

Zeitschrift für Interkulturellen Fremdsprachenunterricht

Didaktik und Methodik im Bereich Deutsch als Fremdsprache

ISSN 1205-6545 Jahrgang 14, Nummer 2 (Oktober 2009)

Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptoren-Tabellen des Gemeinsamen Europäischen Referenzrahmens

Heike Gruhn

Facultad de Lenguas
Benemérita Universidad Autónoma de Puebla, Mexiko
24 Norte 2003
Colonia Humboldt
Puebla, Pue., Mexiko
Telefon: +52 222 2295500 ext. 5820
Fax: +52 222 2979747
e-mail: heike50@hotmail.com

Abstract. Die vorliegende Arbeit wagt einen kritischen Blick auf die Prüfung Zertifikat Deutsch, wobei – vor dem Hintergrund der B1-Deskriptoren des Gemeinsamen Europäischen Referenzrahmens – einzelne Strukturaspekte und Bewertungskriterien der Prüfung herausgegriffen und analysiert werden und die Befunde in den Kontext des spezifisch mexikanischen DaF-Umfeldes gestellt werden. Leitend ist dabei die Frage nach dem Verhältnis zwischen Prüfungsergebnissen und Deskriptoren-Tabellen, was letztendlich auch mit dem Themenkomplex der Verantwortung für die Nutzung von ZD-Prüfungsergebnissen in der mexikanischen Berufswelt in Zusammenhang steht.

Abstract. In this cautiously critical assessment of the „Zertifikat Deutsch“ exam the author selects and analyzes several of its structural aspects and evaluation criteria against the background of the B1-descriptors of the Common European Framework of Reference for Languages and situates the findings in the specific context of German as a Foreign Language in Mexico. At issue is particularly the relationship between the exam results and the tables of descriptors as it concerns the responsibility for the use of ZD examination results in the Mexican business world.

Schlagwörter: Berufliches Sprachhandlungsvermögen, consequential validity, DaF in Mexiko, Deskriptoren, Gemeinsamer Europäischer Referenzrahmen, Profil-Beschreibungen, Zertifikat Deutsch

0. Einleitung

Der in der Europäischen Union zu beobachtende Trend hin zu einem einheitlichen Bildungsraum hat auch vor den fremdsprachlichen Prüfungen nicht halt gemacht. Eines der Anliegen der 2001 vom Europarat veröffentlichten Version des Gemeinsamen Europäischen Referenzrahmens (nachfolgend GeR) ist es, nationale Lehr-, Lern- und somit auch Testtraditionen im Fremdsprachenunterricht zu überdenken und vergleichbar zu machen. Dabei sieht sich der GeR nicht als normative Instanz, sondern möchte lediglich Fragen aufwerfen (vgl. Hinweise für den Benutzer). Dieses Selbstverständnis scheint von der Realität aber schon längst überholt worden zu sein. Kaum ein Lehrbuch, Kursangebot oder Prüfungsanbieter innerhalb der EU (und zunehmend auch außerhalb) kommt heute daran vorbei, seine Anlehnung an den Referenzrahmen und die dort beschriebenen Niveaustufen in großen Lettern auf die Werbefahnen zu schreiben. Prüfungen von offiziellen Prüfungsanbietern im europäischen Raum sind seit

knapp zehn Jahren an diesen Stufen ausgerichtet, wobei es für jedes Niveau spezifische Deskriptorentabellen¹ gibt. Einen normativen Effekt hat das insofern, als letztere auf einem kommunikativen bzw. kompetenzbasierten Ansatz² gründen, der rezeptiven und produktiven Umgang mit Sprache grundsätzlich als Sprachhandlung interpretiert. Prüfungsergebnisse müssen an Hand dieser Tabellen interpretierbar und in diese einordbar sein. Das führte zur Umarbeitung früherer Prüfungsdesigns, die nicht mit diesem System kompatibel waren.

Im Folgenden möchte ich im Hinblick auf die Prüfung „Zertifikat Deutsch“ (nachfolgend ZD) einige kritische Anmerkungen und Fragen aufwerfen und mich dabei auch auf die B1-Deskriptoren des GeR beziehen. Analysiert werden sollen Strukturaspekte der Prüfung unter Einbezug der Bewertungskriterien. Es handelt sich hierbei nicht um die Arbeit einer Prüfungsexpertin, sondern um einen Blick aus der Praxis, in den mitunter auch subjektive Beobachtungen einfließen können. Darüber hinaus möchte ich meine Ausführungen ausdrücklich auf den mexikanischen Kontext beziehen. Zwar ist die ZD-Prüfung weltweit gleich, es gibt jedoch spezifisch regionale Rahmenbedingungen, mit denen das Geschehen rund um eine offizielle Prüfung verzahnt ist. Prüfungen wirken sowohl auf das Unterrichtsgeschehen zurück (Mikroebene), als auch in das gesellschaftliche Umfeld hinein (Makroebene) (vgl. Schifko 2001: 827), wobei die Wirkung aber nicht nur von der Prüfung selbst abhängt, sondern auch durch Eigenheiten der jeweiligen Gesellschaft mitgestaltet wird. Einzelne Aspekte der mexikanischen Bildungslandschaft und Arbeitswelt müssen also in die Überlegungen dieser Arbeit mit einbezogen werden.

1. Das Zertifikat Deutsch als sicherer Nachweis klar definierter Kompetenzen?

1.1 Die B1-Profil-Beschreibungen der Prüfungsanbieter

In der vorliegenden Arbeit soll uns unter anderem die Frage beschäftigen, was für Behauptungen bezüglich des Absolventenprofils der Prüfung aufgestellt werden und ob diese auch eingelöst werden können. Für eine erste Annäherung an diesen Fragenkomplex ist es notwendig, einige Profil-Beschreibungen der Prüfungsanbieter näher zu durchleuchten.

Den Interessenten, die genauer wissen möchten, was von einem B1-Absolventen zu erwarten ist, wird es nicht leicht gemacht: „Mit dem Bestehen des Goethe-Zertifikats B1: Zertifikat Deutsch (ZD) weisen Sie nach, dass Sie über solide Grundkenntnisse in der deutschen Umgangssprache verfügen und dass Sie sich in allen wichtigen Alltagssituationen sprachlich zurechtfinden“, heißt es auf der Homepage des Goethe-Instituts (nachfolgend GI), wobei vor nicht allzu langer Zeit dort auch noch Folgendes stand: „Sie zeigen mit dem ZD, dass Sie die wesentlichen grammatikalischen Strukturen beherrschen [...]“.

Weiterhin wird versprochen, dass man „nach [sic] Bestehen der Prüfung [...] in Alltagssituationen wichtige Informationen aus Gesprächen, öffentlichen Durchsagen und Radiomeldungen auch im Detail verstehen, [sowie] wichtige Informationen aus Zeitungstexten, Statistiken und alltagsbezogenen Beschreibungen genau verstehen [können]“ (<http://www.goethe.de/>).

Das Österreichische Sprachdiplom (nachfolgend ÖSD) erläutert (2004: 7): „Auf dieser Stufe sollten LernerInnen fähig sein, sich in der Alltags- und Berufswelt sprachlich problemlos zurechtzufinden. Im Mittelpunkt dieser Prüfung steht die kommunikative Kompetenz, doch auch auf formale Korrektheit wird bereits geachtet“, und fügt an anderer Stelle (siehe <http://www.osd.at>, ÖSD Prüfungen - Grundlagen) dazu: „Grammatik- und Wortschatzkenntnisse sowie Aussprache und Orthographie werden als der kommunikativen Kompetenz untergeordnete Bereiche angesehen und als solche auch nur implizit mit geprüft und mit bewertet“.

Insgesamt sind die Aussagen beider Anbieter³ nicht deutlich aufeinander abgestimmt und werfen mehr Fragen auf, als sie beantworten. Dass sich Begriffe wie „Alltag“ bzw. „tägliches Leben“ auf sehr Unterschiedliches beziehen können, ist in der Fachliteratur immer wieder angemerkt worden (vgl. z.B. Claluna 2004: 25). Geht es hier um den Alltag eines Arbeiters oder um den eines Wissenschaftlers? Geht es um den Alltag des Lerner oder auch um den Alltag der Menschen, die im (für den Lerner nicht unbedingt alltäglichen) Sprachraum der Zielsprache leben?

Heike Gruhn (2009), Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptorentabellen des Gemeinsamen Europäischen Referenzrahmens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

Wie soll die widersprüchliche Einschätzung der Grammatik aufgefasst werden? Sind kommunikative Kompetenz und formale Korrektheit zwei verschiedene Stiefel (wie das ÖSD suggeriert), während das GI die „kommunikative Kompetenz“ – die ja laut ÖSD bei der Prüfung im Mittelpunkt steht – überhaupt nicht wörtlich anführt?

Um was für eine Berufswelt handelt es sich da eigentlich, die das ÖSD anspricht, das GI aber ausspart? Berufswelten können ja ebenso unterschiedlich sein wie Alltagswelten.

Trotz der nicht immer klar aufeinander abgestimmten Ausgangskonzepte haben sich die beiden Anbieter auf ein einheitliches Prüfungsdesign und einheitliche Lernziele⁴ geeinigt. Der Aufbau von ZD-GI und ZD-ÖSD ist äußerlich identisch, ebenso die jeweiligen Zeitvorgaben und Punktezahlen, auch wenn die konkreten Prüfungssätze der Anbieter nicht dieselben sind. Um von einer vollständigen Gleichwertigkeit ausgehen zu können, müsste auch die Durchführungs- und Auswertungsobjektivität untersucht werden⁵, was in diesem Rahmen allerdings nicht geleistet werden kann.

