / BEISPIELE	♦ An sinnvollen Stellen werden selbstgewählte Beispiele zur Verdeutlichung der Inhalte verwendet
STRUKTUR	
ROTER FADEN	♦ Die Inhalte der Präsentation folgen einem klaren roten Faden, sie sind sinnvoll miteinander verknüpft
KONSISTENZ DER VERWENDETEN BEGRIFFE	♦ Die eingeführten Begriffe werden konsistent und korrekt verwendet
ZEITLICHER RAHMEN	♦ Die Dauer der Präsentation entspricht der Vorgabe

PRÄSENTATIONSKOMPETENZ

Für jede Oberkategorie, in der im Schnitt 3 Punkte (max. 4 Punkte) erreicht werden, wird für die Gesamt-Punktzahl ein Extrapunkt vergeben (→ 0-4 Extrapunkte möglich).

	4 (exzellent)
<u>MEDIEN / HILFSMITTEL</u>	
GESTALTUNG / (AUSWAHL)	♦ Die Gestaltung der Visualisierung ist angemessen und originell; es werden andere Hilfsmittel als PPT passend eingesetzt
VERHÄLTNIS VISUALISIERUNG / WORT	♦ Die zentralen Punkte des Vortrags sind auf den Folien zu finden, der Vortrag greift diese kohärent auf, geht jedoch über das hinaus, was visualisiert wurde
<u>VORTRAG</u>	
SPRACHE UND HALTUNG	♦ Die Geschwindigkeit und Lautstärke der Sprache sind angemessen, das Gesprochene ist gut zu verstehen; Sprach-modulation wird angemessen eingesetzt die Körperhaltung ist passend
FREIES SPRECHEN	♦ Der Vortrag wird völlig frei gehalten
<u>EINBEZUG DER ZUHÖRER</u>	
BLICKKONTAKT	♦ Während des Vortrags besteht die meiste Zeit ein direkter Blickkontakt zu <u>allen</u> Zuhörern (d.h. es wird nicht nur eine Seite des Raums angeschaut o.ä.)
EINBEZUG DURCH FRAGEN	♦ An passenden Stellen werden die Zuhörer durch Fragen aktiviert; es wird genügend Zeit gegeben, um Antworten zu finden; diese werden angemessen und wertschätzend aufgegriffen und weiterverarbeitet
BEANTWORTUNG VON FRAGEN	♦ Die Zuhörer werden ermutigt, Fragen zu stellen; mit gestellten Fragen wird souverän umgegangen
<u>GESCHWINDIGKEIT</u>	
GESCHWINDIGKEIT DER PRÄSENTATION	♦ Die Geschwindigkeit der Präsentation ist passend

ANHANG III: MATERIALIEN ZUR ERHEBUNG DER VERHALTENSANKER

III.1: Mail zur Erklärung des Vorgehens

Liebe/Lieber...

Herzlichen Dank, dass du mich bei der Erstellung von Kriterien zur Bewertung von Präsentationen – und somit auch bei der Erstellung meiner Didactic-Abschlussarbeit - als Expertin/Experte unterstützt!

Wie du ja weisst, versuche ich, im Rahmen dieser Abschlussarbeit ein neues Bewertungsschema für die Evaluation von Präsentationen in Masterseminaren zu entwickeln.

Die Bewertungskriterien habe ich dafür aus Lernzielen abgeleitet. Der nächste Schritt besteht nun darin, Verhaltensanker zu formulieren, anhand derer die einzelnen Kriterien eingeschätzt werden können. Anders ausgedrückt soll die Frage beantwortet werden: Wodurch zeigt sich im konkreten Verhalten eine gute Erfüllung dieses Kriteriums? Wichtig ist hierbei, die Anker der Einstufungsskala als beobachtbares Verhalten zu formulieren. Dabei soll Verhalten identifiziert werden, dass in der jeweiligen Situation besonders ineffektiv oder effektiv war.

