

SZENARIEN DES BENUTZERGENERIERTEN ASSESSMENT IN OPAL UND ONYX

Wenke Müller

HS Zittau/Görlitz
wmueller@hs-zigr.de

Enrico Teich

IHI Zittau
eteich@ihi-zittau.de

Niels Seidel

IHI Zittau
nseidel@ihi-zittau.de

Zusammenfassung

Im heutigen Zeitalter von Web 2.0 sind neben neuen Kommunikationsformen auch neue Lernbedingungen entstanden, welche erweiterte Möglichkeiten des kollaborativen Lernens eröffnen, aber auch ein verändertes Rollenverständnis in Bezug auf Lernen und Lehren erfordern. Auch Assessmentprozesse im Hochschulalltag befinden sich im Wandel. Dazu gehört neben dem zunehmenden Einsatz von E-Assessment auch die Einbeziehung der Studierenden in Bewertungsprozesse (Peer Assessment). Das Projekt „BegA – Benutzergeneriertes Assessment in OPAL“, basierend auf der Kooperation zwischen der Hochschule Zittau/Görlitz, dem IHI Zittau sowie der BPS GmbH, greift diese Entwicklungen und Trends auf. Der Beitrag thematisiert vor diesem Hintergrund Szenarien des benutzergenerierten Assessment und geht dabei auf theoretischen und empirischen Grundlagen, Ziele sowie die technische Umsetzung dieser Szenarien mittels *Onyx* und *OPAL* ein.

1 Grundlagen und Ziele

Ausgehend von wesentlichen Begriffsdefinitionen und theoretischen Grundlagen erläutern wir in diesem Beitrag zunächst die Ziele des Projekts BegA.

Für den Begriff Assessment, weist Reinmann (2011, S. 19) hin, gibt es keine eindeutige Übersetzung aus dem Englischen, "denn dieser umfasst ein ganzes Bündel von Aktivitäten wie: Erfassen, Sammeln, Beschreiben, Aufnehmen, Einschätzen, Bewerten und Interpretieren". Bei diesen Erfassungs- und Bewertungsprozessen steht der Vergleich mit vorab formulierten Lehr- und Lernzielen im Mittelpunkt (vgl. Reinmann, 2011, S.19).

Im Kontext des E-Learning kann unter benutzergeneriertem Assessment eine Form von formativen Lernerfolgskontrollen verstanden werden, deren Überprüfungsaufgaben durch die Lernenden selbst erstellt und auf sie angewendet werden. Aufgaben erfüllen im Lernprozess zwei wichtige Kernfunktionen: Aktivierung und Assessment. Die Aktivierung kann sich dabei sowohl auf den kognitiven Bereich (z.B. Verknüpfung von Informationen) als auch auf den emotional-motivationalen Bereich sowie die „Aktivierung sozialer Interaktion“ beziehen (vgl. Reinmann, 2011, S.64). Daneben spielen Übungsaufgaben eine bedeutende Rolle bei der Selbstregulation des Lernens, indem sich der Lernende selbst kontrollieren kann, z.B. wie gut er sich bereits in den Lerngegenstand vertieft hat. Diese Form des Assessments wird auch als „Assessment for Learning“ bezeichnet (vgl. ebd., 2011, S.64). Davon kann das sogenannte „Assessment of Learning“ abgegrenzt werden, welches den Fokus auf die Beurteilung des Lernenden, in Bezug auf die Erreichung des Lernziels, legt (vgl. ebd., 2011, S.64).

In diesem Beitrag wird ein kombinierter Ansatz beschrieben, in dem Peers die erstellten Aufgaben nicht nur bearbeiten, sondern auch bewerten (Peer Assessment). Die aktive Aufgabenerstellung, -bewertung und -bearbeitung wird durch Aktivitätspunkte honoriert, während aus den qualitativen Bewertungen der Aufgaben und dem errechneten Schwierigkeitsgrad Qualitätspunkte zustande kommen. Der Grad der Aktivität und Qualität im benutzergenerierten Assessment resultiert in einer individuellen und sozialen Bezugsnorm für eine Rangreihe (Schott & Ghanbari 2008) in Form einer Highscoreliste.