1.2 Das ZD und die Kann-Beschreibungen des Gemeinsamen Europäischen Referenzrahmens

Auf den jeweiligen Homepages bzw. in den Werbeschriften der Prüfungsanbieter sind letztendlich weder ausführliche noch präzise Auskünfte darüber zu erhalten, was von einem B1-Absolventen zu erwarten ist. Das mag damit zusammenhängen, dass die Prüfungsanbieter ausdrücklich darauf verweisen, sich auf die Niveaustufen des Referenzrahmens zu beziehen. Dieser sollte also die gesuchten Informationen liefern können.

Der Referenzrahmen selbst schreibt keinerlei Prüfungsinhalte oder Aufgabenformate vor, vielmehr lässt er Raum für unbegrenzt viele verschiedene Prüfungen. Diese sind aber alle an die dort formulierten Grundlagen gebunden, das heißt u.a., dass „Sprachhandlungen“ und „Kompetenzen“ geprüft werden sollen und Prüfungsergebnisse anhand der vorgegebenen Kann-Beschreibungen interpretierbar sein müssen. Damit ist eine gewisse Einschränkung des Spielraums einzelner Prüfungen verbunden. Wollte man die Beziehung zwischen GeR und ZD graphisch darstellen, so würde ein Pfeil vom GeR zum ZD weisen und ein weiterer von diesem wieder zurück zum GeR. Insofern ist das ZD nicht isoliert zu betrachten. Es ist grundsätzlich danach zu fragen, auf welche Weise Grundannahmen des GeR in Aufgabenformate und Prüfungsinhalte eingeflossen sind, und in welchem Umfang die B1-Deskriptorentabellen Auskunft über die Prüfungsergebnisse geben können. In dieser Arbeit können wir uns allerdings nur der letzteren der beiden Teilfragen zuwenden.

Der GeR ist in Mexiko weitgehend bekannt und wird zunehmend auch „institutionell verordnet“ (d.h. Lehrpläne und Lehrbücher sollen auf ihm beruhen, was sich wiederum auf das Unterrichtsgeschehen auswirkt). In einschlägigen Veranstaltungen wird er vorgestellt und erklärt – selten jedoch kritisch hinterfragt. Mit Grotjahn (2007: 29) können wir sagen, dass „der GeR in der Praxis zumeist als eine Art ‚Bibel‘ betrachtet [wird]“.

Für die Stufe B1 stehen in den Kapiteln 4 und 5 rund 60 Deskriptoren zur Verfügung, wobei zwischen globalen und detaillierten Kann-Beschreibungen unterschieden wird⁶. Umfassende Analysen dieser Deskriptoren sind an anderer Stelle bereits geleistet worden, weshalb hier lediglich einige Punkte kurz umrissen werden sollen.

- Kritisiert worden ist wiederholt, dass die dort aufgeschlüsselten Kompetenzen unscharf bleiben, da vielfach mit unklaren Begrifflichkeiten gearbeitet wird (vgl. z.B. Altmayer 2004: 4; Clalüna 2004: 25), sowie dass diese Unschärfen eine begründete Zuordnung von Tests und Prüfungen zu den Kompetenzniveaus des GeR teilweise nicht erlauben (vgl. Grotjahn 2007: 28). Beispiele dafür sind der oben bereits erwähnte Alltagsbegriff, aber auch Begriffe wie „vertraut“, „vorhersehbar“ oder „routinemäßig“.
- Ein weiterer, gewichtiger Kritikpunkt ist, dass die programmatischen Aussagen des GeR bezüglich der inter- bzw. plurikulturellen Kompetenzen in die Deskriptoren praktisch nicht eingeflossen sind (vgl. die Arbeiten der vorgenannten Autoren). Wie soll man sich bei der Messung und Beschreibung sprachlicher Kompetenzen und Handlungsfähigkeit denn auf „Vertrautes“ und „Routinemäßiges“ konzentrieren, wenn

Heike Gruhn (2009), Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptorentabellen des Gemeinsamen Europäischen Referenzrahmens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

Fremdsprachenlernen doch von Anfang an gerade auch die Auseinandersetzung mit dem Fremden – also Unvertrauten – bedeutet? Hierin liegt ein grundsätzliches Dilemma der Deskriptoren!

- Unterstrichen wird immer wieder, dass man dem GeR entnehmen könne, was man braucht, und dass dessen Inhalte auf spezifische Lern- oder Prüfungssituationen abstimmbare seien. Theoretisch können also die Vorgaben des GeR an regionale Bedürfnisse angepasst werden. Diese Möglichkeit wird aber dadurch ausgebremst, dass offizielle Prüfungen, die international dieselben Kompetenzstufen bescheinigen sollen, sich auf die im GeR verankerten Deskriptoren beziehen, und nicht etwa auf regional angepasste.
- Dieselbe Prüfung und dieselben Deskriptoren sollen für alle Prüflinge jeglichen Bildungsstandes und jeglichen sozio-kulturellen Erfahrungshintergrundes aussagekräftig sein. Hier wird also stillschweigend vorausgesetzt, dass alle Lerner, unabhängig von ihrem Bildungsstand, ihrer Vorerfahrung im Fremdsprachenlernen oder den Zielen, die sie mit dem Lernen der Sprache verbinden, dieselben Lern- und Kompetenzstufen durchlaufen. Doch nicht nur zwischen den Kulturen, auch unter Muttersprachlern einer selben Sprachgemeinschaft weichen die kommunikativen Fertigkeiten und die Kommunikationsbedürfnisse miteinander erheblich voneinander ab.
- Nicht zuletzt lässt sich auch das auf den mexikanischen Kontext übertragen, was Schmidt (2007: 127) in Bezug auf Japan mit „Trockenschwimmen“ bezeichnet. Wenn sprachliche Aktivitäten als „Sprachhandeln“ gedeutet werden und die Deskriptoren und sich daran orientierende Prüfungen ausschließlich dieses Handeln zu bewerten vorgeben, so stellt sich doch die Frage, wie das dort möglich ist, wo die Lerner kaum jemals auf authentische Kommunikationssituationen in der Fremdsprache treffen können, d.h. wo jedes „Sprachhandeln“ lediglich simuliert sein wird.

Festzuhalten sei an dieser Stelle, dass die Deskriptoren, die Auskunft über die Kompetenzen von Prüfungsabsolventen geben sollen, selbst in der Kritik stehen.

1.3 Zum Verhältnis von Kann-Beschreibungen und Prüfungsinhalten

Die Vielzahl an Deskriptoren zu B1 erwecken den Anschein, man hätte hier ein Instrument vorliegen, das detaillierte Auskunft über sämtliche Kompetenzen eines B1-Absolventen zu geben vermöge: „Die Summe aller Kann-Beschreibungen auf einem bestimmten Niveau kann als Beschreibung dieses Niveaus gelesen werden“, heißt es bei „Profile Deutsch“⁷. Doch neben der bereits angesprochenen inhaltlichen Unschärfe wird auch die Quantität selbst zum Problem, zumindest dann, wenn Prüfungsergebnisse damit beschrieben werden sollen. Es ist nämlich keine Prüfung denkbar, die alle in den Skalen aufgeführten Teilaspekte prüfen und deren Erfüllung verbürgen könnte.

Manche der in den B1-Skalen des GeR aufgelisteten Kompetenzen lassen sich nur ansatzweise oder so punktuell in den Prüfungsanforderungen des ZD ausmachen, dass eine teilweise oder völlige Nichterfüllung praktisch keine Auswirkungen auf das Prüfungsergebnis haben wird, und die entsprechende Kompetenz auf Grund der Prüfung nicht verbürgt werden kann.

Andere werden beim ZD überhaupt nicht geprüft, sondern bestenfalls aufgrund weiterer, geprüfter Kompetenzen als wahrscheinlich angenommen⁸. Hier kurz ein Beispiel unter vielen:

5.2.2.5. Soziolinguistische Angemessenheit B1:

[...] Ist sich der wichtigsten Unterschiede zwischen den Sitten und Gebräuchen, den Einstellungen, Werten und Überzeugungen in der betreffenden Gesellschaft und in seiner eigenen bewusst und achtet auf entsprechende Signale.

Heike Gruhn (2009), Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptoren-Tabellen des Gemeinsamen Europäischen Referenzrahmens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

Es handelt sich hier um den einzigen Deskriptoren auf B1-Niveau, der Aussagen über soziokulturelle Kompetenzen des Lerners trifft. Entworfen wird ein sehr anspruchsvolles Profil, veranschlagt er doch – sozusagen in einem Atemzug – ein klares Bewusstsein bezüglich einer Vielzahl an soziokulturellen Eigenheiten der „Zielsprachengesellschaft“ (müsste es nicht besser „Gesellschaften“ heißen?). Doch „Sitten, Gebräuche, Einstellungen, Werte und Überzeugungen“ sind ein sehr weites Feld, zumal sie sich auch innerhalb einer Sprachgemeinschaft stark unterscheiden können. Beim ZD konnte ich keine Stelle ausmachen, die eine Aussage über eine diesbezügliche Kompetenz des Prüflings erlauben würde. In der Praxis sieht es so aus, dass die Prüfung auch dann problemlos bestanden werden kann, wenn das einzige landeskundliche Programm aus den (meist wenig aufschlussreichen) Bildern und Fakten besteht, die im Lehrbuch enthalten sind.