Im Anhang sende ich dir eine Tabelle mit den relevanten Bewertungskriterien. Diese sind in verschiedene Bereiche gegliedert, denen jeweils ein Lernziel zugrunde liegt. Nun geht Dich möchte ich nun um Folgendes bitten:

Stelle dir zu jedem dieser Bewertungskriterien einen Studenten/eine Studentin vor, die während einer Präsentation dieses Kriterium besonders schlecht oder aussergewöhnlich gut erfüllt haben. Welches konkrete Verhalten haben diese Studierenden gezeigt, damit ihr Verhalten besonders ineffektiv oder effektiv für die Erfüllung des Kriteriums war? Beschreibe dieses Verhalten bitte so genau wie möglich und gib auch an, warum dieses Verhalten besonders ineffektiv/effektiv war und welche konkreten Konsequenzen das jeweilige Verhalten hatte.

Eventuell kann es hilfreich sein, sich eine mittlere Ausprägung des Verhaltens vorzustellen, um besonders ineffektives/effektives Verhalten zu identifizieren. Wenn dir dies hilft, schreibe diese mittlere Ausprägung gerne auf, dann kann ich auch sie später als Verankerung verwenden.

Um dir eine Vorstellung von dem «Zielprodukt» zu geben, füge ich dir ein Beispiel eines alten Bewertungskriteriums mit einer fünf-stufigen Bewertungsskala ein. Die Abstufung der Skala ist für die Formulierung der Extrempunkte jedoch unerheblich.

	0 (ungenügend)	1 (eher schwach)	2 (mittel)	3 (gut)	4 (exzellent)
EINFÜHRUNG IN DAS THEMA	<ul style="list-style-type: none"> Der/Die Studierende steigt sofort in das Thema ein, ohne den Inhalt, den Ablauf und/oder das Ziel der Präsentation zu verdeutlichen. 				<ul style="list-style-type: none"> Der/Die Studierende präsentiert prägnant den Inhalt, den Ablauf sowie das Ziel der Präsentation.

Solltest du Fragen haben, melde dich gerne jederzeit bei mir!!

Viele liebe Grüsse und ganz herzlichen Dank für deine Unterstützung,

Regina

P.S. Wenn du daran Interesse hast, stelle ich dir das finale Bewertungsschema selbstverständlich gerne zur Verfügung!

III.2: Erhebungsbogen für den Bereich «Inhalt»

Stelle dir zu jedem dieser Bewertungskriterien einen Studenten/eine Studentin vor, die während einer Präsentation dieses Kriterium besonders schlecht oder aussergewöhnlich gut erfüllt haben. Welches konkrete Verhalten haben diese Studierenden gezeigt, damit ihr Verhalten besonders ineffektiv oder effektiv für die Erfüllung des Kriteriums war? Beschreibe dieses Verhalten bitte so genau wie möglich und gib auch an, warum dieses Verhalten besonders ineffektiv/effektiv war und welche konkreten Konsequenzen das jeweilige Verhalten hatte.