Ausgangspunkt bzw. theoretische Grundlagen für die Idee der Veranschaulichung der Teilnehmerergebnisse durch eine Highscoreliste bilden u.a. psychologische Theorien des sozialen Vergleichs sowie das erweiterte Erwartungs-Wert-Modell der Leistungsmotivation nach Eccles (1994). Die Hauptmotive für den sozialen Vergleich bestehen in der Selbsterkenntnis, Selbsterhöhung und Selbstverbesserung. Mussweiler (2006) führt an, dass für die verschiedenen Vergleichsmotive, auch unterschiedliche Standards selektiert werden. Zur Aufrechterhaltung eines stabilen Selbstbildes werden in der Regel laterale Vergleiche, d.h. Vergleiche mit ähnlichen Personen durchgeführt. Um das Selbstwertgefühl zu erhöhen, werden bevorzugt abwärts gerichtete Vergleiche vorgenommen. Aus Vergleichen mit „überlegenen Standards“ können dagegen wichtige Informationen zur Verbesserung der eigenen Fähigkeiten entnommen werden (vgl. Mussweiler, 2006, S.104ff). In Bezug auf das benutzergenerierte Assessment ist besonders das letztgenannte Motiv zur Selbstverbesserung in der Lern- und Leistungssituation interessant. Daraus kann beispielsweise geschlossen werden, dass qualitativ gute Aufgaben und Lösungen die anderen Studierenden dazu anregen können ihre Fähigkeiten diesbezüglich zu verbessern.

Nachfolgend soll in diesem Zusammenhang auch kurz auf das Erwartungs-Wert-Modell der Leistungsmotivation nach Eccles (1994) sowie die Bedeutung der sozialen Bezugsnorm im Lernprozess eingegangen werden. Die Lern- und Leistungsmotivation wird nach diesem Modell situationsspezifisch von der Erwartung („Kann ich es lernen?“) und dem Wert („Will ich es lernen?“), der Aufgabenbearbeitung bestimmt (vgl. Möller, 2008, S.267ff). Dabei beeinflussen die „individuellen Interessen und Zielorientierungen“ des Lernenden, welchen Wert beispielsweise der Lösung einer Aufgabenstellung beigemessen wird. Die Zielorientierung kann auf die Tätigkeit selbst oder an deren Folgen, wie im Sinne des Erhaltens einer guten Note oder Anerkennung, ausgerichtet sein (vgl. Möller, 2008, S.267ff).

Die Erwartungskomponente wird durch das Selbstkonzept, die Selbstwirksamkeit sowie die Schwierigkeit der Aufgabe beeinflusst. Das Selbstkonzept basiert auf „konkreten Leistungsrückmeldungen, anschließenden sozialen Vergleichen und Kausalattributionen“ (vgl. Möller, 2008, S.274). Zur Einschätzung der eigenen Lernleistung stellen „Peers“ daher eine wichtige „Vergleichs- oder Bezugsgruppe“ dar (vgl. Möller, 2008, S.286). Insbesondere leistungsstarke Lernende profitieren von der Orientierung an der sozialen Bezugsnorm (vgl. ebd, 2008, S.290).

Hier gilt es bei der Umsetzung von Peer – und benutzergeneriertem Assessment zu beachten, dass eine Demotivierung eher leistungsschwacher Teilnehmer vermieden wird. In einigen Untersuchungen zum sogenannten „Big-Fish-Little-Pond-Effekt“ zeigte sich, dass Schüler in sehr leistungsstarken Klassen, im Vergleich zu gleich begabten Schülern in eher leistungsschwachen Klassen, oft ein niedrigeres Selbstkonzept ausbilden (vgl. Möller, 2008, S.274). Die Studierenden sollten daher beim Peer-Assessment dazu angeregt werden, vor allem Aufgaben mit einem mittleren Schwierigkeitsgrad zu entwickeln. Einer eventuellen Demotivierung kann auch durch differenzierte Rückmeldungen zur Verbesserung der eigenen Leistung, beispielsweise anhand bestimmter Bewertungskriterien, entgegen gewirkt werden.