„Él que mucho abarca poco aprieta“⁹, lautet ein mexikanisches Sprichwort. Wird wie im GeR versucht, sämtliche Elemente aufzuschlüsseln und abzustufen, die beim „Sprachhandeln“ eine Rolle spielen und auf dieses einwirken, führt das nicht unbedingt zu einem präziseren Profil von Prüfungsabsolventen. Nicht nur, weil Prüfungen ohnehin nur Ausschnitte abbilden können, sondern auch weil eine Prüfung unmöglich so viele einzelne Elemente erfassen und auf eine Weise prüfen kann, dass allgemeingültige Aussagen bezüglich des Prüflings daraus ableitbar wären.

1.4 Die verschiedenen sprachlichen Fertigkeiten: Rechenspiele

Weitere Gründe, warum die B1-Deskriptoren nicht als Katalog sämtlicher Kompetenzen des Zertifikats-Absolventen dienen können, ergeben sich aus dem Aufbau und den Bewertungsmaßstäben der Prüfung. Bestanden ist sie bereits ab 60% der erreichbaren Gesamtpunktzahl, 40% der Aufgaben darf der Prüfling also nicht oder falsch beantworten. Problematisch ist hierbei, dass die verschiedenen Teile des Zertifikats, unterschiedliche Fertigkeitsbereiche¹⁰ abdecken und Kompetenzen in all diesen Bereichen (s. GeR: Kapitel 4, dort Aktivitäten genannt) von den B1-Deskriptoren verbürgt werden.

Das ZD sieht jedoch keine Mindestpunktzahlen für einzelne Fertigkeitsbereiche vor (einmal abgesehen davon, dass sowohl der schriftliche als auch der mündliche Teil bestanden sein müssen). Fehlende Punkte in einem Prüfungsteil können beliebig durch einen anderen Prüfungsteil ausgeglichen werden. Das bedeutet, dass jemand, der zum Beispiel beim Lese- oder Hörverstehensteil nur sehr wenige oder keine Punkte erzielt, die Prüfung immer noch mit „befriedigend“ bestehen kann; wer beim einzigen produktiven Teil der schriftlichen Prüfung, dem Brief, keine Punkte erzielt, kann die Prüfung noch mit „gut“ bestehen.

Das heißt wiederum, dass es keine Garantie gibt, dass die einzelnen in den Deskriptorentabellen von Kapitel 4 des GeR beschriebenen Fertigkeitsbereiche von dem jeweiligen Absolventen auch abgedeckt werden. Im Falle des letzten Beispiels müssten sämtliche Aktivitäten des schriftlichen Ausdrucks (GeR Kapitel 4.4.1.2) in Bezug auf den entsprechenden Prüfling gestrichen werden.

Natürlich wird selten jemand einen Prüfungsteil ganz unbeantwortet lassen und in allen anderen Teilen glänzen. Wahrscheinlicher ist, dass der Prüfling auch in den Bereichen, die er kaum beherrscht, einige Punkte erzielt und somit einen größeren Fehlerspielraum für die übrigen Prüfungsteile hat.

Zu Hilfe kommen kann ihm dabei das Zufallsprinzip: Der gesamte Prüfungsteil „Hörverstehen“ (Global-, Selektiv- und Detailverstehen – ein Drittel der Gesamtpunktzahl der schriftlichen Prüfung!) besteht aus 20 richtig-falsch-Aufgaben, d.h. die Wahrscheinlichkeit, die richtige Antwort anzukreuzen, liegt durchweg bei 50%. Von diesem Aufgabentyp wird heute aufgrund von Validitätsproblemen allgemein abgeraten (vgl. Adkins Wood 1990: 40f.). Im Prinzip kann man bei einem Prüfling, der in etwa die Hälfte der Antworten richtig ankreuzt hat, nicht mit Sicherheit sagen, ob er überhaupt etwas von dem Hörtext verstanden hat. Da drängt sich doch der Eindruck auf, dass hier ein Mechanismus zur Erhöhung der Absolventenzahlen eingesetzt wird. Das Aufgabenformat erlaubt jedenfalls kaum Rückschlüsse auf die im GeR (unter 4.2.2.1) beschriebenen Fähigkeiten des Lerners, ebenso wenig kann es die vom GI proklamierte Fähigkeit des detaillierten Hörverstehens bestimmter Textsorten garantieren (vgl. Zitat Kapitel 1.1).

Heike Gruhn (2009), Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptorentabellen des Gemeinsamen Europäischen Referenzrahmens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

Nicht nur die Ratewahrscheinlichkeit spricht gegen das *multiple-choice*-Format. Manfred Schifko (2001: 833) merkt an, dass diese Aufgabenform „eine äußerst unauthentische Form der Sprachverarbeitung“ ist, „da man in realen Kommunikationssituationen selten drei oder vier Antwortalternativen präsentiert bekommt, von denen eine ausgewählt werden muss, wenn man Verständnis signalisieren möchte.“ Mit anderen Worten: Gerade bei sich als kommunikativ verstehenden Prüfungen müsste dieser Aufgabentyp eigentlich vermieden werden.

Zweifeln lässt sich auch daran, dass der Prüfungsteil „Leseverstehen Teil 2“ als Nachweis der versprochenen Kompetenz des detaillierten Leseverständnisses genügen kann. Nicht nur, dass diese Kompetenz ohne geforderte Mindestpunktzahl im Prinzip ganz ausfallen darf, es muss auch grundsätzlich gefragt werden, ob infolge der Beantwortung von insgesamt nur 5 Mehrfachwahl-Aufgaben zu ein oder zwei längeren Texten (von denen zwischen 0 und 5 richtig beantwortet sein dürfen) und einer Ratewahrscheinlichkeit von 33,3% (*multiple-choice*-Aufgaben mit drei Wahlmöglichkeiten) davon ausgegangen werden kann, dass der Prüfungstext und andere vergleichbare Texte im Detail verstanden werden.

Meine Erfahrungen bei der Vorbereitung dieses Prüfungsteils laufen darauf hinaus, dass a) die Aufgaben mit Techniken des Selektivverstehens beantwortet werden können (die jeweils vorgegebene Antwortmöglichkeit wird gesucht, nicht aber unbedingt der Text in seiner Gesamtheit gelesen und verstanden); b) dieselben Studenten bei verschiedenen Probeprüfungssätzen teilweise recht unterschiedliche Ergebnisse erzielen; aber auch: c) die Studenten oftmals die richtige Antwort angekreuzt haben, sich bei der Besprechung aber herausstellt, dass sie etwas ganz anderes verstanden hatten.

Die mündliche Prüfung enthält Routineaufgaben, auf die man sich sehr gezielt vorbereiten kann. Der erste Teil (Vorstellung) kann quasi vorher schon auswendig gelernt werden, und auch bei Teil 3 geht es immer wieder um dasselbe: Die Vorbereitung eines Treffens, bei dem verschiedene Aufgaben abgesprochen und verteilt werden müssen. Um auf vergleichbare Leistungen bei authentischen Sprechanlässen schließen zu können, müssten die Sprechaufgaben weniger vorhersehbar sein.

Um sicherzustellen, dass die Absolventen in allen Fertigkeiten die veranschlagten Kompetenzen besitzen, wäre auf jeden Fall das Ansetzen von Mindestpunktzahlen (auf der 60%-Marke) für alle Prüfungsteile sinnvoll. Und natürlich der Ersatz von Prüfungsaufgaben mit hoher Ratewahrscheinlichkeit. Denkbar wäre auch, für einzelne Fertigungsbereiche unterschiedliche Niveaustufen zu zertifizieren (z.B. Rezeption schriftlich B1, Rezeption mündlich A2), ähnlich dem Test-DaF-Modell.

Ein potentieller Arbeitgeber wird sich mit dem Zertifikatsabsolventen jedenfalls ein Überraschungsei einhandeln: Etwas ist drin, aber möglicherweise kommen gerade die Teilkompetenzen oder Fertigkeiten, auf die er besonderen Wert gelegt hätte, nicht zum Vorschein.

Prüfungen können in der Regel nur Ausschnitte eines Lernstoffes abfragen und bilden insofern auch nur Ausschnitte des veranschlagten Wissens oder Könnens ab. Das ist nicht neu und das lässt sich nicht ändern. Die Validität einer Prüfung hängt unter anderem davon ab, wie repräsentativ die gewählten Ausschnitte bezüglich der abzu prüfenden Inhalte bzw. des veranschlagten Absolventenprofils sind und wie die jeweiligen Aufgabenformate und Bewertungsparameter aussehen. Beim ZD sind diesbezüglich, wie wir gesehen haben, Mechanismen wirksam, die im Rahmen der geprüften Ausschnitte in Hinsicht auf jede einzelne der zu erbringenden Leistungen und in Hinsicht auf ganze Fertigungsbereiche größere Lücken erlauben. Sicher, Prüfungen werden allgemein mit einer Mindestpunktzahl bestanden, die unterhalb von 100% liegt. Das Problem liegt hier aber im Verhältnis zwischen Prüfungsleistungen und Aussagen bezüglich derselben. Mit anderen Worten: Es geht um die Frage, wie sich solche potentiell lückenhaften Leistungen zu der Vielzahl an Deskriptoren mit detaillierten Kann-Beschreibungen verhalten, die ja Auskunft über diese Leistungen sowie daraus abgeleitete weitere Kompetenzen geben sollen. Bei früheren Prüfungen stellte sich die Frage nicht in dieser Form, da sie von ihren Versprechungen her (bestimmter Wortschatz, bestimmte Strukturen) unverfänglicher waren. Wenn Prüfungsanbieter auf die GeR-Skalen verweisen, welche beschreiben sollen, was der Prüfling in der Fremdsprache tun kann, so impliziert das doch, dass dabei

