

## ABSCHLUSSBERICHT DER AG 9: QUALITÄTSKRITERIEN FÜR UNICERT®-PRÜFUNGEN

*Fazit der Teilnehmerinnen und Teilnehmer: „Wir sind UNICert.“*

*Leitung der AG:*

Barbara Amling, Sprachenzentrum der Universität Rostock, barbara.amling@uni-rostock.de

Nicole Kruczek, Zentrale Einrichtung für Sprachen und Schlüsselqualifikationen (ZESS) der Georg-August-Universität Göttingen, nicole.kruczek@zess.uni-goettingen.de

INHALT:

1. In Vorbereitung des Workshops
2. Allgemeine Qualitätskriterien von Prüfungen
3. Qualitätskriterien für UNICert®-Prüfungen
  - 3.1. Qualitätskriterium Gerechtigkeit - Transparenz/ Objektivität/ Fairness
  - 3.2. Qualitätskriterium Hochschulrelevanz und Berufsbezogenheit
  - 3.3. Qualitätskriterium Niveaustufenadäquatheit
4. Leitfragen/ Maximen für die Erstellung von UNICert®-Prüfungen
5. Danksagung

### 1. IN VORBEREITUNG DES WORKSHOPS

*Leitfrage der AG:*

Welche Qualitätskriterien sind im Kontext einer UNICert®-Prüfung nicht nur sinnvoll, sondern im Sinne von Mindeststandards unbedingt einzuhalten?

*Grundideen der AG:*

- Erfahrungen aus den im Rahmen von Akkreditierungsverfahren begutachteten Prüfungen
- Heterogenität der Prüfungen und Ziel der Vergleichbarkeit innerhalb des UNICert®-Verbundes
- Wunsch der akkreditierten Einrichtungen nach gründlichem Feedback zu den Prüfungen
- Bedeutung der UNICert®-Zertifikate im Studienverlauf und weiteren Berufsleben

*Zielstellung/ Grenzen:*

- Gemeinsame Entwicklung von Kriterien/ Maximen für eine gute UNICert®-Prüfung
- Leitfaden für die tägliche Arbeit bei der Organisation, Erstellung und Bewertung von UNICert®-Prüfungen
- Was bedeuten Kriterien wie Gerechtigkeit/ Fairness, Hochschul- und Berufsspezifisch, Handwerk für die Planung, Erstellung, Durchführung und Bewertung von UNICert®-Prüfungen?
- Sicherung der Qualität/ von allgemeingültigen Qualitätsstandards
- Sprach-, niveaustufen- und institutionsübergreifende Diskussion anhand der vorliegenden Prüfungsaufgaben und Prüfungsarbeiten (Dank an die Einrichtungen, die Prüfungen zur Verfügung gestellt haben) und der für die einzelnen Gruppen entwickelten Leitfragen

*Spezifik des UNICert®-Konzepts*

Das UNICert®-Prüfungskonzept orientiert sich an den konkreten hochschulspezifischen Anforderungen und Bedingungen der einzelnen Einrichtungen und erfordert deshalb – im Vergleich zu standardisierten Prüfungen - spezifische Qualitätssicherungsmaßnahmen und Kriterien, die diesem Kontext gerecht werden. So sind z.B. Vorerprobung der Prüfungsaufgaben oder statistische Auswertung des Bewerberverhaltens sinnvoll, lassen sich jedoch im Alltag der meisten Sprachenzentren kaum umsetzen. Diese AG beschäftigte sich daher mit der Frage, welche Qualitätskriterien UNICert®-Prüfungen erfüllen müssen und wie diese in die alltägliche Arbeit der Erstellung und Bewertung von Prüfungsaufgaben integriert werden können.

#### *Leitfragen der Arbeit in Arbeitsgruppen*

Ausgehend von der Leitfrage der Arbeitsgruppe, welche Qualitätskriterien im Kontext einer UNICert®-Prüfung nicht nur sinnvoll, sondern im Sinne von Mindeststandards unbedingt einzuhalten sind, hatte die AG-Leitung im Vorfeld drei Bereiche (Gerechtigkeit, Niveaustufenadäquatheit sowie Hochschul- und Berufsrelevanz) als wichtige UNICert®-spezifische Qualitätskriterien ausgewählt, die unter verschiedenen Leitfragen näher beleuchtet werden sollten:

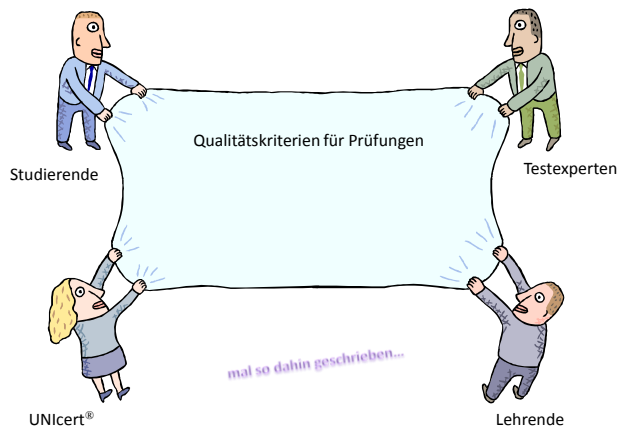
- Wie können diese Qualitätskriterien bei der Erstellung, Gestaltung und Bewertung von UNICert®-Prüfungen gesichert werden?
- Was sind typische hochschul- und berufsspezifische Inhalte und Aufgabenformate?
- Welche Maximen sollten für die Auswahl der Inhalte und Aufgabenformate (hochschul- und berufsbezogener) zugrunde gelegt werden?
- Welche Informationen müssen die Aufgabenblätter enthalten?
- Wie können Bewertungskriterien entwickelt werden? Wie können Bewertungskriterien dargestellt werden? Wann werden Bewertungskriterien bekannt geben?
- Wann ist das Qualitätskriterium Transparenz erfüllt?
- Wann ist das Kriterium der Objektivität/ Fairness (nicht) erfüllt?
- Wann ist eine Prüfungsaufgabe niveaustufenadäquat?
- Wie lassen sich insbesondere die Niveaustufen II und III hinsichtlich Inhalt und Aufgabenformaten voneinander unterscheiden?
- Wann ist das Kriterium der Hochschul- und Berufsspezifität erfüllt?
- Wie können diese Qualitätskriterien auch in die untersten Stufen eingebracht werden?