Eines der Hauptziele des benutzergenerierten Assessment besteht in der direkten Förderung des Lernprozesses durch eine vertiefende Beschäftigung mit dem Lerngegenstand, welche durch die selbstständige Generierung von Aufgaben und Antworten entsteht. Insbesondere die eigene Erstellung von Aufgaben erfordert von den Studierenden eine kreative Eigenleistung, welche neben dem Durchdringen des Lernstoffes auch Kompetenzen zur Formulierung und Gestaltung konsistenter Fragestellungen erfordert. Die aus der Bewertung von Aufgaben entstehenden Feedbackprozesse verlangen zudem ein gewisses Maß an Team-, Kommunikations-, Empathie- und Selbstkritikfähigkeit von den Studierenden (vgl. Topping, 1998, S.256). Ein weiterer Nutzen besteht darin, dass aus den Ergebnissen des Peer- und benutzergenerierten Assessment sowohl die Studierenden selbst als auch die

Lehrenden eine Rückmeldung über die aktuellen Fortschritte im Lernprozess erhalten, sozusagen als eine formative Art der Leistungsüberprüfung während des Semesters. Der mit der Aufgabengenerierung und -bewertung einhergehende Perspektivwechsel von Dozent und Student soll weiterführend die Akzeptanz von akademischen Bewertungsprozessen verbessern und nach Möglichkeit transparenter gestalten (vgl. Topping, 1998, S.256).

2 Umsetzung der Lehr- und Lernszenarien

Der für dieses Projekt maßgebliche Ansatz besteht in der Identifizierung von didaktischen und technischen Szenarien des benutzergenerierten Assessment. Die Grundlage dafür bildet die in der ersten Projektphase durchgeführte Bedarfsanalyse. Diese umfasst, neben einer ausführlichen Sekundäranalyse, auch eine Befragung von Lehrkräften und Studierenden der Hochschule Zittau/Görlitz. Zur Erhebung der Anforderungen, Hoffnungen und Befürchtungen bezüglich der Einführung von Peer- und benutzergeneriertem Assessment wurde mit insgesamt acht Lehrenden (fünf Frauen und drei Männern) sowie 17 Studierenden eine qualitative Befragung durchgeführt. Die Erhebungen, in Form von Leitfadeninterviews sowie Fokusgruppen, haben dabei ergeben, dass beide Zielgruppen sehr ähnliche Anforderungen an entsprechende Einsatzszenarien stellen. Gemeinsame Hoffnungen bestehen etwa hinsichtlich potenzieller Lernerfolge aufgrund der vertieften Auseinandersetzung mit den Vorlesungsinhalten sowie den kontinuierlichen Rückmeldungen zum eigenen Lernfortschritt. Einen weiteren Vorteil sehen Studierende und Lehrende in der Zeit- und Ortsunabhängigkeit der Nutzung, wenn diese über die Lernplattform OPAL ermöglicht werden kann. Befürchtungen, die dagegen in den Interviews sowie Fokusgruppen geäußert wurden, beziehen sich insbesondere auf einen eventuellen Mehraufwand, unfaire Bewertungen sowie technische Schwierigkeiten. Um diese Hoffnungen und Befürchtungen bei der Umsetzung zu beachten, sollte eine entsprechende Vorbereitung, Schulung und Unterstützung der Mitglieder der Hochschule sowie eine benutzerfreundliche technische Realisierung gewährleistet sein.

Didaktische Szenarien müssen zusammengefasst, dem übergeordneten Ziel des effizienteren Lernens gerecht werden. Der hier verfolgte Ansatz gründet auf der Annahme, eine intensivere Auseinandersetzung mit den Lerninhalten führe dazu. Möchte man dies mit Mitteln des E-Learning erreichen, müssen Benutzerdialoge in einem hohen Grad bedienbar sein (*Usability*) und nach Möglichkeit eine gewisse *User Experience* vermitteln. Im Rahmen des Projektes BegA wurden verschiedene Techniken des *Cognitive Walkthrough* genutzt, um vorhandene Systeme insbesondere für unsere Anforderung zu optimieren und die Anzahl der notwendigen Arbeitsschritte zu minimieren. Dies gilt auch für die Umsetzung der Szenarien durch die Lehrenden. Hinsichtlich der *User Experience* wird hier, soweit dies innerhalb des

existierenden Lernmanagementsystems möglich ist, auf persuasive Technologien (Fogg 2003) gesetzt, welche hier in Form einer speziellen Highscoreliste (Abb. 1) motivierend auf die Lernenden wirken sollen. Einerseits bildet die Highscoreliste eine soziale Rangreihe innerhalb der Peergruppe ab, andererseits spiegelt sie auch der individuelle Bezugsnorm, d.h. die Verbesserungen bzw. Verschlechterungen bezüglich Engagements und Güte der erstellten, bearbeiteten und bewerteten Assessment-Aufgaben wider.