Lernziel	Bewertungskriterien	besonders schlechte Erfüllung des Kriteriums			aussergewöhnlich gute Erfüllung des Kriteriums			ggf. mittlere Ausprägung
		gezeigtes Verhalten	warum ineffektiv?	Konsequenz des Verhaltens?	gezeigtes Verhalten	warum effektiv?	Konsequenz des Verhaltens?	
1 Wissen	Wissen über die aufbereiteten Inhalte zeigt sich durch...							
Die Studierenden sind in der Lage, vorgegebene oder selbstgewählte Inhalte so aufzubereiten, dass sie sie anschliessend kennen.	Kenntnis der für das Thema relevanten Begriffe							
	Kenntnis der fachspezifischen Methodik							
	Kenntnis der themenspezifischen Theorien und/oder Klassifikationen							
2 Verständnis	Verständnis der aufbereiteten Inhalte zeigt sich durch...							
Den Studierenden gelingt es, während der Aufbereitung der Inhalte ein vertieftes Verständnis dieser zu erlangen.	Korrekte Wiedergabe / Paraphrasierung der Inhalte in eigenen Worten							
	Korrekte Einbettung der Inhalte in die Präsentation							
	korrekte Interpretation der Inhalte							
3 Anwendung	Eine Anwendung der aufbereiteten Inhalte zeigt sich durch...							
Die Studierenden können die aufbereiteten Inhalte anwenden. Zur Anwendung der Inhalte gehört im Rahmen der Präsentation auch die zielgruppengerechte Kommunikation der Inhalte.	korrekte Verwendung der relevanten Begriffe und Theorien							
	Anpassung der aufbereiteten Inhalte auf den Kontext							
	Anpassung der aufbereiteten Inhalte auf die Zielgruppe							
4 Analyse	Die Fähigkeit zur Analyse zeigt sich durch...							
Den Studierenden gelingt es, die inhaltlichen Zusammenhänge des aufbereiteten Stoffs zu analysieren.	korrekte Analysen der Zusammenhänge innerhalb der aufbereiteten Inhalte							
	korrekte Analysen der Zusammenhänge zwischen den aufbereiteten Inhalten und den Inhalten des Seminars							
5 Synthese	Die Fähigkeit zur Synthese zeigt sich durch...							
Die Studierenden schaffen eine eigene Synthese der erarbeiteten Inhalte.	die zielgerichtete Auswahl der zentralen Elemente des aufbereiteten Inhalts							
	die zielgerichtete Zusammenfassung der aufbereiteten Inhalte die Ableitung (abstrakter) Implikationen und weiterführender Gedanken							
6 Bewertung	Die Fähigkeit zur Bewertung zeigt sich durch...							
Die Studierenden sind zu einer eigenständigen Bewertung der aufbereiteten Inhalte in der Lage.	eine kritische Reflexion der aufbereiteten Inhalte hinsichtlich der Inhalte und der darin enthaltenen Aussagen an sich							
	eine kritische Reflexion der aufbereiteten Inhalte hinsichtlich fachspezifischer Konventionen, Methodik und Theorien							

III.3: Erhebungsbogen für den Bereich «Präsentationskompetenz»

Stelle dir zu jedem dieser Bewertungskriterien einen Studenten/eine Studentin vor, die während einer Präsentation dieses Kriterium besonders schlecht oder aussergewöhnlich gut erfüllt haben. Welches konkrete Verhalten haben diese Studierenden gezeigt, damit ihr Verhalten besonders ineffektiv oder effektiv für die Erfüllung des Kriteriums war? Beschreibe dieses Verhalten bitte so genau wie möglich und gib auch an, warum dieses Verhalten besonders ineffektiv/effektiv war und welche konkreten Konsequenzen das jeweilige Verhalten hatte.

Lernziel	Bewertungskriterien	besonders schlechte Erfüllung des Kriteriums			aussergewöhnlich gute Erfüllung des Kriteriums			ggf. mittlere Ausprägung
		gezeigtes Verhalten	warum ineffektiv?	Konsequenz des Verhaltens?	gezeigtes Verhalten	warum effektiv?	Konsequenz des Verhaltens?	
7 Klare Kommunikation	Eine klare Kommunikation zeigt sich durch...							
Die Studierenden sind in der Lage, die Inhalte mit einer klaren Struktur und einer klaren Sprache zu präsentieren.	eine klare und nachvollziehbare Struktur							
	freies und gut verständliches Sprechen							
8 Zielgruppen-Spezifität	Zielgruppenspezifität der Präsentation zeigt sich durch...							
Die Studierenden sind in der Lage, die Präsentation auf die jeweilige Zielgruppe und deren Kenntnisstand anzupassen.	Berücksichtigung des Vorwissens der Zielgruppe							
	Anpassung der Geschwindigkeit der Präsentation an die Zielgruppe							
	Anpassung der Wortwahl an die Zielgruppe							
9 Gestaltung visueller Hilfsmittel	Eine zielführende Gestaltung der Hilfsmittel zeigt sich durch...							
Die Studierenden sind in der Lage, die visuellen Hilfsmittel zielgerichtet und angemessen einzusetzen.	ein angemessenes Verhältnis von visualisierten Inhalten und gesprochenem Wort							
	eine für den Kontext angemessene Gestaltung der Hilfsmittel							
10 Interaktion mit den Zuhörern	Eine souveräne Interaktion mit den Zuhörern zeigt sich durch...							
Die Studierenden sind in der Lage, mit den Zuhörern souverän zu interagieren.	Blickkontakt mit allen Zuhörern							
	souveräne Reaktion auf Fragen von den Zuhörern							
	Einbindung der Zuhörer							
11 Zeitlicher Rahmen	Die Einhaltung des zeitlichen Rahmens zeigt sich durch...							
Die Studierenden sind in der Lage, den vorgegebenen zeitlichen Rahmen einzuhalten.	keine wesentliche Über- oder Unterschreitung der vorgegebenen Zeit							