Hierzu haben eine Vielzahl der 25 Teilnehmerinnen und Teilnehmer der AG anonymisierte Prüfungsaufgaben und –arbeiten sowie Fragen, Ideen und Anregungen aus der eigenen Praxis eingebracht. Ausgehend von einem kurzen Input zu relevanten Gütekriterien und den mitgebrachten Beispielen wurden die oben aufgeführten Leitfragen diskutiert und daraus gemeinsam Maximen zur Erstellung einer guten UNICert®-Prüfung entwickelt, die als Leitfaden dienen und dazu beitragen können, die Vergleichbarkeit der UNICert®-Prüfungen zu verbessern. Die Untergruppe zum Bereich der Hochschul- und Berufsrelevanz arbeitete auf Englisch. Deshalb ist der vorliegende Abschlussbericht zum Teil zweisprachig.

## 2. ALLGEMEINE QUALITÄTSKRITERIEN VON PRÜFUNGEN

(erstellt von Nicole Kruczek)

*Input zu den Gütekriterien (als PowerPoint Präsentation)*



erstellt von N. Kruczek



### Objektivität, Reliabilität und Validität



- Voraussetzung für eine zufriedenstellende Objektivität ist die Standardisierung.
- Pilotierung, statistische Überprüfung, große Stichproben
- Je mehr Items, desto zuverlässiger.
- Geschlossene Aufgaben – schnell und objektiv auszuwerten
- Wiederholbarkeit, Messfehler vermeiden, Normierung, Datenbank
- Trennwerte
- Vermeidung von Tabuthemen, Bias und struktureller Benachteiligung
- Teststrategien
- Impact/Effekt/Auswirkung - Testethik

High, Low and Middle Stakes



#### Teil 2 – Verpflichtungen der Prüfungsabnehmer

Prüfungsabnehmer haben ein Anrecht darauf, vom Prüfungsentwickler Informationen über die Prüfungen zu bekommen. Für Prüfungsabnehmer ist es wichtig, dass die Prüfungsentwickler auch: sind sie den Kandidaten gegenüber verpflichtet, haben Standards im Hinblick auf fairen Verhalten zu setzen und einzuhalten. Diese Verpflichtungen werden im folgenden beschrieben: Auswahl geeigneter Prüfungen, Interpretation der Prüfungsergebnisse, Bewerten um Fairness.

##### Auswahl geeigneter Prüfungen

Prüfungsabnehmer sollten Prüfungen auswählen, die dem Zweck entsprechen, für den sie eingesetzt werden und die für die vorgesehene Zielgruppe der Prüfungskandidaten angemessen sind.

##### Interpretation der Prüfungsergebnisse

Prüfungsabnehmer sollten die in den Prüfungen erzielten Ergebnisse korrekt interpretieren.

##### Bewerten um Fairness

Prüfungsabnehmer sollten Prüfungen auswählen, die so entwickelt wurden, dass sie möglichst frei sind gegenüber Kandidaten mit unterschiedlichen Voraussetzungen (ethnischer oder kultureller Zugehörigkeit, Alter, Geschlecht, Behinderungen, etc.).

##### Informationen für Prüfungsteilnehmer

Prüfungsabnehmer, die in direktem Kontakt mit den Kandidaten stehen, sollten davon ausgehen, dass viele der für die Mitglieder der ALTE geltenden Verpflichtungen, die in Teil 1 unter der Überschrift „Informationen für Prüfungsabnehmer“ genannt werden, auch für sie gelten. Darübereinstimmend gilt der Code of Fair Testing Practices in Einklang mit dem von Washington U.S. Joint Committee on Testing herausgegebenen Code.



EALTA  
European Association  
for Language Testing  
and Assessment



Code of Practice...  
Richtlinien für Testautoren...  
Testspezifikationen...  
rater training...



Transparenz  
Objektivität ? Gute  
Noten!  
Nützlichkeit

„eigentlich hätte ich  
gegen alle  
Prüfungen aus  
formalen Gründen  
Einspruch erheben  
können“



Besuche den Kurs bei Dozent  
xy, da bekommst Du bessere  
Noten.

Die Note in Englisch zieht  
meinen ganzen Schnitt nach  
unten... Die Seminare sind  
doch auf Englisch, reicht das  
nicht?

Wie kann ich für die Prüfung  
lernen, ich habe noch zwei  
Tage Zeit?

Durchführbarkeit  
Möglichst wenig Aufwand  
Authentizität – Schwierigkeit  
prozessorientiert vs. Punktuell

Aber er kann doch was. Ich  
kann doch keine 5,0 geben.

Sie spricht zwar ganz gut, aber  
das ist nicht C1-Niveau.



Ich will wissen, ob die  
Studierenden  
etwas gelernt haben.

Die Studierenden kommen  
nicht mit der Zeit aus.

Offene Aufgaben – die  
Studierenden sollen den  
Transfer leisten.

Authentische Quellen – das ist  
zu schwer.