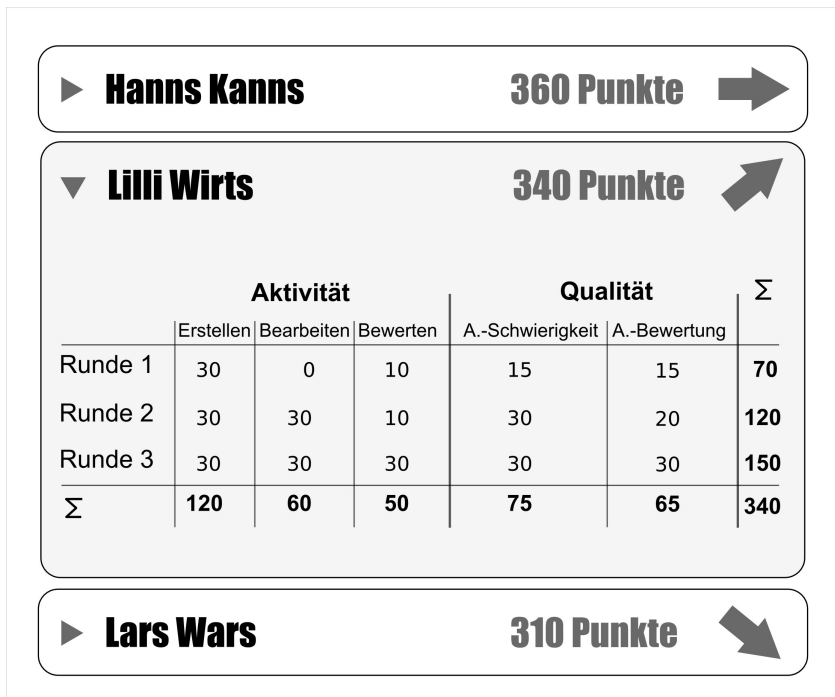


Abb. 1: Highscoreliste mit individueller Auswertung

Aus technischer Perspektive zielt das Projekt auf die Integration von Peer-Assessment-Funktionen in die Lernplattform OPAL ab. Dazu bedarf es der Heranziehung eines weiteren Werkzeugs zur Erstellung von Assessmentaufgaben (*Onyx Testsuite*) sowie eines Repositoriums (*edu-sharing*) zur Speicherung der Aufgaben. *Onyx* soll auch hier, wie bereits in vorangegangenen Projekt (Seidel 2010), hinsichtlich der Bedienbarkeit verbessert werden.

Das für die Testphase im Wintersemester 2011/2012 an der Hochschule Zittau/Görlitz geplante Einsatzszenario besteht aus den vier Phasen Information, Schulung, Workflow sowie Evaluation. Um die intrinsische Motivation der Teilnehmer zu erhöhen, ist es wichtig, dass sowohl Studierende als auch Lehrende zunächst über den Nutzen sowie die Ziele einer Einführung von

Peer- und benutzergeneriertem Assessment aufgeklärt werden. Desweiteren sind zudem organisatorische Aspekte, wie beispielsweise der zeitliche Rahmen, zu klären. Sowohl Studierende als auch Lehrende müssen vor dem Beginn der Testphase eine Schulung hinsichtlich der technischen Umsetzung, insbesondere der Arbeit mit dem *Onyx-Editor*, erhalten. Die Studierenden müssen zudem hinsichtlich der Auswahl eines geeigneten Aufgabenformates sowie der Gestaltung von Test- und Übungsaufgaben geschult werden. Außerdem ist eine kurze Einführung zum Thema Feedback- und Bewertungsprozesse vorgesehen.

Nach der Einführungs- und Schulungsveranstaltung beginnt das eigentliche benutzergenerierte Assessment. Nach Abschluss einer ersten Vorlesung erhalten die Studierenden die Möglichkeit zu dem Thema dieser Veranstaltung innerhalb einer Woche jeweils eine Aufgabe mit dem *Onyx-Editor* zu erstellen. Die Auswahl eines Aufgabenformats ist hier durch die technischen Bedingungen und Auswertungsprozesse eingeschränkt, so dass in der Testphase zur Vereinfachung nur Zuordnungs-, Auswahl- und Reihenfolgeaufgaben erstellt werden sollen. Der lizenzpflichtige *Onyx-Editor* ist, ebenfalls aufgrund einer technischen Restriktion dieses Tools, als lokale Desktopanwendung auf den Rechnern in den PC-Pools der Hochschule installiert. Die erwartete Ortsunabhängigkeit kann folglich nicht realisiert werden.