ANHANG IV: BEWERTUNGSSCHEMA FÜR PRÄSENTATIONEN VERSION 3**IV.1: Bewertungsbogen**

Referenten: _____ **Datum:** _____
Thema: _____ **Gesamtpunktzahl:** _____

INHALT

1 Wissen							
Die Studierenden sind in der Lage, vorgegebene oder selbstgewählte Inhalte so aufzubereiten, dass sie sie anschließend kennen.							
Kenntnis der für das Thema relevanten Begriffe			0	1	2	3	4
Kenntnis der fachspezifischen Methodik			0	1	2	3	4
Kenntnis der themenspezifischen Theorien und/oder Klassifikationen			0	1	2	3	4
2 Verständnis							
Den Studierenden gelingt es, während der Aufbereitung der Inhalte ein vertieftes Verständnis dieser zu erlangen.							
korrekte Wiedergabe/Paraphrasierung der Inhalte in eigenen Worten			0	1	2	3	4
korrekte Einbettung der Inhalte in die Präsentation			0	1	2	3	4
korrekte Interpretation der Inhalte			0	1	2	3	4
3 Anwendung							
Die Studierenden können die aufbereiteten Inhalte anwenden. Zur Anwendung der Inhalte gehört im Rahmen der Präsentation auch die zielgruppengerechte Kommunikation der Inhalte.							
korrekte Verwendung der relevanten Begriffe und Theorien			0	1	2	3	4
Anpassung der aufbereiteten Inhalte auf den Kontext			0	1	2	3	4
Anpassung der aufbereiteten Inhalte auf die Zielgruppe			0	1	2	3	4

4 Analyse					
Den Studierenden gelingt es, die inhaltlichen Zusammenhänge des aufbereiteten Stoffs zu analysieren.					
korrekte Analysen der Zusammenhänge innerhalb der aufbereiteten Inhalte					
	0	1	2	3	4
korrekte Analysen der Zusammenhänge zwischen den aufbereiteten Inhalten und denen des Seminars					
	0	1	2	3	4
5 Synthese					
Die Studierenden schaffen eine eigene Synthese der erarbeiteten Inhalte.					
zielgerichtete Auswahl der zentralen Elemente des aufbereiteten Inhalts					
	0	1	2	3	4
zielgerichtete Zusammenfassung der aufbereiteten Inhalte					
	0	1	2	3	4
Ableitung (abstrakter) Implikationen und weiterführender Gedanken					
	0	1	2	3	4
6 Bewertung					
Die Studierenden sind zu einer eigenständigen Bewertung der aufbereiteten Inhalte in der Lage.					
kritische Reflexion der aufbereiteten Inhalte hinsichtlich der Inhalte und der darin enthaltenen Aussagen an sich					
	0	1	2	3	4
kritische Reflexion der aufbereiteten Inhalte hinsichtlich fachspezifischer Konventionen, Methodik und Theorien					
	0	1	2	3	4