Heike Gruhn (2009), Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptoren-Tabellen des Gemeinsamen Europäischen Referenzrahmens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

schon mitbedacht wurde, dass die Prüfung lediglich Ausschnitte erfassen kann. Darüber hinaus muss angenommen werden, dass die Aussagen der Skalen auch dann zutreffend sein sollen, wenn die Prüfung mit der Mindestpunktzahl bestanden wurde. „Die detaillierten Kann-Beschreibungen haben vor allem die Aufgabe, das typische oder wahrscheinliche Verhalten von Lernenden auf den verschiedenen Niveaus zu umschreiben“, heißt es bei Profile Deutsch (2.0, Info zu Kann-Beschreibungen [CD-ROM]). Je größer aber die Lücken sein dürfen, die die Prüfung erlaubt, und je zahlreicher die Detailaussagen für die Niveaustufe sind, desto wahrscheinlicher wird es, dass viele der letzteren auf einzelne Prüflinge gar nicht zutreffen. Mit anderen Worten: Die Deskriptoren können bezüglich der Ergebnisse des ZD nur bedingt aussagekräftig sein. In welchem Ausmaß sie im Falle welcher Prüfungsergebnisse zutreffen sollen, darüber müssten Nutzer eigentlich genauer unterrichtet werden. Denn dass diesen solche Zusammenhänge bewusst sind, kann nicht erwartet werden.

2. Das Zertifikat Deutsch und die kommunikative Kompetenz

2.1 Anmerkungen zu den Bewertungsmaßstäben beim Prüfungsteil „Schriftlicher Ausdruck“

Geht es bei der Auswertung der meisten Prüfungsteile lediglich um die Auszählung von richtig angekreuzten Lösungen, so finden sich mit dem schriftlichen und mündlichen Ausdruck zwei produktive Prüfungsteile, bei denen die Subjektivität der Prüfer in die Auswertung mit einfließt. Das lässt sich bei Prüfungen zur freien Sprachproduktion generell nicht völlig vermeiden (man denke nur an die Bewertung schulischer Aufsätze in der Muttersprache).

Nachfolgend möchte ich am Beispiel der sprachsystematischen und hier insbesondere der grammatischen Kompetenzen die Bewertungskriterien des Prüfungsteils „Schriftlicher Ausdruck“ etwas genauer betrachten. Vorange stellt sei ein einschlägiger Deskriptor des GeR, mit dem diese Kriterien abgeglichen werden sollen:

5.2.1.2 Grammatische Korrektheit B1:

Kann sich in vertrauten Situationen ausreichend korrekt verständigen; im Allgemeinen gute Beherrschung der grammatischen Strukturen trotz deutlicher Einflüsse der Muttersprache. Zwar kommen Fehler vor, aber es bleibt klar, was ausgedrückt werden soll.

Der GeR ist hier in sich schlüssig: Da er die grammatische als integralen Bestandteil der kommunikativen Kompetenz betrachtet, soll dieser Deskriptor eine gute Beherrschung der wichtigsten grammatischen Strukturen bescheinigen. Fehler dürfen vorkommen, aber eben nur dann, wenn sie nicht verständnisstörend sind, d.h. die Kommunikation nicht beeinträchtigen.

Die „formale Richtigkeit“ ist eines von drei (GI) bzw. eines von vier (ÖSD) Teilkategorien bei der Bewertung des ZD-Briefes, wobei der Prüfling in dieser Rubrik mehr als 0 Punkte erzielen muss (sonst ist der Brief nicht bestanden, die Gesamtprüfung kann trotzdem bestanden werden).

Für die Bewertung von Fehlern in Syntax, Morphologie, Orthografie und Interpunktion steht eine Skala¹¹ zur Verfügung, die Fehler folgendermaßen abstuft: „keine/nur vereinzelte, einige, an mehreren Stellen, an vielen Stellen“. Ein weiteres Kriterium ist die Verständlichkeit des Textes, es wird abgestuft nach: Fehlern, die „das Verständnis nicht stören, nur wenig beeinträchtigen, an mehreren Stellen erheblich beeinträchtigen, an vielen Stellen erheblich beeinträchtigen, Brief kaum noch verständlich“.

Die Interpretation von Kriterien wie „einige“, „mehrere“, „viele“ usw., aber auch Einschätzungen bezüglich der Verständlichkeit können je nach Prüfer erheblich variieren. Eine Textstelle, aus der ein Prüfer noch etwas Verständliches herauszulesen glaubt, mag einem anderen schon zusammenhanglos erscheinen. Was der eine schon als „viel“ empfindet, kann für den anderen noch akzeptabel sein. Die vorgegebenen Bewertungsraster des Briefes eröffnen einen Spielraum, der von der Subjektivität, d.h. den individuellen Kriterien der einzelnen Prüfer, ausgefüllt werden muss (und das nicht nur hinsichtlich der „Formalen Richtigkeit“). Das hat zur Folge, dass die Beno-

Heike Gruhn (2009), Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptoren-Tabellen des Gemeinsamen Europäischen Referenzrahmens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

tung derselben Arbeit je nach Prüfer oder Prüfungskommission unterschiedlich ausfallen kann. Klar operationalisierbare Bewertungskriterien (so z.B. Vorgaben bezüglich der Anzahl oder Schwere der Fehler) gibt es nicht.

Wiederholungsfehler sollen – so die Vorgaben – prinzipiell nur einmal angerechnet werden. Dass bedeutet aber, dass nicht unterschieden wird, ob es sich um Fehler handelt, die sich auf Einzelfälle beziehen (ein falscher Artikel, ein falsch gebildeter Plural, eine falsche Deklination), oder ob es um grundlegende Strukturen geht, die nicht beherrscht werden (z.B. die Nebensatzstellung oder die Unterscheidung von Akkusativ und Dativ). Das ist unbefriedigend. Auch wenn die Schwere von Fehlern Meinungsverschiedenheiten unter Prüfern auslösen mag, sollten Fehler nicht erst dann unterschiedlich gewichtet werden, wenn sie Verständnisprobleme verursachen.

Die auf dieser Stufe verlangten Sprachhandlungen können i.d.R. auch mit wenig gefestigten Strukturkenntnissen so bewältigt werden, dass sie noch verständlich sind. Wird das Verständnis bereits etwas beeinträchtigt, erhält der Prüfling 40% Punkteabzug in dieser Kategorie. Beeinträchtigen Fehler an mehreren Stellen das Verständnis erheblich, sind es 60%. Beeinträchtigen sie an vielen Stellen das Verständnis erheblich (wobei in der Schwebe bleibt, worin genau der Unterschied zwischen „mehrere“ und „viele“ liegt), sind es 80%. Erst wenn „der Brief kaum noch verständlich ist“, erhält der Kandidat 0 Punkte.

Die Aussagen des oben zitierten Deskriptors des GeR sind aber bereits dann nicht mehr auf einen Prüfling zutreffend, wenn diesem 40% der Punkte abgezogen werden, da hier eine teilweise Beeinträchtigung des Verständnisses in Kauf genommen wird. Mit anderen Worten: Jegliche Bewertung der „formalen Richtigkeit“, die Verständnisstörungen explizit mit einbezieht, deckt den oben zitierten Deskriptoren (5.2.1.2) für Grammatik nicht mehr ab.¹²

Festzuhalten bleibt, dass der Fehlerspielraum, der dem Prüfling beim Prüfungsteil „Schriftlicher Ausdruck“ des ZD im Bereich „Formale Richtigkeit“ zugestanden wird, insgesamt sehr groß ist.

Weitere Bewertungsskalen des Briefes wie die zur kommunikativen Gestaltung oder zum Textaufbau weisen ebenfalls erhebliche Unschärfen auf. Was genau ist der Unterschied zwischen einer „guten“, einer „angemessenen“ und einer „zum Teil noch angemessenen“ kommunikativen Gestaltung? Was ist, wenn ein Prüfling nur zwei oder drei Konnektoren beherrscht, diese aber richtig anwendet, während ein anderer mit einem größeren Repertoire mehr Fehler macht? Dazu ist den Skalen nichts zu entnehmen. Die Bewertung hängt auch hier vom Fingerspitzengefühl der Prüfer ab. Und von deren Vorbildung, d.h. in diesem Fall von ihrer Erfahrung im Umgang mit Texten. Eine solche ist unerlässlich, wenn es darum geht, Fehler im Textaufbau zu erkennen.

Die Korrekturvorgaben müssten insgesamt eindeutiger und detaillierter sein, um mehr Objektivität erreichen zu können.

2.2 Anmerkungen zu sprachsystematischen Kompetenzen

Die eben konstatierte hohe Fehlertoleranz in Hinblick auf sprachsystematische Gesichtspunkte kommt allerdings nicht von ungefähr. Sie gründet sich auf die jüngere Spracherwerbsforschung, die gezeigt hat, dass das Auftreten von Fehlern während des Fremdspracherwerbs etwas völlig Normales ist und dass bestimmte Fehler in bestimmten Lernphasen regelrecht zu erwarten sind. Man spricht heute von Lernaltersprachen (*interlanguages*), die im Laufe des Aneignungsprozesses immer wieder modifiziert werden. Apeltauer (1997: 115) drückt das folgendermaßen aus: „Die Entwicklung einer Lernaltersprache ist im Idealfall ein Prozess fortschreitender Differenzierung und Restrukturierung, in dessen Verlauf die Fähigkeiten, zielsprachige Äußerungen zu verstehen und zu produzieren, beständig verbessert werden“. Diese Einsicht ist zunächst einmal unumstößlich, kontrovers diskutiert werden kann aber dennoch die Frage, wie viel sprachsystematische Kompetenz zu welchem Zeitpunkt vom Lerner erwartet werden sollte.