Hochschulspezifisches Zertifikat

Europaweite Qualitätsstandards und Vergleichbarkeit

Anpassung an den jeweiligen Bedarf und die Bedürfnisse

Offenheit für sinnvolle Konzepte

Testziel?  
good practice

Kommunikativ,  
Kriterienorientiert

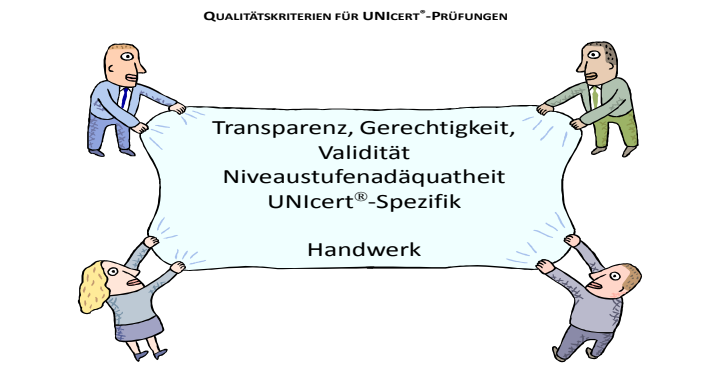


Was ist der Sinn von UNICert?

UNICert "verlangt" und beurteilt -z.B. die  
eingereichten Klausuren- und bietet  
selbst kein Muster bzw. konkreteren  
Leitfaden.

Gibt es mehrere Konzepte von UNICert?





### Aus der Rahmenordnung UNICert®

Es geht bei diesem Programm um eine hochschulspezifische und hochschuladäquate Fremdsprachenausbildung, die die Besonderheiten der Teilnehmer, der Zielsetzungen und der Arbeitsformen an Hochschulen angemessen berücksichtigt und die sich am Gemeinsamen europäischen Referenzrahmen für Sprachen (GeR) orientiert.

Ziele der dem UNICert®-System unterliegenden Fremdsprachenausbildung sind:

(a) die Befähigung zur Bewältigung hochschulbezogener allgemeiner wissenschafts-, berufs- und fachsprachlicher Situationen, wie sie im Kontext eines Studiums sowohl an einer deutschen wie auch an einer Hochschule im Land der Zielsprache erwartet werden müssen. Dazu gehört auch die Vertrautheit mit interkulturellen Problemstellungen sowie mit den kulturellen Gegebenheiten des Ziellandes [Stichwort „Mobilität“]; (b) die Vorbereitung auf die sprachlichen Anforderungen entsprechender akademischer Berufe im In- und Ausland.

Alle Teile der Prüfungen gehen gleichwertig (ohne vorherige Rundung) in die Endnote ein, die dann auf eine der in der jeweiligen Prüfungsordnung aufgeführten Noten gerundet wird.

### Mögliche Gütekriterien für UNICert®

UNICert® versteht sich als Qualitätssiegel für das Lehren und Testen / Überprüfen von Fremdsprachenkenntnissen an Hochschulen. UNICert® ermöglicht eine transparente Zertifizierung hochschulspezifischen Fremdsprachenunterrichts sowie fremdsprachlicher Leistungen. Dabei werden die Besonderheiten von Studierenden sowie Zielsetzungen und Arbeitsformen an Hochschulen besonders berücksichtigt. Hierfür müssen UNICert®-Prüfungen **qualitativen Mindeststandards** genügen. Anlehnend an Arbeiten verschiedener Autoren würden wir gern folgende Bereiche zur Diskussion stellen: Gerechtigkeit, UNICert®-Spezifik und Handwerk.

#### Gerechtigkeit

##### Fairness/ Objektivität

*Negativbeispiel: Sitzposition eines Studierenden im Abstand zur Tonquelle hat Einfluss auf seine Hörverstehensleistung → Keine systematische Benachteiligung.*

Für die Studierenden in einer UNICert®-Ausbildung sollte es egal sein, wann, wo und bei wem sie die Prüfung ablegen → Kriteriumsorientierte Bewertungen, einheitliche Regeln für die Durchführung, Auswertung und Interpretation eines Tests. Keine Normierung oder Standardisierung, aber einheitliche Festlegungen und Maßnahmen wie z.B. kollegiale Korrekturen.

### **Transparenz**

*Negativbeispiel: scheinbar willkürliche Testergebnisse aus Sicht der Studierenden hemmt evtl. weiteren Lernerfolg*

- Wahrgenommene Gerechtigkeit kann beeinflusst werden z.B. durch Vermittlung der Bewertungskriterien.
- Studierende sollen selbst Verantwortung für Ihr Lernen übernehmen. Hierzu müssen Lehrziele und am Ende nachzuweisende Kompetenzen bekannt sein.

### **Reliabilität/Zuverlässigkeit**

*Fragen:* Sind die Ergebnisse im Test zuverlässig genug, um präzise Entscheidungen zu treffen? Werden die Kompetenzen der/des Studierenden verlässlich in Hinblick auf die Niveaustufe abgebildet? → Testbeschreibung erstellen (Inhalt, Niveau, Aufgabenstellung und Aufgabentypen), Beispielaufgaben, problematische Items (z.B. raten) identifizieren, Klarheit schaffen (Ziel, Aufgabe); *problematisch:* fehlende Erprobung, Kürze des Tests...

### **UNICert®-Spezifik**

Eine UNICert®- Fremdsprachenausbildung zielt auf die Sprachverwendungssituation der Studierenden in Hochschule und späterem Berufsleben ab. Diese soll möglichst realistisch abgebildet und erfasst werden und Bezug nehmen auf die didaktischen Prinzipien des Unterrichts an Hochschulen. Sie ist kein „hartes“, sondern eher ein fließendes Kriterium, aber von zentraler Bedeutung. In der Testliteratur fallen hierunter vor allem drei Kriterien: **Authenzität und Interaktivität** (siehe Bachmann und Palmer) sowie **Aktivität** (siehe Dlaske und Krekeler).