Nach der Aufgabenerstellung wird diese Ressource vom jeweiligen Studierenden in das Repositorium *edu-sharing* in einem sogenannten *shared-folder* (gemeinsamer Ort im Workspace) bis zu einem festgelegten Termin abgelegt. Der nächste Schritt besteht nun darin, dass der Lehrende bzw. eine autorisierte studentische Hilfskraft alle in *edu-sharing* abgelegten Aufgaben der Studierenden in den *Onyx-Editor* reimportiert und daraus einen gemeinsamen Test erstellt, welcher dann lokal abgespeichert wird und anschließend in einem OPAL-Kurs als neuer Testbaustein einzubinden ist. Dieser Test ist dann in OPAL für einen definierten Zeitraum (etwa 1,5 Wochen) für die Studierenden zugänglich, um die Aufgaben zu beantworten und über eine Kommentarfunktion eine qualitative Aufgabenbewertung (z.B. hinsichtlich Verständlichkeit) vorzunehmen. Die Anmerkungen zur Qualität einer Aufgabe sind insbesondere für den Ersteller der Aufgabe, aber auch für den Lehrenden interessant. Eine besonders valide und verständliche Aufgabe könnte so in ähnlicher Weise in einer Klausur Verwendung finden. Die Testergebnisse der Studierenden spiegeln sich anschließend in einer Highscoreliste wider, welche ebenfalls in dem OPAL-Kurs der jeweiligen Veranstaltung zur Verfügung steht. OPAL selbst konnte jedoch nicht dahingehend erweitert werden, die Testergebnisse zu einem solchen Ranking zusammenzuführen, weshalb wir die erweiterten Testergebnisse in Form einer XML-Datei in einem *jQuery*-Script auswerten und rendern.

Neben den bereits im Assessment-Prozess ablaufenden Maßnahmen zur Qualitätssicherung, beispielsweise bezogen auf die Verständlichkeit der Aufgaben, sollte auch am Ende des Peer- und benutzergenerierten Assessment eine Evaluation des Gesamtprozesses, bezüglich der drei Aspekte Nutzerakzeptanz, Lerneffekte sowie *Usability*, erfolgen. Zur Erhebung der Akzeptanz sollen die Teilnehmer des Assessmentprozesses, d.h. sowohl Studierende als auch Lehrende, hinsichtlich ihrer positiven und negativen Erfahrungen befragt werden. Sie werden dabei vor allem angeregt konkrete Verbesserungshinweise zu geben, die sie zu einer verstärkten Nutzung motivieren würden. Des Weiteren muss überprüft werden, inwieweit das Hauptziel des Projektes, d.h. eine Förderung des Lernprozesses durch eine vertiefte Auseinandersetzung mit dem Lerngegenstand, realisiert werden konnte. Diese Erfahrungsberichte der Nutzer sind vor allem auch für weitere zukünftige Teilnehmer interessant. Neben diesen Aspekten, sollte zudem auch die technische Umsetzung und *Usability* evaluiert werden. Dies sollte nicht nur am Ende des Assessmentprozesses, sondern auch bereits während der Umsetzung kontinuierlich erfolgen und etwaige Verbesserungsprozesse direkt initiiert werden. Insbesondere die im Projekt durchgeführten Befragungen der Zielgruppen heben die wichtige Bedeutung der Benutzerfreundlichkeit hervor. *Usability* definiert sich hier durch die Effektivität, Effizienz und Zufriedenheit der Benutzer bei der Erledigung definierter Aufgaben im *Onyx* als auch *OPAL*.