Heike Gruhn (2009), Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptoren-Tabellen des Gemeinsamen Europäischen Referenzrahmens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

Die klassischen Theoretiker der kommunikativen Kompetenz (Hymes, Canale u.a.), aber auch die Autoren des GeR (Kapitel 5.2) subsummieren die grammatische und lexikalische Kompetenz unter die linguistische. Diese wiederum bildet im GeR neben der soziolinguistischen und der pragmatischen Komponente einen integralen Bestandteil der kommunikativen Kompetenz, d.h. all die genannten Komponenten müssen ineinandergreifen, um erfolgreiches „Sprachhandeln“ zu gewährleisten. Das heißt aber, dass im GeR die Sprachsystematik der kommunikativen Kompetenz nicht untergeordnet wird, sondern zu den Bausteinen zählt, die diese konstituieren. Letztendlich erfüllen nicht nur der Wortschatz, sondern auch die Grammatik sinnstiftende Funktionen. Butzkamm & Butzkamm (2004: 274) gehen sogar so weit, in der Grammatik überhaupt den Schlüssel zu menschlicher Phantasie und Freiheit zu sehen. Sie hat eine Art Scharnierfunktion, die es uns erlaubt, „dass wir die fasslich gewordenen Denkinhalte fast nach Gutdünken hin- und herbewegen können.“ „Ohne angemessene Sprachsystematik können die sprachlichen Fertigkeiten nicht entwickelt werden“, möchte ich mit Alyaz & Freibichler (2005: 16) behaupten, für welche „die entscheidende Rolle der Sprachsystematik beim [Fremd]Spracherwerb nach wie vor nicht zu leugnen ist“.

In diesem Zusammenhang ist allerdings einzuräumen, dass eine fehlerhafte Sprachsystematik nicht in allen Fällen zum Fehlschlagen der Kommunikation führen muss: Viele grammatikalisch, lexikalisch, orthographisch oder phonetisch unkorrekte Äußerungen können, wie oben schon angesprochen, verstanden werden und eine hinreichend klare kommunikative Funktion in bestimmten Situationen erfüllen.

Wovon es nun im Einzelnen abhängt, ob bzw. wann Fehler im Bereich der Sprachsystematik die Kommunikation zum Scheitern bringen, ist schwer festzulegen (zur Grammatik vgl. Völzing 1995: 509-527). Sicher wird diese umso wichtiger, je komplexer die Sachverhalte sind, mit denen der Lerner umzugehen hat (dies wurde in die Deskriptoren aufgenommen: Die Komplexität von Kommunikationssituationen wird je nach Kompetenzstufe erhöht). Sie erhält einen höheren Stellenwert beim „Heraustreten aus dem Zeigefeld“, bei Telefongesprächen zum Beispiel, beim Sprechen über Vergangenes, Zukünftiges oder räumlich Entferntes, und allgemein bei abstrakten Inhalten. Sie hängt ganz allgemein vom Präzisionsbedarf einer bestimmten Kommunikationssituation ab.

Wird die kommunikative Kompetenz im produktiven Bereich daran gemessen, ob Äußerungen produziert werden können, die verständlich oder gerade noch verständlich sind, so bedeutet das zwar nicht, dass sprachsystematisch fehlerhafte Aussagen angestrebt würden, es impliziert aber auf jeden Fall, dass solche weitgehend toleriert werden.

Ein Sprecher, der sich nur um die kommunikative Funktion kümmert, kann bei seinen Aussagen alle linguistischen Elemente aussparen, die für die Kommunikation nicht unerlässlich sind. Er kann zum Beispiel die grammatikalischen Morpheme für Geschlecht und Anzahl bei den Adjektiven weglassen, da die Substantive ja bereits Geschlecht und Anzahl angeben[.]

heißt es bei Da Silva Gomes & Signoret Dorcasberro (2005: 124, Übers. H.G.) wobei anzumerken ist, dass sich die jeweils erlässlichen Elemente von Sprache zu Sprache und von Fall zu Fall unterscheiden können. Sprache, die nur zu kommunikativen Zwecken verwendet wird, erfährt – so ist aus dem Text herauslesen – eine „funktionale Verarmung“. Die Autorinnen (vgl. ebd.: 213) zählen in diesem Zusammenhang die Sprachfunktionen nach J.H. Schumann auf, der neben der kommunikativen auch eine integrative (Zugehörigkeit zu einer sozialen Gruppe) und eine expressive Funktion (z.B. Sprachkunst) erwähnt.

Was die systematische Grammatik auf B1-Niveau betrifft, so treffen wir auf folgende Situation: Über sprachliche Mittel, die für dieses Niveau im Allgemeinen (in Curricula oder Lehrbüchern) veranschlagt werden, muss der ZD-Prüfling nicht unbedingt verfügen, um die in der Prüfung geforderten produktiven Sprachhandlungen erfolgreich ausführen zu können. So ist z.B. bei Profile Deutsch (CD-ROM Version 2.0) für das Niveau B1 unter „systematische Grammatik“ die Deklination der Adjektive im Dativ mit dem Beispielsatz „Das Haus ist in einem schlechten Zustand, [...]“ aufgeführt, hat aber ein Prüfling den Dativ grundsätzlich nicht verstanden und schreibt – „Die Haus ist in schlechte Zustand“, so ist dieser Satz vollkommen verständlich und wird die grammatikalischen Ansprüche der schriftlichen Prüfungsaufgabe erfüllen (vgl. Kapitel 2.1).¹³

Heike Gruhn (2009), Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptoren-Tabellen des Gemeinsamen Europäischen Referenzrahmens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

Bezüglich der rezeptiven Sprachsystematik liegen die Ansprüche etwas höher; sie wird in einem eigenen Prüfungsteil mit 20 Multiple-Choice-Aufgaben zu Wortschatz und Grammatik geprüft. Auch dieser kann allerdings kein Garant für gute sprachsystematische Kompetenzen sein: Eine geforderte Mindestpunktzahl gibt es nicht, und von der Gesamtpunktzahl der Prüfung entfallen lediglich 10% auf diesen Teil.

Prüft man, wie beim ZD, schwerpunktmäßig nach funktionalen Kriterien, so wird dabei der Tatsache nicht genügend Rechnung getragen, dass sprachsystematische Kompetenzen aufeinander aufbauen. Der Erwerb solcher Kompetenzen in der Grundstufe sollte nicht nur in Hinblick auf deren unmittelbaren Nutzen betrachtet werden, er dient auch als Grundlage für die langfristige Entwicklung einer differenzierteren Sprachbeherrschung.

Ein nicht verstandener Dativ zum Beispiel wirkt sich auf die Deklination von Artikeln, Pronomen oder Adjektiven aus und wird spätestens dann auch die kommunikative Kompetenz beeinträchtigen, wenn es darum geht, in komplexeren Satzgefügen und Texten Bezüge zu erkennen oder herzustellen.

Manche Fachleute, wie z.B. die Prüfungsforscherin Rea-Dickins (2001: 29), vertreten die Auffassung, dass Sprachsystematik speziell in früheren Lernerstadien geprüft werden sollte:

In certain contexts – for example in the earlier stages of language learning [...] it may be highly relevant to have explicit system-focused testing. In these situations, it may be crucial to uncover whether learners have the right packaging at their disposal in order to convey messages appropriately. [...] The more indirect approach to assessing grammatical ability, observed above, may be premised on the assumption that the learners (should) have already acquired an adequate command of the language system to package messages appropriately.

Hier könnte man nun einwenden, dass das ZD ja ein Sprachstandstest ist und dass Institutionen, die mehr Wert auf Sprachsystematik legen, diese in eigenen Lernfortschrittstests prüfen können. Das ist sicher richtig, unterschätzt wird bei diesem Argument aber der *washback*-Effekt, der von Sprachstandstests ausgehen kann. Rückwirkungen der neuen GeR-basierten Prüfungskonzepte auf Lehrmaterialien und Unterricht sind gegenwärtig überall zu beobachten. Wir leben in einer Bildungslandschaft (und das nicht nur in Mexiko), bei der statistische Werte aller Art immer mehr ins Blickfeld geraten, da Mittel und Ruf (Rankings!) einzelner Institutionen zunehmend von diesen abhängen. „Teaching to the test“ wird da für viele ein Muss, Lernen als „l’art pour l’art“ ein reiner Luxus. Im Zuge dieser Entwicklung ist Quantität (Absolventenzahlen!) oftmals mehr wert als Qualität. Für den Sprachunterricht kann das dann leicht bedeuten, dass nicht gefördert wird, was – in offiziellen Prüfungen – nicht gefordert wird.

Sehr pessimistisch in dieser Hinsicht ist McNamara (2008: 11, Übers. H.G.), der bezüglich des GeR apodiktisch anmerkt: „Und wenn jemand versucht, hinsichtlich Curriculum oder Bewertung irgendetwas in sein Sprachlehrprogramm hineinzunehmen, was nicht in diese Begriffe [des GeR] übersetzt werden kann, so wird dieser Versuch unsichtbar und völlig wertlos“.