- *Fragen:* Genuine, nicht spezifisch für den Test produzierte Aufgaben? Einbeziehung von Studierenden? Thema (und damit Aufgabe) interessant? Sprachbenutzung im Test = reale Sprachbenutzung? Test passend zum Unterricht? Bildet die Prüfung die Inhalte des Curriculums ab? Vorbereitung Test = | ≠ Vorbereitung auf reale Sprachsituation? Authentisches Sprachhandeln?

Aber auch: “Any language test is by its very nature inauthentic, abnormal language behaviour ...”

Spolsky (1985, S. 39, zitiert nach Grotjahn)

- Hochschul- und Berufsspezifisch
- Kompetenz- und Handlungsorientierung
- Rückmeldung, Übereinstimmung mit Lernzielen, Niveau der Aufgaben, Transparenz, Gerechtigkeit
- Beachtung der Auswirkung/Effekt (Impact) für Beteiligte

### **Validität**

Die Testergebnisse können auf eine bestimmte Fähigkeit interpretiert und übertragen werden; Beurteilung der Validität je nachdem, was mit dem Test gemacht wird (Messick).

- Maßnahmen zur Sicherstellung der Validität (Frage nach Ziel des Testes, Inhalt, Kriterium, Augenschein, Konstrukt, Vorhersage, Identifizierung irrelevanter Aspekte) und Stabilität (interne Konsistenz, Wiederholbarkeit, Vergleichbarkeit verschiedener Versionen)

### **Handwerk - Niveaustufenadäquatheit**

### **Angemessene formale Gestaltung**

- keine zusammenkopierten Prüfungen, Kopf und Deckblatt
- Angaben zu Kurs, Ausrichtung, Niveaustufe/UNICert®-Stufe, Hilfsmittel, zeitlichem Rahmen, erwartete Länge der Antwort, Lesbarkeit der Texte und Arbeitsanweisungen, Transparenz der Anforderungen der einzelnen Teilaufgabe, Bewertungsraster bzw. Lösungsschlüssel, Quellenangaben

### **Angemessene Aufgabenformate**

- Angemessenheit der Prüfungsaufgaben und -texte für die jeweilige Kursausrichtung, die Niveaustufe und die universitären Rahmenbedingungen, keine Betroffenheit oder Bevorzugung, Passung
- Unterschiedliche Aufgaben fordern die Aktivierung unterschiedlicher Kompetenzen und unterschiedliche kognitive Arbeitsschritte
- Unterschiede in der Vertrautheit mit den verwendeten Formaten können Einfluss auf die Aufgabenschwierigkeit haben
- Aufgaben und Texte müssen geeignet sein, um die Kompetenzen der entsprechenden Niveaustufe und des Kontextes abzubilden

### **Bewertung**

- Klare Bewertungskriterien: Lehrende müssen die Prinzipien, die der Bewertungsskala zugrunde liegen, gründlich verstehen, da sie mit dieser Bewertungsskala arbeiten müssen.
  - Einstellung und Training von Prüfern und Beurteilern, Verfahren um eine einheitliche Bewertung/Interpretation der Kriterien zu gewährleisten.
- Kriteriumsorientiert, Gewichtung von Aufgaben, Passung zur Niveaustufe etc. jeweils in Abhängigkeit von der spezifischen Testsituation
- Zusammenspiel verschiedener Faktoren im Bezug auf die Schwierigkeit

### **Quellen:**

- Bachman, L.F., Palmer, A.S. (2010). Language Testing in Practice. Oxford: Oxford University Press.
- Dlaske, Andrea; Krekeler, Christian (2009). Sprachtests. Leistungsbeurteilungen im Fremdsprachenunterricht evaluieren und verbessern. Baltmannsweiler: Schneider Verlag Hohengehren.
- Grotjahn, Rüdiger (2003). Leistungsmessung und Leistungsbewertung. Fernstudienbrief für den Weiterbildungs-Masterstudiengang „Deutschlandstudien. Schwerpunkt: Deutsche Sprache und ihre Vermittlung“. FernUniversität – Gesamthochschule in Hagen.
- Handreichungen für Testautoren der Association of Language Testers in Europe (ALTE), <http://www.alte.org/resources>
- Rahmenordnung UNICert® aktualisierte Fassung 23.02.2012, [http://www.unicert-online.org/sites/unicert-online.org/files/ro\\_unicert\\_2012.pdf](http://www.unicert-online.org/sites/unicert-online.org/files/ro_unicert_2012.pdf)

### 3. QUALITÄTSKRITERIEN FÜR UNICERT®-PRÜFUNGEN

Im Folgenden werden die Ergebnisse der Arbeit in den Kleingruppen sowie der anschließenden Diskussion in der AG zusammenfassend dargestellt.

#### 3.1. QUALITÄTSKRITERIUM GERECHTIGKEIT – TRANSPARENZ / OBJEKTIVITÄT / FAIRNESS

Prüfungen erfüllen das Kriterium der Transparenz/ Objektivität/ Fairness, wenn folgende Aspekte berücksichtigt werden:

- bei der Auswahl von Prüfungsinhalten
  - Die Thematik hat einen hochschulspezifischen bzw. berufsspezifischen Bezug und ruft nie persönliche oder ethnische *Betroffenheit* hervor (keine religiösen Themen, keine nationalen Themen etc.).
  - Stil, Diktion und Register der Texte sollten mindestens der gehobenen Umgangssprache, besser der allg. Wissenschaftssprache entsprechen, in Fachsprachenkursen der Fachsprache.
  - Die Prüfungstexte und –aufgaben stehen thematisch in Beziehung zur inhaltlichen Ausrichtung des Kurses.
  - Bei der Aufgabenstellung wird die *Darbietung des Textes* berücksichtigt, d.h. für gesprochene Texte eignen sich andere Aufgaben als für Lesetexte.
  - Aufgaben sind nicht durch Weltwissen lösbar, sondern erfordern die Beschäftigung mit Hör- bzw. Lesetext.
  - Die Aufgabeformate werden bereits während der Ausbildung zur geprüften Stufe erprobt. Dies bedeutet jedoch nicht, dass es eine gezielte Vorbereitung zur Prüfung geben soll, vielmehr sollten Studierende in der Lage sein, Gelerntes zu übertragen.
- bei der formalen Gestaltung der Prüfungsunterlagen
  - Es gibt für jede Teilnehmerin und jeden Teilnehmer ein Deckblatt mit folgenden Informationen :  
Kopfzeile Universität/Einrichtung; UNICert®-Prüfung/Kursabschlussprüfung auf Niveau UNICert®...Niveaustufe (z.B. UNICert® Basis, I, II oder III), Kursbezeichnung, Semester, Dozent/in; Inhaltliche Ausrichtung der Prüfung (z.B. Allgemeinsprache oder Fachsprache Wirtschaftswissenschaften), Platz für Namen und Matrikelnummer des Prüflings, Datum der schriftlichen Prüfung, Teilfertigkeit/en und Prüfungsdauer, Platz für das Gesamtergebnis und die Note, Platz für Datum und Unterschrift von Erst- und Zweitbewerter
  - Die Aufgabenblätter enthalten folgende Informationen:  
Namen bzw. Matrikelnummer, eine durchgehende Nummerierung der Aufgaben, Prüfungsbezeichnung (Verstehendes Lesen, Verstehendes Hören, schriftlicher Ausdruck); Zeitlicher Umfang der Prüfung bzw. Teilprüfungen; Information über zugelassene Hilfsmittel; Häufigkeit der Präsentation und Länge der Audiodateien; Quellen der Hör- und Leseverstehenstexte; Angabe ob/dass/dass nicht die Fragen in chronologischer Reihenfolge sind, Bewertungskriterien bzw. Punktzahl für die einzelnen Aufgaben sowie unter der letzten Aufgabe die Gesamtpunktzahl; genügend Schreibzeilen bei offenen Aufgaben.

- Prüfungstexte und –aufgaben sind mit dem PC geschrieben, d.h., es werden z.B. keine aufgeklebten Zeitungsartikel verwendet. Nur so ist sicherzustellen, dass alle UNICert®-Prüfungen auch in digitaler Form vorliegen und sie auch bei künstlichem Licht gut lesbar sind. Alle Schriftgrößen und Typen sind so zu wählen, dass sie gut lesbar sind. *Arbeitsanweisungen* wie „Bitte beantworten Sie die folgenden Fragen zum Text“ sind graphisch anders darzustellen (z.B. kursiv) als die *Fragen* selbst.
- bei der Erstellung von Arbeitsanweisungen
  - Die Arbeitsanweisungen sind für alle leicht verständlich, eindeutig und klar formuliert.
  - Sie sind in der Zielsprache und ein Niveau tiefer als die Prüfungsaufgaben formuliert und es wird explizit, gegebenenfalls mit Beispiellösungen, erklärt, was zu tun ist/erwartet wird.

*Anmerkungen der Arbeitsgruppe:*

Aus Gründen der Fairness sind zweisprachige bzw. deutschsprachige Arbeitsanweisungen sehr bewusst zu betrachten und nur überlegt anzuwenden. In internationalen Studierendengruppen z.B. könnten Deutsche es etwas leichter haben, wenn die Anweisungen in Deutsch sind.

Bei der Auswahl und Erstellung der Aufgaben sollte man sich immer wieder fragen, ob die Aufgabe in den Diskurs passt. (Does the task fit the given discourse? Is it authentic for this particular context?)

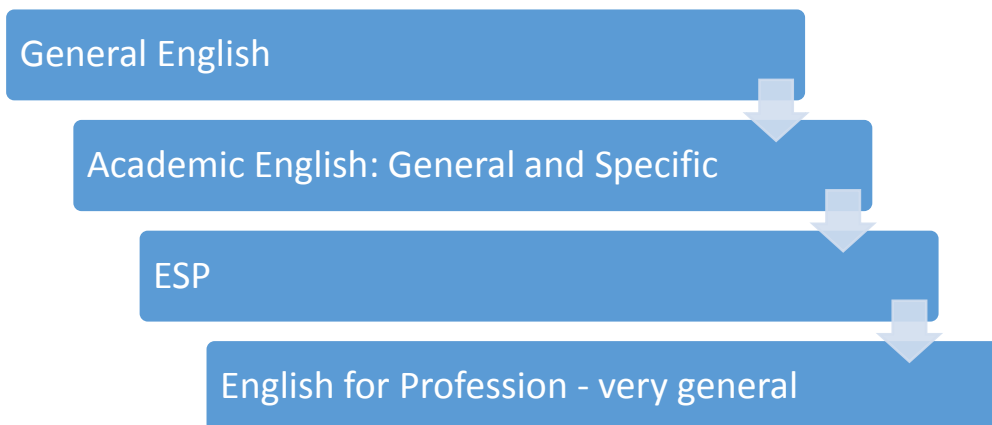
Hilfreich sind sowohl einrichtungsspezifische Durchführungsbestimmungen als auch eine Mini-Erprobung. Erste bieten Richtlinien zur Durchführung der einzelnen Prüfung, z.B. Abspielhäufigkeit der Hördateien, und dienen vor allem bei Parallelkursen oder der Arbeit mit vielen Lehrbeauftragten als Sicherung der Vergleichbarkeit. Sie können zudem den Studierenden in eventuell gekürzter Form zur Verfügung gestellt werden, um Diskussionen zu vermeiden, Transparenz zu erhöhen und Vorbereitung zu ermöglichen. Um Schwächen oder eventuelle Unklarheiten in einer Prüfung zu erkennen, ist es sinnvoll, erstellte UNICert®-Prüfungen von Kolleginnen oder Kollegen erproben zu lassen. Dies können Lehrende der eignen Einrichtung, Hilfskräfte aber auch Lehrende aus anderen Einrichtungen sein.