3 Ergebnisse und Ausblick

Im Rahmen dieses noch laufenden neunmonatigen Projektes (bis 12/2011) konnten bereits wichtige Ergebnisse generiert werden, aus denen sich insbesondere Empfehlungen für weitere derartig gelagerte Vorhaben ableiten lassen. Die ersten Ergebnisse belegen zunächst den Bedarf der Lehrkräfte und Studierenden der Hochschule Zittau/Görlitz an benutzergenerierten Assessment. Die aus den durchgeführten Befragungen sowie einer ausführlichen Sekundäranalyse resultierenden Anforderungen der Studierenden und Dozenten sollten weiterführend bei der Entwicklung entsprechender Einsatzszenarien Beachtung finden. Dabei war festzustellen, dass Szenarien, welche für die teilnehmende Studentenschaft einen hohen Teilnahmeanreiz als auch eine gute *Usability* bieten, aber von den involvierten Dozenten nur mit erhöhten Aufwand gesteuert und kontrolliert werden können, bei einer potenziellen Anwendung beim Lehrkollegium schnell an Akzeptanz verlieren. Andererseits schaffen einfach zu steuernde und zu kontrollierende Szenarien aufgrund ihrer oft inhärenten Trivialität für die Studierenden wenig Anreize für eine andauernde aktive Teilnahme an dem benutzergeneriertem Assessment. Hier den sprichwörtlichen goldenen Mittelweg zu finden, entwickelte sich zu einer der zentralen Herausforderungen dieses Projektes.

Des Weiteren galt es bei der Implementierung der entwickelten Lehr- und Lernszenarien, wie bereits in mehreren Vorgängerprojekten, eine ganzheitliche in OPAL integrierte Systemlösung anzustreben, da ein kombinierter Einsatz von Desktop- und Web-Anwendungen und die damit einhergehenden Medienbrüche die Komplexität der Szenarien wiederum erhöhen und somit folglich Akzeptanzprobleme provozieren. Aufgrund der bereits gesammelten software-technischen Entwicklungserfahrung bleibt diesbezüglich allerdings festzuhalten, dass eine zentralistische Entwicklerinstanz, wie sie in diesem Projekt vorliegt, für derartige innovative Vorhaben nicht unbedingt von Vorteil ist. Die finanziellen Zwänge sowie gesonderte Vermarktungsstrategien, sind nicht die besten Voraussetzungen zur Erprobung innovativer Konzepte. Begrüßenswert wären deshalb offene Programmierschnittstellen (API), um kleine Erweiterungen direkt testen zu können, ohne dass sie sich sofort auf das Gesamtsystem bzw. alle Hochschulen auswirken.

Abschließend ist zu sagen, dass benutzergeneriertes Assessment über das Potenzial verfügt sowohl für Studierende als auch Dozenten eine interessante Bereicherung des akademischen Alltags zu bewirken. Es kann einen Beitrag zur Lern- und Lehrzielerreichung leisten. Inwieweit sich diese Erwartungen erfüllen, wird die anstehende modellhafte Anwendungsphase der entwickelten Szenarien in der Lehre der Hochschule Zittau/Görlitz zeigen.

4 Literatur

- [1] *Eccles, J. S.* (1994): "Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices". *Psychology of Women Quarterly*, 18, 585-609.
- [2] *Fogg, B. J.* (2003): "Persuasive Technology - Using Computers to Change What We Think and Do". San Francisco, Morgan Kaufmann Publishers, Elsevier.
- [3] *Mussweiler, T.* (2006). "Sozialer Vergleich". In H.-W. Bierhoff & D. Frey (Eds.), *Handbuch der Sozialpsychologie und Kommunikationspsychologie* (S. 103-112). Hogrefe Verlag GmbH & Co.KG.
- [4] *Möller, J.* (2008). "Lernmotivation". In A. Renkl (Ed.), *Lehrbuch Pädagogische Psychologie* (S. 263-298). Bern: Verlag Hans Huber, Hogrefe AG.
- [5] *Reinmann, G.* (2011). "Studientext Didaktisches Design". München. Verfügbar unter: <http://lernen-unibw.de/studientexte> [2011/10/09].
- [6] *Schott, F. & Azizi Ghanbari, S.* (2008). "Kompetenzdiagnostik, Kompetenzmodelle, kompetenzorientierter Unterricht. Zur Theorie und Praxis überprüfbarer Bildungsstandards ComTrans ein theoriegeleiteter Ansatz zum Kompetenztransfer als Diskussionsvorlage". Münster: Waxmann Verlag.

- [7] *Seidel, N.* (2010). "Usability und Didaktik des E-Assessment in OPAL". WEL'10 Hochschule Zittau Görlitz.
- [8] *Topping, K.* (1998). "Peer Assessment Between Students in Colleges and Universities". *Review of Educational Research*, 68(3), 249-276.