Die Tendenz, sprachsystematischen Kompetenzen nicht nur bei Prüfungen, sondern bereits im Unterricht der Grundstufe wenig Bedeutung zukommen zu lassen, lässt sich auch bei anderen Fremdsprachen beobachten. Bezüglich des fremdsprachlichen Englischunterrichts im Schulbereich gibt es eine linguistische Denkrichtung, die offen für „Globalesisch“ als Unterrichtsprinzip eintritt, d.h. für das Erlernen und Prüfen einer Lernaltersprache mit einfachen kommunikativen Funktionen statt des „Englischen“, wie Schmundt (2008) in seinem Spiegel-Artikel „Die Kunst des Stammelns“ berichtet. Die Frage, die sich angesichts dieses Konzeptes stellt, ist doch, ob Kinder, die angehalten werden, lediglich solch ein rudimentäres Englisch zu erlernen, später ein besseres Englisch noch nachlernen können, falls sie das wollen bzw. eine Notwendigkeit dazu besteht.

Dieselbe Frage muss auch im DaF-Bereich gestellt werden. Können für die Grundstufe veranschlagte sprachsystematische Kompetenzen, die vom Lerner in dieser Phase nicht oder nur bruchstückhaft erworben wurden, zum Bestehen des ZD aber nicht zwingend notwendig sind, später noch nachgelernt oder korrigiert werden?

Heike Gruhn (2009), Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptoren-Tabellen des Gemeinsamen Europäischen Referenzrahmens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

Erfahrungswerte zeigen, dass eine oft nach jahrelangem Lernen auf Grundstufenniveau fehlerhaft eingeschlossene Sprachsystematik später nur sehr schwer – manchmal auch gar nicht mehr – korrekturfähig ist. Lernaltersprachen können steckenbleiben, d.h. sich nicht mehr weiterentwickeln, und „unvollständige oder falsche Formen werden sich insbesondere dann verfestigen, wenn ein Lerner seine kommunikativen Absichten mit Hilfe dieser Formen durchsetzen kann“ (Apeltauer 1997: 115) – hinzuzufügen ist wohl auch (in Bezug auf den FU) „seine Prüfungsziele durchsetzen kann“. Bei manchen Lernern wird die Sprachentwicklung auf die unerlässlichen Elemente beschränkt bleiben bzw. fehlerhafte Formen umfassen, weil einfache Kommunikationsziele auf diese Weise erreichbar sind. Solche sich nicht weiterentwickelnden Strukturen der Lernaltersprache nennt man „Fossilisierungen“. Selinker (1972), Schöpfer des Konzepts, geht davon aus, dass die Lernaltersprache von rund 95% der Zweit- oder Fremdsprachenlerner langfristig Fossilisierungen aufweist und sich nicht bis zur Zielsprachlichen Norm weiterentwickelt. Unterschiede zwischen Lernern liegen dann weniger in der Tatsache begründet, dass sie Fossilisierungen aufweisen, sondern eher in der Frage, in welchem Umfang diese auftreten und in welchem Stadium ihrer *interlanguage* sie einsetzen.

Die Gründe für das Auftreten von Fossilisierungen mögen vielfältig und nicht in abschließender Weise erforscht sein. Sicher sind sie im Lerner selbst und ebenso im engeren und weiteren Lernumfeld zu suchen (soziopsychologische Aspekte). Sicher kann auch unqualifizierter Sprachunterricht einen Einfluss auf ihr Entstehen ausüben.

Wichtig ist letztendlich, welche Ziele mit dem Fremdsprachenlernen verbunden sind. So mag das ZD m.E. überall dort ein angemessenes Prüfungskonzept bieten, wo die Fremdsprache zu einfachen kommunikativen Zwecken gelernt wird, für private Kontakte zum Beispiel, für Reisen oder auch für bestimmte Migrationskontexte. Vorbehalte habe ich in Bezug auf den akademischen Bereich, da dort langfristig eine differenziertere Sprachbeherrschung zu Studien- oder Berufszwecken angestrebt wird bzw. angestrebt werden sollte, auf die von Anfang an hingearbeitet werden müsste¹⁴. Das ZD selbst gibt zwar keine akademische Zielsetzung vor, ist aber die einzige DaF-Prüfung für Erwachsene auf B1-Niveau, die in Mexiko im universitären Bereich überhaupt angeboten wird – Alternativen scheint es für diese Stufe ohnehin nicht zu geben. Die Mehrheit der Prüfungszentren des GI und des (billigeren) ÖSD (8 von insgesamt 13) befindet sich an Universitäten.

2.3 Kann „Kommunikative Kompetenz“ in Anteilen bescheinigt werden?

Sehen wir uns das Bewertungsraster für formale Richtigkeit beim mündlichen Ausdruck noch einmal an: Der Prüfling erhält 0 Punkte, wenn er „so viele Fehler in Syntax und Morphologie [macht], dass die Kommunikation zu scheitern droht bzw. scheitert“. Erhält er einen oder zwei Punkte – nämlich wenn das Verständnis an „mehreren“ oder „vielen“ Stellen erheblich beeinträchtigt ist – wird also nicht davon ausgegangen, dass die Kommunikation scheitert oder zu scheitern droht.

Diese Aussage ist denkwürdig, legt sie doch nahe, dass Kommunikation etwas sei, das aus mehreren Einzelkomponenten bestünde, von denen ein bestimmter Anteil einfach ausfallen dürfe. Wirklich scheitern tut sie nach diesem Verständnis erst dann, wenn der Pegel auf Null gesunken ist.

Kommunikative Kompetenz mag, wie oben ausgeführt, verschiedene einzelne Kompetenzen umfassen, und Kommunikationssituationen aus mehreren hintereinander folgenden Stellen bestehen. Nur dass diese Einzelkompetenzen ineinander übergreifen und einzelne Stellen miteinander verflochten sind. So kann nicht prinzipiell davon ausgegangen werden, dass Kommunikationssituationen auf zufriedenstellende Weise bewältigt werden, wenn in Teilen ihres Umfangs Falschinterpretationen, Missverständnisse, Unverständnis oder Unverständlichkeit auftreten.

In Lese- und Hörtexten sind einzelne Elemente oftmals so miteinander verzahnt, dass ein einziges Missverständnis bereits das Gesamtverständnis in Frage stellen kann. Eine einzige Fehlinterpretation kann auch einen Gesprächsverlauf sensibel stören. Das weiß jeder, der schon einmal einem Gespräch gelauscht hat, bei dem zwei Gesprächspartner – ohne es zu merken – über völlig unterschiedliche Dinge sprechen.

Heike Gruhn (2009), Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptoren-Tabellen des Gemeinsamen Europäischen Referenzrahmens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

In diesem Zusammenhang sind natürlich Strategien von Bedeutung (die in den Deskriptoren des GeR auch berücksichtigt werden). So kann z.B. durch gezielte Rückfragen das Scheitern der Kommunikation an einer bestimmten Stelle abgewendet werden. Der Einsatz solcher Strategien setzt allerdings immer schon voraus, dass Interaktionspartner sich potentieller Missverständnisse bewusst sind. Darüber hinaus wird er von Persönlichkeitsmerkmalen (schüchterne Personen werden seltener nachhaken) oder kulturellen Eigenheiten abhängen (manche Lerner erachten es als unhöflich, ältere oder ranghöhere Gesprächspartner zu unterbrechen) und ist insofern nicht einfach zu prüfen.

Auch aus den Deskriptoren des GeR lässt sich ersehen, dass Kommunikationssituationen häufig nicht aufteilbar sind. Viele Kann-Bestimmungen lassen nämlich nur zwei Möglichkeiten offen: Entweder man kann das, was dort beschrieben wird, oder man kann es nicht. So z.B.:

4.4.2.1 Hörverstehen allgemein.

Ankündigungen, Durchsagen und Anweisungen verstehen. B1: [...] Kann detaillierten Wegbeschreibungen folgen.

Versteht der Lerner nur ein Detail der Wegbeschreibung falsch (verwechselt er zum Beispiel die Wörter „rechts“ und „links“), wird er wahrscheinlich nicht ans Ziel gelangen.

Hier zeigt sich, was für ein schwieriges Unterfangen es ist, kommunikative Kompetenz als solche prüfen und anhand von Deskriptorentabellen beschreiben und verbürgen zu wollen. Prüfungen mögen Kompetenzen überhaupt nur ausschnittsweise abfragen können, doch wie verhalten sich dann zum Beispiel die Ergebnisse eines Kandidaten, der die Prüfung mit 60% besteht, der also nur Ausschnitte des Ausschnitts beherrscht bzw. erhebliche „Kommunikationslücken“ aufweist, zu Deskriptoren, deren Aussagen entweder zutreffen oder nicht, aber keine Halbwerte zulassen? Kommunikative Kompetenz auch dort verbürgen wollen, wo nicht alle ihrer Komponenten sicher gewährleistet werden können, ist eigentlich ein Widerspruch.

2.4 Braucht der Lerner professionellen DaF-Unterricht? Implikationen des Faktors „Zeit“

Wenden wir uns nun dem Faktor „Zeit“ zu, das heißt, der Frage, wie viele Kursstunden ein Lerner benötigt, um das ZD bestehen zu können. Nicht möglich ist, eine bestimmte Anzahl an Stunden allgemeingültig vorgeben zu wollen, da das Lernumfeld sowie die soziokulturellen und persönlichen Voraussetzungen der einzelnen Lerner sich stark voneinander unterscheiden.