- bei der Bewertung von Prüfungen
  - die Bewertung ist mitgedacht, d.h. die Entscheidung für einen bestimmten Ansatz (holistisch oder analytisch), für bestimmte Instrumente (z.B. Raster, Deskriptoren, Punkteskalen) und Kriterien der Bewertung wird bewusst getroffen, die ausgewählten Ansätze und Instrumente scheinen geeignet.
  - wird ein einheitliches Punkte-Rasters (kriterienorientiert), nach den Deskriptoren des GER und von UNICert® je nach Stufe unterschiedlich tarisiert, verwendet.
  - Die Bewertungskriterien werden den Studierenden bereits während der Ausbildung bekannt gegeben, zusätzlich werden sie direkt vor/während der Prüfung bekannt gegeben.
  - Die Bewertungskriterien und der Erwartungshorizont für die einzelnen Teilprüfungen sind transparent und nachvollziehbar, ebenfalls die gesamte Bewertung der Prüfungsarbeiten.  
(Es ist erkennbar, dass die Prüfungsarbeiten von zwei Personen bewertet werden/wurden), die Punktzahl für jede Teilprüfung und die mögliche Gesamtpunktzahl sind angegeben.

- Die Gewichtung der Prüfungsteile ist erkennbar → gleiche Gewichtung der Fertigkeiten
- Bewertungskriterien wie die erwartete Wortanzahl für den schriftlichen Ausdruck und (große) Kriterien wie Inhalt, Register, Korrektheit sind angegeben.

### 3.2. Qualitätskriterium Hochschulrelevanz und Berufsbezogenheit

Diskutiert wurde insbesondere der Unterschied zwischen allgemein wissenschaftssprachlichen Kursen (English for General Academic Purposes) und fachspezifischen Sprachkursen (English for Specific Academic Purposes) und die Frage ob die Prüfung auch Kursinhalte testen kann und soll. Da UNICert® ein Ausbildungs- und Zertifizierungssystem ist, das die Kursteilnahme erfordert, um zur Prüfung zugelassen zu werden, wurde hier ein deutlicher Unterschied zu anderen Prüfungen wie z.B. TOEFL, IELTS oder Cambridge English gesehen und als besonderes Merkmal einer UNICert®-Prüfung verstanden. Anders ausgedrückt, UNICert®-Prüfungen sind in hohem Maße mit den sprachlichen und fachlichen Ausbildungsinhalten und –zielen verknüpft und erlauben/erfordern daher Formen des handlungsorientierten Prüfens, die über die reinen sprachlichen Kompetenzen hinaus gehen. (UNICert® is not content-independent but there should be more emphasis on the independent parts; „language plus exam“)



Prüfungen erfüllen das Kriterium der Hochschul- und Berufsrelevanz, wenn folgende Aspekte berücksichtigt werden:

- Die Inhalte und Formen, Aufgabenstellungen sind für den akademischen und beruflichen Kontext der Studierenden relevant; (What is academic and professional content and format? What tasks are relevant ?).
- Die Aufgabe/n passt passen in den beruflichen bzw. fachlichen Kontext.
- Die Auswahl der Inhalte und Aufgabenformate basiert auf den Ergebnissen einer Expertenbefragung im akademischen und späteren Berufsfeld, sodass davon ausgegangen werden kann, dass die Aufgaben einen tatsächlichen Bezug zum späteren Berufsleben haben. (→ conduct a **needs analysis** – do not ask the students, ask professionals)
- Die Aufgaben kommen *studienpezifischen* Handlungen nahe, simulieren diese (→ as authentic as possible, knowing that some of the tasks are only approximations).

Anders ausgedrückt: Die Aufgabenstellung(en) sollten in allen Niveaustufen handlungsorientiert/ situativ eingebettet sein und einem kommunikativen Ansatz folgen. Die Aufgaben sind in einen Kontext eingebettet, d.h., sie haben eine Verbindung zum

studentischen und akademischen Leben (UNICert®-Basis, I und II), in Fachsprachenkursen zum beruflichen und fachbezogenen Umfeld (vorrangig UNICert® III) (→ hochschulbezogen means student/university-related)

- Bei fachsprachlich ausgerichteten Prüfungen wird das Wechselspiel von Sprache und Inhalt angemessen berücksichtigt. Reine Grammatikaufgaben sollten in der Abschlussprüfung nicht eingesetzt werden, sondern vielmehr sollten die fachspezifische Lexik und Syntax etc. angemessen in die Aufgaben des schriftlichen und mündlichen Ausdrucks integriert werden.
- Hochschul- und Berufsbezogenheit bedeuten, dass die Erfüllung der Aufgabenstellung aus pragmatischer Sicht genauso stark gewichtet wird wie die rein sprachliche Erfüllung. (pragmatic fulfilment of a task is as important as the linguistic fulfilment) → relevant, meaningful testing, Adäquatheit der Sprachproduktion (what is s/he talking about, would this be useful? D.h. mehr als Sprache, auch Kontext und Inhalt (as if we were a professional in this field) – Reaktion in Wirklichkeit (Aufgabenerfüllung) → language validity and content validity must be distinguished and give us the right to impose a veto, (z.B. letter of advice – Wirkung auf Adressaten).