PRÄSENTATIONSKOMPETENZ

7 Klare Kommunikation							
Die Studierenden sind in der Lage, die Inhalte mit einer klaren Struktur und einer klaren Sprache zu präsentieren.							
eine klare und nachvollziehbare Struktur			0	1	2	3	4
freies und gut verständliches Sprechen			0	1	2	3	4
8 Gestaltung visueller Hilfsmittel							
Die Studierenden sind in der Lage, die visuellen Hilfsmittel zielgerichtet und angemessen einzusetzen.							
ein angemessenes Verhältnis von visualisierten Inhalten und gesprochenem Wort			0	1	2	3	4
eine für den Kontext angemessene Gestaltung der Hilfsmittel			0	1	2	3	4
9 Interaktion mit den Zuhörern							
Die Studierenden sind in der Lage, mit den Zuhörern souverän zu interagieren.							
Blickkontakt mit allen Zuhörern			0	1	2	3	4
souveräne Reaktion auf Fragen von den Zuhörern			0	1	2	3	4
Einbindung der Zuhörer			0	1	2	3	4
10 Zeitlicher Rahmen							
Die Studierenden sind in der Lage, den vorgegebenen zeitlichen Rahmen einzuhalten.							
keine wesentliche Über- oder Unterschreitung der vorgegebenen Zeit			0	1	2	3	4

IV.2: Verhaltensanker

Leitfrage für die Referatsausarbeitung : was sollen die Mit-Studierenden an zentralen Erkenntnissen und Ergebnissen mitnehmen?

Leitmotto für die Referatsausarbeitung: *Eigenständige* Bearbeitung des vorgegebenen Papers

INHALT

Lernziel	Bewertungskriterien	besonders schlechte Erfüllung des Kriteriums	aussergewöhnlich gute Erfüllung des Kriteriums
1 Wissen Die Studierenden sind in der Lage, vorgegebene oder selbstgewählte Inhalte so aufzubereiten, dass sie sie anschliessend kennen.	Wissen über die aufbereiteten Inhalte zeigt sich durch...		
	Kenntnis der für das Thema relevanten Begriffe	Die Studierenden verwenden zentrale Begriffe gar nicht oder falsch, vertauscht sie oder setzen sie in den falschen Kontext und führen sie nicht korrekt ein.	Die Studierenden präsentieren und verwenden die zentralen Begriffe korrekt und erklären sie präzise.
	Kenntnis der fachspezifischen Methodik	Die Studierenden beschreiben und erklären die verwendete Methodik nicht korrekt (z.B. <i>wissen die Studierenden nicht, welche Methode angewendet wurde</i>) und verwechseln methodische Termini (z.B. <i>Experiment vs Feldforschung</i>).	Die Studierenden beschreiben die verwendete Methodik einfach, korrekt und prägnant; Sie verknüpfen sie korrekt mit den inhaltlichen Aspekten.
2 Verständnis Den Studierenden gelingt es, während der Aufbereitung der Inhalte ein vertieftes Verständnis dieser zu erlangen.	Verständnis der aufbereiteten Inhalte zeigt sich durch...		
	korrekte Wiedergabe / Paraphrasierung der Inhalte in eigenen Worten	Die Studierenden geben die Inhalte mit Wörtern aus dem Text wieder und sind nicht in der Lage, sie in eigenen Worten zu beschreiben; Bei Wiedergabe mit eigenen Worten stellen die Studierenden die Inhalte falsch dar.	Die Studierenden geben die Inhalte in für sie und das Publikum (und den Kontext) typischer Sprache und mit eigenen Worten wieder.