In einer prüfungsvergleichenden Zeittabelle geht Bauer (2007: 31, 38-39) für das Zertifikat von 360 bis 450 Unterrichtseinheiten (à 45 Minuten) beim Lernen im Zielsprachenland (Immersion) aus, von rund 650 Einheiten im Land der Lerner (keine Immersion) und von mehr als 1200 Einheiten unter problematischen Bedingungen. In Mexiko wird das ZD, je nach Institution und Lerner, wenn überhaupt, dann oft erst nach 750-950 Zeitstunden (à 60 Minuten) abgelegt¹⁵, was 1000 oder mehr Unterrichtseinheiten à 45 Minuten entspricht. Das Pendel schlägt also schon in Richtung „problematische Bedingungen“ aus, zu denen Bauer (ebd.: 38) u.a. nicht ausgebildete Lehrkräfte, ungeeignete Lehrstrukturen, nicht aktualisierte Materialien und ungeeignete Methoden zählt.

Die Interlanguage-Theorie zeigt uns, dass Lernaltersprachen (mit den dazugehörigen Fehlern) etwas Natürliches sind. Sie macht aber keine Aussagen darüber, für wen und über welchen Lernzeitraum welcher Stand der Lernaltersprache noch als Übergangsphase zu betrachten ist, bzw. ab wann bereits von Fossilisierungen gesprochen werden kann, von Lernaltersprachen also, die sich nicht mehr weiterentwickeln.

Zu fragen ist, ob nach einem unverhältnismäßig langen Lernzeitraum auf Grundstufenniveau ein hohes Fehlerniveau noch als Durchgangsstadium betrachtet werden kann. Möglicherweise handelt es sich dabei schon längst um Fossilisierungen, d.h. um Fehler, die sich dauerhaft verfestigt haben.

Heike Gruhn (2009), Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptorentabellen des Gemeinsamen Europäischen Referenzrahmens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

Mit anderen Worten: Angesichts eines funktionalen Ansatzes, bei dem die Bewältigung von kommunikativen Aufgaben geprüft wird, welche auf diesem Niveau auch ohne gefestigte sprachsystematische Kompetenzen erfolgreich gelöst werden können, muss angenommen werden, dass die ZD-Prüfung in einigen Fällen ein bereits fossilisiertes Lernerstadium zertifiziert, das ein weiteres Voranschreiten zu anspruchsvolleren sprachlichen Kompetenzen und Sprachprüfungen gar nicht mehr erlaubt. Diesen Zusammenhang vermute ich vor allem dort, wo nach langen Lernzeiten lediglich ein ausreichender oder befriedigender ZD-Abschluss erreicht wird. Gestützt wird meine Vermutung von der Sachlage, dass in Mexiko nur ein verschwindend geringer Anteil an ZD-Absolventen nach der Prüfung weiterlernen bzw. eine Prüfung auf höherer Kompetenzstufe (B2)¹⁶ erfolgreich ablegen.

Das ZD macht keine Prognosen über zukünftige Entwicklungen der Lerner, einen Einfluss auf institutionelle Lernumfelder übt es nichtsdestotrotz aus. Keine Aussagen machen ist gleichbedeutend mit gewähren lassen. Eine B1-Prüfung, die bezüglich der Beherrschung der für das Grundstufenniveau veranschlagten sprachsystematischen Kenntnisse strengere Maßstäbe ansetzen würde, würde der Zertifizierung von Fossilisierungen zumindest stärker entgegenwirken, was wiederum einen Einfluss auf das Unterrichtsgeschehen hätte.

Meine These lautet, dass von der ZD-Prüfung in ihrer aktuellen Form keine Impulse zur Professionalisierung des Fachbereichs DaF in Mexiko ausgehen. Denn eine Professionalisierung¹⁷ mag aus institutioneller Sicht überflüssig erscheinen, da Teilerfolge¹⁸ auch unter Beibehaltung des Status Quo erzielt werden können.

2.5 Die Einbettung von Prüfungen in ideologische Interessen

Die Rolle von offiziellen Sprachprüfungen – das ist heute unter Fachleuten ein Gemeinplatz – geht weit über ihre Qualifikations- und Selektionsfunktionen hinaus. „Tests, [...] are not neutral but rather embedded in political, social, educational, ideological and economic contexts“, konstatiert Shohamy (2001: 114).

Sucht man nach solchen ideologischen Kontexten, so stößt man angesichts der gegenwärtigen Tendenz, das Prüfungsgeschehen weltweit auf die Standards des GeR einzuschwören, bald auf die Frage nach demokratischen Entscheidungsprozessen. Haben denn Lehrkräfte und Institutionen sich überall freiwillig für diese Standards entschieden? Oder bedarf es dazu bildungspolitischer Überzeugungsarbeit bzw. schlichtweg des Zwangs von oben? Sehr deutliche Worte findet McNamara vom *Language Testing Research Center, University of Melbourne*, Australien. Er zeigt sich äußerst skeptisch, was die Überzeugungsarbeit im Bildungssektor angeht:

Wenn die Lehrkräfte über die Standards [des GeR] unterrichtet werden, wird ihnen gesagt, dass es deren Zweck sei, den Lehrer zu unterstützen und die Studenten zu unterstützen. Das stimmt aber nicht: Ihr Zweck besteht darin, die Verwaltung des Systems zu unterstützen (2008: 8, Übers. H.G.¹⁹).

Und mit System – so meine Interpretation – sind hier globale politische und wirtschaftliche Interessen gemeint, die auf die Hilfe von bildungspolitischen Weichenstellungen angewiesen sind. Der GeR ist folglich ein Instrument, das laut McNamara (vgl. ebd.: 11) denjenigen Macht verleiht, die darauf bestehen, dass Prüfungsergebnisse in Form einer Reihe an Behauptungen präsentiert werden.

Und dieses Instrument, so konstatiert er, mutiert zunehmend zum einzigen, was beim Sprachunterricht und folglich bei der Leistungsbewertung – weltweit – noch zählt! „Das ist gar nicht gesund; überhaupt nicht gesund“, klagt der Forscher.

Dieser Einschätzung möchte ich mich anschließen. Es wird zwar immer wieder beteuert, dass der GeR nichts vorschreibe und die Entscheidungsgewalt bezüglich Lehrplänen und Unterrichtsgestaltung bei den Fachbereichen bzw. Kursanbietern bleibe. Das stimmt so aber nicht. Es gibt ja gar keine offizielle DaF-Prüfung mehr, deren Inhalte und überprüftes Kompetenzniveau nicht mit den Kann-Beschreibungen des GeR in Beziehung stehen würden – allein das schränkt den Entscheidungsspielraum von vorneherein ein. Wenn sich außerdem auf institutioneller Ebene ein Druck hin zu mehr vorzeigbarer Effizienz – in Form von bestandenen offiziellen Prüfungen²⁰ - abzeich-

Heike Gruhn (2009), Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptoren-Tabellen des Gemeinsamen Europäischen Referenzrahmens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

net, so ergibt sich daraus praktisch zwangsweise, dass Lehrpläne und Unterricht weitgehend auf die Prüfungsansprüche zugeschnitten und an die dort zugrundeliegenden (GeR-basierten) Kriterien angepasst werden. Entscheidungsfreiheit bleibt den einzelnen Lehrkörpern dann nur noch in Hinsicht auf Details.

Es geht hier nicht darum, den Deskriptorentabellen Nützlichkeit abzusprechen, der GeR und dessen Standards sind sicher in vielerlei Hinsicht und in vielen Kontexten nützlich und anregend. Das Erschreckende ist wohl, dass der Rahmen im Begriff ist, sein erklärtes Ziel der Vergleichbarkeit nationaler Lehr-, Lern- und Testtraditionen im Fremdsprachenunterricht, so umfassend zu erreichen, dass – trotz gegenteiliger Beteuerungen – eine Tendenz hin zur Gleichschaltung besteht. Die Frage ist auch, wer in Zukunft noch die Macht haben wird, etwaige Änderungen am GeR und den Deskriptoren (mit Auswirkungen auf offizielle Prüfungen und daran angelehnte Lehrpläne) durchzusetzen. Im Moment sieht es so aus, als ob dies nur noch von oben, das heißt seitens des Europarates geschehen könne.

Für McNamara (ebd.: 12, Übers. H.G.) bedeutet all das eine Art geistige Verengung im Bereich der Fremdsprachenlehre: „Es begrenzt unser Nachdenken bezüglich der Ziele des Fremdsprachenunterrichts. Es begrenzt unser Nachdenken bezüglich der Dinge, die wir sonst noch von den Lernern erwarten möchten.“

Das ist der ideologische Rahmen, innerhalb dessen sich auch das ZD bewegt, ja in der Gegenwart bewegen muss, um bestehen zu können.

Auch die finanziellen Interessen, mit denen offizielle Prüfungen verbunden sind, liegen jenseits von deren Selektions- oder Qualifikationsfunktion. Eine Frage, die selten gestellt wird und die wir in diesem Zusammenhang lediglich in den Raum stellen möchten, ist die folgende: Wie verhält sich die Tatsache, dass offizielle Prüfungen „verkäufliche Produkte“ sind, zum Design, zur Durchführungsobjektivität und zu den Bewertungsparametern dieser Prüfungen, aber auch zur Vorbildung der Prüfer? Anzunehmen ist doch, dass es für Prüfungsanbieter betriebswirtschaftlich erstrebenswert ist, an möglichst vielen Orten möglichst viele Prüflinge zu prüfen. Und dabei mag es auch vorteilhaft sein, wenn möglichst viele Anwärter die Prüfungen bestehen. Hier gibt es im DaF-Bereich sicher noch Forschungsbedarf.