*Anmerkungen der Arbeitsgruppe:*

How do you know your papers meet the criteria of being academically and professionally relevant? Start your course with a needs analysis, do some research in the field, find out what's relevant for a particular profession, a person from the professional background can tell whether a task is relevant in their context we as course tutors and exam writers have to know whether a task or exam paper is relevant → talk to the professionals there is not something as a typical task, a lawyer would never describe a chart or a bar graph...

„General academic not only general“ – UNICert® unterstützt akademische Sprache, nicht Allgemeinsprache. Dies ist zwar je nach Niveau oder Ausrichtung leichter umzusetzen, sollte aber bereits bei Anfängerkursen berücksichtigt werden. Hier ist die Art und der Kontext der Aufgabenstellung relevant, z.B. Studierendenwohnheim nicht Hotel, small talk in Konferenz, etc.

Writing section: one component in which no previous content necessary – free writing essay (essay is not particularly task-oriented)

### **3.3. Qualitätskriterium Niveaustufenadäquatheit**

Die Niveaustufenadäquatheit spielt eine wichtige Rolle und spielt in alle bereits genannten Überlegungen hinein. UNICert®-Prüfungen testen auf verschiedenen Niveaus unterschiedliche Kompetenzen, die sich nicht nur aus den Niveaustufen des Gemeinsamen europäischen Referenzrahmens für Sprachen, sondern auch aus den Ausbildungsprofilen der einzelnen UNICert®-Stufen ableiten lassen.

Prüfungen erfüllen das Kriterium der Niveaustufenadäquatheit, wenn folgende Aspekte berücksichtigt werden:

- Die Prüfung und ihre Bewertungskriterien sind anhand eindeutiger Deskriptoren (nach GER und UNICert®) konzipiert.
- Es wird originales (d.h. möglichst nicht didaktisiertes) authentisches Material, das in sich schlüssig ist, mit einer dem Niveau adäquaten Komplexität (auf Satz- und Text- und Wortschatzebene), Länge und Sorte (Abstraktionsgrad) verwendet. Die „Schwierigkeit“ wird durch die Art der Aufgabenstellung reguliert. Die Texte sind nicht Lehrbüchern entnommen, d.h., es werden keine didaktisierten Texte verwendet. Leichte Vereinfachungen originalsprachlicher Texte sind möglich.
- Die Texte sind geeignet, um die Kompetenzen der entsprechenden Stufen des GER (z.B. Stufe II/ B2 - „komplexe Texte zu konkreten und abstrakten Themen“; Stufe III/ C1 – „anspruchsvolle längere Texte“) zu überprüfen bzw. nachzuweisen → bezogen auf Schwierigkeit der Texte/ Textsorte (problemorientiert vs. beschreibend), Länge und Inhalt
- Verschiedene Aufgabenformate sind miteinander verbunden (offen, geschlossen, MC, ja/nein-Fragen, richtig/falsch/warum falsch, Zuordnung, Lückentext, Titel zuschreiben oder erfinden, Reihenfolge von Paragraphen rekonstruieren, stichwortartige Wiedergabe / Zusammenfassen gehörter Texte etc.); Die *Aufgabentypen* werden mit zunehmendem Niveau der Prüfung offener. Ausschließlich MC- oder r/f-Aufgaben sind nicht adäquat. Bilder können besonders im Hörverstehen Aufgabenstellungen ergänzen oder ersetzen, um sprachliche Hürden oder das Vorwegnehmen von Sprachmaterial zu vermeiden. Bilder können auch als Sprech Anlass dienen.
- Falls möglich sollten möglichst übertragbare Aufgaben gewählt werden, z.B. Statistik beschreiben, Abstrakt erstellen, Texte in verschiedenen Sprachmodi (neutral, überzeugend, argumentierend...) → Content/task format – not specific to a particular level, more specific to a subject area.
- Bei der Auswahl von Verstehenstexten wird beachtet, dass Textsorte und –komplexität für das Niveau adäquat, die Texte inhaltlich und sprachlich auf dem richtigen Niveau sind, und einen studien- bzw. berufsspezifischen Bezug haben.
- Die Aufgaben sind auf das geprüfte Niveau kalibriert worden (Steuerung durch die Komplexität der Fragen bzw. Aufgabentypen und -stellung).
- Der Erwartungshorizont der schriftlichen Produktion ist dem Niveau quantitativ (erwartete Textlänge) und qualitativ (Textsorte, Inhalte, Komplexität der Sprache...) angepasst.
- Auch wenn nur eine Auswahl von Aufgaben gestellt werden kann, können die Studierenden zeigen, dass Sie alle Kompetenzen besitzen, die auf der jeweiligen Niveaustufe gefordert sind, z.B. Verstehen impliziter Informationen auf C1.
- Eine Erprobung und Korrekturlesen durch Kolleginnen und Kollegen zur abschließenden Sicherung der Niveauadäquatheit ist erfolgt.

### *Anmerkungen der Arbeitsgruppe*

Diskutiert wurden im Bereich *Inhalte*, ob es sinnvoll ist, mehrere kleinere (Hör-)Texte oder einen längeren Text für die Prüfung zu wählen. Beide Alternativen wurden als möglich diskutiert, die Entscheidung sollte je nach Zielsetzung der Aufgabe und der inhaltlichen Dichte der gewählten Texte angepasst werden. Mehrere kleinere Texte haben z.B. den Vorteil, dass man verschiedene Aspekte eines Themas behandeln kann, jedoch den Nachteil, dass ein erneutes Einhören in den jeweiligen Sprecher und das Thema erforderlich ist.