	korrekte Einbettung der Inhalte in die Präsentation	Die Studierenden folgen bei der Argumentation einer falschen oder keiner Logik (z.B. <i>Elemente fallen vom Himmel</i>), führen Inhalte an unpassenden Stellen ein oder präsentieren Inhalte, die für das eigentliche Thema irrelevant sind.	Die Studierenden stellen korrekte Zusammenhänge zwischen den verschiedenen inhaltlichen Elementen (z.B. <i>zentrale Begriffe und Theorien</i>) her und bauen die inhaltlichen Elemente sinnvoll aufeinander auf.
	korrekte Interpretation der Inhalte	Die Studierenden interpretieren die dargestellten Inhalte gar nicht oder falsch oder ziehen aus den Inhalten unzulässige Schlussfolgerungen oder Verallgemeinerungen.	Die Studierenden beschreiben verschiedene Interpretationen der und Sichtweisen auf die Inhalte und bieten eine eigene Interpretation der Inhalte / eine eigene Schlussfolgerung aus den Inhalten an; ggf. werden diese zur Diskussion gestellt.
3 Anwendung Die Studierenden können die aufbereiteten Inhalte anwenden. Zur Anwendung der Inhalte gehört im Rahmen der Präsentation auch die zielgruppengerechte Kommunikation der Inhalte.	Eine Anwendung der aufbereiteten Inhalte zeigt sich durch...		
	korrekte Verwendung der relevanten Begriffe und Theorien	Die Studierenden wählen für die Erklärung der Inhalte die falschen theoretischen (Teil-) Aspekte aus, die gewählten Beispiele sind nicht korrekt und verwandte Begriffe werden nicht differenziert verwendet.	Die Studierenden wählen für die Erklärung der Inhalte die richtigen Begriffe und theoretischen Konzepte sowie theoretischen Teil-Aspekte aus; Begriffe sowie verwandte Begriffe werden korrekt, differenziert und im passenden Kontext verwendet.
	Anpassung der aufbereiteten Inhalte auf den Kontext	Die Studierenden nehmen keine Auswahl der Inhalte und Beispiele passend zur Fragestellung oder Schwerpunktsetzung vor, sondern übernehmen die Inhalte 1:1 aus der Literatur.	Die Studierenden akzentuieren die Inhalte gemäss der konkreten Fragestellung/inhaltlichen Schwerpunktsetzung. Es wird eine eigene Auswahl der Inhalte deutlich.
	Anpassung der aufbereiteten Inhalte auf die Zielgruppe	Die Studierenden erklären die Inhalte auf einem zu komplex oder zu einfach und nutzen das vorhandene Vorwissen der Zuhörer nicht.	Die Studierenden holen ihre Zuhörer ab, indem sie vorhandenes Vorwissen aktivieren (z.B. <i>Inhalte interaktiv erarbeiten</i>) und Bezug zu bekannten Inhalten herstellen und neue Inhalte auf angemessenem Niveau erklären.

Lernziel	Bewertungskriterien	besonders schlechte Erfüllung des Kriteriums	aussergewöhnlich gute Erfüllung des Kriteriums
4 Analyse Den Studierenden gelingt es, die inhaltlichen Zusammenhänge des aufbereiteten Stoffs zu analysieren.	Die Fähigkeit zur Analyse zeigt sich durch...		
	korrekte Analysen der Zusammenhänge innerhalb der aufbereiteten Inhalte	Die Studierenden präsentieren die Inhalte zum Thema als einzelne Blöcke und thematisieren die Verbindungen zwischen den einzelnen Themen nicht.	Die Studierenden wählen eine kohärente Argumentation für die Übergänge zwischen den Inhalten und zeigen Gemeinsamkeiten und Unterschiede auf.
	korrekte Analysen der Zusammenhänge zwischen den aufbereiteten Inhalten und den Inhalten des Seminars	Die Studierenden ziehen keinerlei oder falsche Querverweise zu den anderen Themen des Seminars.	Die Studierenden stellen Zusammenhänge zwischen der aufbereiteten Inhalten und den Inhalten des Seminars / dem Seminarverlauf durch einfache aber prägnante Erklärungen her (z.B. durch Verweise auf bereits diskutierte Seminarinhalte oder andere Referate) und beurteilen dabei, wie die Inhalte zueinander stehen (kohärent, inkonsistent o.ä.).
5 Synthese Die Studierenden schaffen eine eigene Synthese der erarbeiteten Inhalte.	Die Fähigkeit zur Synthese zeigt sich durch...		
	zielgerichtete Auswahl der zentralen Elemente des aufbereiteten Inhalts	Die Studierenden nehmen keine passende Gewichtung der einzelnen Inhalte vor, präsentieren unwichtige Elemente und stellen prägnante Punkte der Präsentation nicht als bedeutsam heraus.	Die Studierenden nehmen eine Gewichtung der Inhalte vor, idealerweise orientiert an formulierten Lernzielen. Dabei wird deutlich, was sie ihren Zuhörern vermitteln wollen.
	zielgerichtete Zusammenfassung der aufbereiteten Inhalte	Die Studierenden ziehen keine (Zwischen-)Fazits oder wählen nicht die wesentlichen Punkte aus, um ihre Botschaft klar zu vermitteln.	Die Studierenden stellen präzenterte oder erarbeitete (Zwischen-)Fazits in den Dienst ihrer eigentlichen Botschaft.
	Ableitung (abstrakter) Implikationen und weiterführender Gedanken	Die Studierenden beschreiben keine Implikationen, Ideen für weiterführende Arbeiten oder für den Transfer in die Praxis oder diese Implikationen "fallen vom Himmel".	Die Studierenden präsentieren Ideen für weitere Forschung oder den Transfer in die Praxis, die sich klar aus der vorangegangenen Argumentation ergeben.