Bei der Frage nach der Schwierigkeit wurde die Frage besprochen, ob Regeln aufgestellt werden sollten, wie viele unbekannte Wörter ein Text auf einer Stufe enthalten darf oder soll. Konsens war, dass eine solche Festlegung keinen Sinn macht, da die Schwierigkeit eines Textes nicht allein vom Wortschatz abhängt und es gerade im hochschulischen Fremdsprachenunterricht auch auf die Simulation authentischer Situationen ankäme. So versteht auch ein muttersprachlicher Leser selten 100% der Wörter eines akademischen Textes auf Anhieb und muss vieler Lese-, Kompensations- und Texterschließungskompetenzen anwenden → Ambiguitätstoleranz, Bedeutungerschließung aus dem Kontext als wichtige akademische Kompetenzen.

Man sollte sich dessen bewusst sein, dass die Schwierigkeit einer Prüfung von folgenden Aspekten und ihrem Zusammenspiel abhängig ist:

- Text (Thematik, Komplexität der Sprache und der Textstruktur)
- Darbietung des Hörtextes (Sprecherqualität und Natürlichkeit)
- Aufgabenstellungen (eindeutig, sinnenthüllend etc.)
- Lösungsschlüssel und seine Handhabung durch jeden Korrektor (Absprache, Prüferschulung)
- Vergabe der Punktzahl im Verhältnis zur Aufgabenschwere
- Notenschlüssel (Relation von Gesamtpunktzahl und Note)

## **4. LEITFRAGEN/ MAXIMEN FÜR DIE ERSTELLUNG VON UNICERT®-PRÜFUNGEN**

Die in Abschnitt 3. beschriebenen Qualitätskriterien sollen an dieser Stelle in Maximen zur Erstellung einzelner Prüfungsteile übertragen werden.

### *4.1. Leitfragen für die Erstellung der Lese-und Hörverstehensprüfung*

- Steht der Text/ Stehen die Texte thematisch in Beziehung zur inhaltlichen Ausrichtung des Kurses bzw. zu hochschul- bzw. berufsrelevanten Themen? Ist er/ sind sie originalsprachlichen Quellen entnommen?
- Wie oft soll der Text gehört werden? Sollen die Fragen zum HV vor dem Hören durchgelesen werden?
- Nutze ich eine Audiodatei oder ein Video? (Hör-Seh-Verstehen vs. Hörverstehen) → Einfluss der visuellen Komponente und Bedeutung für Testschwierigkeit nicht wirklich geklärt
- Bezieht sich die Mehrheit der Aufgaben auf das Lese- Hörverstehen? (Im Unterschied zu Aufgabenstellungen, die eher Kenntnisse des Sprachsystems überprüfen wollen)

- Erfassen die Aufgaben zum Lese- Hörverstehen sowohl das globale als auch das Detailverstehen?
- Stehen offene und geschlossene Aufgabenstellungen in einem angemessenen Verhältnis? (Stufe III sollte in größerem Umfang als Stufe II offene Aufgabenstellungen enthalten. Auf Stufe III sollten t/f und multiple choice Aufgaben ein Maximum von 10% der Aufgaben nicht überschreiten.)
- Können die Aufgaben nur gelöst werden, wenn der Text/ die Texte gelesen/ gehört wurden?
- Harmonisieren der Schwierigkeitsgrad und die Länge der Texte mit den Aufgabenstellungen?
- Erscheinen die Aufgabenstellungen geeignet, um die Kompetenzen der entsprechenden Stufen des GER zu überprüfen bzw. nachzuweisen? (z.B. „kann ein breites Spektrum anspruchsvoller, längerer Texte verstehen und auch implizite Bedeutungen erfassen“ - Stufe III/ C1; „kann die Hauptinhalte komplexer Texte zu konkreten und abstrakten Themen verstehen – Stufe II/ B2)
- Sind die Aufgaben eindeutig und verständlich formuliert?
- Ist die Bewertung klar erkennbar?

#### *4.2. Leitfragen für die Erstellung der Klausur schriftlicher Ausdruck*

- Wurde ein kommunikativer oder handlungsorientierter Ansatz für die Aufgabenstellung gewählt?
- Hat die Aufgabenstellung inhaltlich/ thematisch einen hochschulrelevanten bzw. fachlichen Bezug?
- Hat die Aufgabenstellung einen Praxisbezug, d.h., müssen die Studierenden in der Praxis (in der studien- bzw. berufsbezogenen Kommunikation) derartige Kommunikationsaufgaben bewältigen?
- Handelt es sich bei dem zu schreibenden Text um eine Replik auf einen vorgegebenen Text (z.B. Anfrage – Antwort oder Texterstellung nach Vorgaben) oder um eine eigenständig zu entwickelnde Argumentation?
- Verlangt die Aufgabenstellung eher die Produktion eines deskriptiven (Stufe II) oder eher eines argumentativen Textes (Stufe III)?
- Erscheint die Aufgabenstellung geeignet, die Kompetenzen der entsprechenden Stufen des GER zu überprüfen bzw. nachzuweisen (z.B. Stufe II/ B2 – „kann sich klar und detailliert ausdrücken, einen Standpunkt zu aktuellen Fragen erläutern und die Vor- und Nachteile verschiedener Möglichkeiten angeben“; Stufe III/ C1 – „kann sich klar, strukturiert und ausführlich zu komplexen Sachverhalten äußern und dabei verschiedene Mittel der Textverknüpfung angemessen anwenden“)?
- Ist die Bewertung klar erkennbar?

## **5. DANKSAGUNG**

Wir danken den Kolleginnen und Kollegen der UNICert®-Kommission, namentlich vor allem Brigitte Krefting, aber besonders den Teilnehmerinnen und Teilnehmern der Arbeitsgruppe 9 des 14. UNICert®-Workshop für Ihre Anregungen und hoffen, dass die hier zusammengestellte Dokumentation neue Impulse für die weitere Arbeit bringt.

*Barbara Amling und Nicole Kruczek, Rostock und Göttingen im Dezember 2013*