6 Bewertung Die Studierenden sind zu einer eigenständigen Bewertung der aufbereiteten Inhalte in der Lage.	Die Fähigkeit zur Bewertung zeigt sich durch...		
	kritische Reflexion der aufbereiteten Inhalte hinsichtlich der Inhalte und der darin enthaltenen Aussagen an sich	Die Studierenden nehmen nicht kritisch Stellung zu den präsentierten Inhalten, sondern nehmen diese als gegeben hin; Widersprüche mit anderen Theorien/Inhalten oder ethische Probleme werden nicht angesprochen.	Die Studierenden stellen die Inhalte sowie die Belastbarkeit der Aussagen kritisch in Frage und nehmen dabei Bezug auf die Aussagen der Autoren, formulieren aber auch eigene Gedanken; bei fraglichen Interpretationen formulieren sie alternative Interpretationen.
	kritische Reflexion der aufbereiteten Inhalte hinsichtlich fachspezifischer Konventionen, Methodik und Theorien	Die Studierenden formulieren keine Kritik hinsichtlich der gewählten Methodik (z.B. <i>hinsichtlich der Passung von Methodik und Fragestellung, den Vor- und Nachteilen verschiedener Methoden</i>) oder sie präsentieren eine Liste an Stärken/Schwächen, die jedoch losgelöst von den fokussierten Inhalten sind.	Die Studierenden nehmen mit einer passenden Auswahl von Stärken und Schwächen Stellung zur gewählten Methode, formulieren Limitationen, deren Konsequenzen und entwickeln ggf. Vorschläge zu alternativen Vorgehensweisen.

PRÄSENTATIONSKOMPETENZ

Lernziel	Bewertungskriterien	besonders schlechte Erfüllung des Kriteriums	aussergewöhnlich gute Erfüllung des Kriteriums
7 Klare Kommunikation Die Studierenden sind in der Lage, die Inhalte mit einer klaren Struktur und einer klaren Sprache zu präsentieren.	Eine klare Kommunikation zeigt sich durch...		
	eine klare und nachvollziehbare Struktur	Die Studierenden folgen in der Präsentation keiner klaren Struktur (z.B. <i>springen in der Argumentation hin und her, haben keine Reihenfolge vorbereitet</i>) oder machen diese für die Zuhörer nicht nachvollziehbar (z.B. <i>keine Präsentation einer Agenda, keine Verwendung von sprachlichen Strukturierungshilfen</i>).	Die Studierenden folgen in ihrer Präsentation einem durchdachten und klaren roten Faden, der den Zuhörern an den notwendigen Stellen durch adäquate Mittel verdeutlicht wird.
	freies und gut verständliches Sprechen	Die Studierenden lesen ihre Präsentation vollständig ab und sprechen dabei nicht in angemessener Lautstärke oder deutlich genug und verwendet viele unnötige Füllwörter ("äh").	Die Studierenden präsentieren bei angemessener Modulation und Lautstärke sowie natürlicher Geschwindigkeit so, dass Blickkontakt zu den Zuhörern und Bewegung im Raum möglich sind (mit oder ohne Hilfsmittel).
8 Gestaltung visueller Hilfsmittel Die Studierenden sind in der Lage, die visuellen Hilfsmittel zielgerichtet und angemessen einzusetzen.	Eine zielführende Gestaltung der Hilfsmittel zeigt sich durch...		
	ein angemessenes Verhältnis von visualisierten Inhalten und gesprochenem Wort	Die Studierenden sagen ausschliesslich das, was auf den Folien steht oder verwenden keine Visualisierung, durch komplexe Visualisierungen führen sie nicht angemessen durch (z.B. <i>keine Fokussierungshilfe bei komplexen Tabellen</i>) und verwenden Medien, wo diese nicht sinnvoll sind (z.B. <i>Videos mit fraglicher Relevanz für Thema</i>).	Die Studierenden nutzen Visualisierungen um Verständnis zu fördern (z.B. <i>Kreation von eigenen Abbildungen, wo keine vorhanden sind</i>) und Vortrag zu beleben, allerdings in einem angemessenen Masse (z.B. <i>keine "Materialschlacht", keine Videos an Stellen, wo sie nicht zum Verständnis beitragen</i>).

	eine für den Kontext angemessene Gestaltung der Hilfsmittel	Die Hilfsmittel sind unübersichtlich (z.B. <i>komplexe Abbildungen vollständig aus Publikation kopiert und zu komplex</i>) oder nicht leserlich gestaltet (z.B. <i>zu geringe Kontraste, zu kleine Schrift</i>).	Die Hilfsmittel sind übersichtlich und leserlich gestaltet (z.B. <i>Schriftgrösse und Kontraste</i>), Informationen, die länger zugänglich sein sollen, werden visuell zur Verfügung gestellt (z.B. <i>auf Whiteboard</i>).
9 Interaktion mit den Zuhörern Die Studierenden sind in der Lage, mit den Zuhörern souverän zu interagieren.	Eine souveräne Interaktion mit den Zuhörern zeigt sich durch...		
	Blickkontakt mit allen Zuhörern	Die Studierenden suchen keinen Blickkontakt zu den Zuhörern oder nur zu einzelnen wenigen, drehen sich mit dem Rücken zum Publikum oder etablieren starren Blickkontakt, der für das Gegenüber unangenehm wird.	Die Studierenden schauen oft ins Publikum zu allen anwesenden Personen (auch nach rechts und links, hinten etc) mit angemessener Länge des Blickkontakts zu einzelnen Personen
	souveräne Reaktion auf Fragen von den Zuhörern	Die Studierenden verlieren nach einer Frage den roten Faden und reagieren persönlich angegriffen (z.B. <i>patzige Antwort auf wahrgenommene Kritik</i>) oder zu defensiv oder Fragen werden ignoriert, nicht aufgegriffen.	Die Studierenden werden durch Fragen nicht aus dem Konzept gebracht, klären diese sachlich und versichern sich beim Fragesteller, ob die Frage geklärt sei; Wenn sie an anderer Stelle angemessener zu beantworten sind, stellen die Studierenden die Frage mit einem entsprechenden Verweis zurück.
	Einbindung der Zuhörer	Die Studierenden entwickeln keinen Kontakt mit den Zuhörern durch (Rück-)Fragen oder persönliche Ansprachen.	Die Studierenden stellen durch (Rück-)Fragen (ggf. auch rhetorische Fragen), Sammeln von Meinungen, eigenen Erfahrungen o.ä. Kontakt zu den Zuhörern her.
10 Zeitlicher Rahmen Die Studierenden sind in der Lage, den vorgegebenen zeitlichen Rahmen einzuhalten.	Die Einhaltung des zeitlichen Rahmens zeigt sich durch...		
	keine wesentliche Über- oder Unterschreitung der vorgegebenen Zeit	Die Studierenden unter- oder überschreiten die vorgegebene Zeit ohne zu Zögern oder werden zum Ende der Präsentation erheblich schneller, um pünktlich fertig werden zu können.	Die Studierenden halten sich an die vorgegebene Zeit, haben zum Schluss noch Zeit, um Fragen zu klären sowie eigene Diskussionsfragen, falls die Präsentation früher als vorgesehen zu Ende ist.