3. Wer trägt die Verantwortung für die Nutzung von ZD-Prüfungsergebnissen in der mexikanischen Berufswelt?

Die positiv formulierten Kann-Beschreibungen als zentrale Parameter beim Prüfungsgeschehen dienen u.a. dem Zweck, das Selbstvertrauen des Prüflings zu stärken, indem man sich nicht auf noch vorhandene Defizite, sondern eben auf das konzentriert, was schon „gekonnt“ wird. So weit, so gut. Gefragt werden muss dennoch, ab welchem Punkt das Zertifizieren früher Stadien einer *interlanguage* sinnvoll ist bzw. welchem Zweck diese dienen soll oder wird. Prüfungen werden ja in der Regel nicht nur zur Selbstbestätigung oder Selbstmotivierung abgelegt, sondern sind – das zeigen zumindest informelle Umfragen in Mexiko – mit der Hoffnung auf eine bessere Positionierung auf dem Arbeitsmarkt verbunden. Nun stellt das ZD für viele mexikanische Lerner den glorreichen Abschluss ihres Lernweges der deutschen Sprache dar, nur ein Bruchteil legt auch noch eine Prüfung auf höherer Kompetenzstufe (B2) ab. Mit anderen Worten: Der Arbeitsmarkt trifft vor allem auf A2- oder B1-Absolventen. Infolge eines Mangels an qualifizierten deutschsprachigen Fachkräften in Sprach- bzw. Sprachmittlerberufen (und natürlich auch aus finanziellen Gründen) werden diese unter anderem mit folgenden Aufgaben betraut: Erteilung von Deutschkursen, Firmenkorrespondenz oder Übersetzungen. Mir ist sogar ein Fall bekannt, in dem ein B1-Absolvent zum Leiter der Deutschabteilung einer großen Privatuniversität ernannt wurde.

Sprachprüfungen und deren Ergebnisse wirken nicht nur auf den Sprachunterricht zurück (*washback*-Effekt), sondern auch in das soziale Umfeld hinein, was vor allem in der nordamerikanischen Forschung unter dem Stichwort *consequential validity* (Messick 1989) diskutiert wird: „[...] Messick also explicitly advanced a critical role for value implications and social consequences, particularly evaluation of intended and unintended social consequences of test interpretation and use, as part of test validity“ referiert Kunnan (2005: 780). In diesem Zu-

Heike Gruhn (2009), Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptorentabellen des Gemeinsamen Europäischen Referenzrahmens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

sammenhang ist nicht nur in Betracht zu ziehen, wer sich der Prüfungsergebnisse zu welchem Zweck bedient, es muss auch der Frage nach der Verantwortlichkeit für beabsichtigte und nicht beabsichtigte Folgeerscheinungen nachgegangen werden:

„The responsibility of the language testers ist clear: we must accept responsibility for all those consequences which we are aware of“, lautet die Stellungnahme von Liz Hamp-Lyons zu diesem Thema (zit. nach Shohamy 2001: 117). Diese Argumentationsweise impliziert, laut Shohamy (ebd.), auch die Möglichkeit „to impose sanctions, punish and forbid the use of tests which are misused [...]“.

Das geht meiner Meinung nach zu weit. In vielen Ländern müssten offizielle Sprachprüfungen dann gleich ganz abgeschafft werden, weil ihre Ergebnisse regelmäßig zweckentfremdet genutzt werden. Doch es kommt ja auch niemand auf die Idee, von Autoherstellern zu verlangen, keine Autos mehr zu produzieren, nur weil manche Leute mit diesen Unfälle verursachen.

In Bezug auf das ZD und dessen Nutzung in Mexiko möchten wir uns an die Ausführungen von Davies (1997: 335) und Farhady (1998) anlehnen, die nicht davon ausgehen, dass Prüfungsanbieter für alle möglichen sozialen Folgen rund um die von ihnen angebotenen Prüfungen aufkommen können. „What kind of use the agencies would make from the test scores is often a matter of sociopolitical concerns rather than an academic endeavor“, konstatiert Farhady (1998: 11).

Ein Nachweis über die Fähigkeit als Sprachlehrer oder Sprachmittler zu arbeiten ist mit dem ZD sicher nicht verbunden und wird seitens der Prüfungsanbieter nicht behauptet. Die Art der Nutzung, die die Prüfungsergebnisse des ZD erfahren und die in dieser Form von den Anbietern nicht beabsichtigt ist, ergibt sich aus der Eigendynamik eines spezifischen soziokulturellen Kontextes. Nachfolgend möchte ich einige Gründe auflisten, warum die Prüfungsanbieter nicht für alle in Mexiko mit dem ZD verbundenen Folgen einstehen können:

- Die Nutzung der Ergebnisse findet innerhalb eines spezifischen politisch-rechtlichen, d.h. nationalstaatlichen Systems statt, innerhalb dessen die europäischen Prüfungsanbieter keine Entscheidungsbefugnis haben.
- Entscheidungen bezüglich der Einstellung von ZD-Absolventen für bestimmte berufliche Tätigkeiten werden auf der Ebene einzelner Institutionen (Bildungsinstitutionen oder Unternehmen) und dort oftmals von Einzelpersonen getroffen. Es handelt sich also um eine Vielzahl an individuellen Entscheidungsträgern. Überinstitutionelle Entscheidungsträger stehen insofern nicht als Ansprechpartner zur Verfügung.
- Die Tatsache, dass ZD-Absolventen zu den oben erwähnten Tätigkeiten herangezogen werden, hat in der Regel wenig oder gar nichts mit konkreten Merkmalen des ZD zu tun. Angesichts eines Marktes, auf dem das Angebot an qualifizierten Fachkräften der Nachfrage an Deutschkursen deutlich hinterherhinkt, geht es vielen Kursanbietern darum, überhaupt jemanden zu finden, der einen Nachweis über Deutschkenntnisse vorlegen kann. Hat man die Wahl zwischen einem B1- und einem A2-Absolventen, wird man ersteren bevorzugen. In einigen Fällen wird auf einen solchen Nachweis aber auch ganz verzichtet.

„In spite of the fact that language testers are at the center of the process of language testing, they do not have direct access to the parameters of decision-making. Politicians, educators, bureaucrats, and so many other sectors in the society are in the position to make decisions“, können wir das Vorgesagte mit Farhady (1998: 16) zusammenfassen.

Ganz aus der Mangel genommen werden können die Prüfungsanbieter dennoch nicht. In ihrer Auflistung verschiedener wissenschaftlicher Positionen zum Thema, nennt Shohamy (2001: 116) auch die folgende, der ich mich anschließen möchte:

Heike Gruhn (2009), Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptoren-Tabellen des Gemeinsamen Europäischen Referenzrahmens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

[...] the only responsibility a tester has with regards to the use of tests is to *point out* to society the intentions, effects and consequences of tests. Accordingly, the role of the tester is to collect data about dimensions and aspects of use and make this information known and accessible to the users and the public at large.

Es gilt also, einer inadäquaten Nutzung von Prüfungsergebnissen so weit wie möglich vorzubeugen, indem aussagekräftige und detaillierte Beschreibungen des entsprechenden Kompetenzniveaus verbreitet werden. Auch über Sinn und Zweck der Prüfung sollten die Nutzer sehr konkrete Informationen erhalten.

Wird denn in den Werbeschriften der Prüfungsanbieter eine berufliche Verwertbarkeit der B1-Sprachkompetenzen versprochen und wenn ja, in welchem Rahmen? Zu Beginn dieser Arbeit sahen wir bereits, dass sich das GI diesbezüglicher Aussagen ganz enthält, das ÖSD im ZD aber einen „Nachweis der Fähigkeit zur selbstständigen Sprachverwendung in Alltags- und Berufssituationen“ (<http://www.osd.at>) sieht. Diese Aussage wird nicht weiter erläutert, was vielerlei Interpretationsmöglichkeiten offen lässt. Es sind sicher Berufssituationen denkbar, die ein Lerner auf B1-Niveau meistern kann, so etwa Empfangs- oder Bedienungstätigkeiten im touristischen Bereich oder Gaststättengewerbe. Es sind aber eben auch viele andere Berufssituationen denkbar, in denen die vom ZD bescheinigten Kompetenzen durchaus nicht ausreichend sein werden, denn Berufssituationen können ja – wie wir eingangs schon feststellten – völlig Unterschiedliches beinhalten.

Rückendeckung erhält das ÖSD in diesem Zusammenhang von den Deskriptorentabellen des Referenzrahmens, denen in Bezug auf die Nutzung der Prüfungsergebnisse ein besonderer Stellenwert zukommt: Für viele potentielle Arbeitgeber sind sie nämlich die einzige Informationsquelle, die in einer ihnen verständlichen Sprache (hier Spanisch) abgefasst ist. Tatsächlich wird man in diesen Tabellen an vielen Stellen fündig, immer wieder ist die Rede von „berufsbezogenen Themen“ „eigenem Fachgebiet“ usw., wobei auch die dortigen Aussagen oft wenig präzise sind. Hier nur ein Beispiel von vielen:

4.4.1.2 B1: Berichte und Aufsätze schreiben: [...] Kann im eigenen Sachgebiet mit einer gewissen Sicherheit größere Mengen von Sachinformationen über vertraute Routineangelegenheiten und über weniger routinemäßige Dinge zusammenfassen, darüber berichten und dazu Stellung nehmen.

Was bedeutet „mit einer gewissen Sicherheit“? Verwirrend ist auch, dass Sachinformationen über routinemäßige und weniger routinemäßige Dinge gleichzeitig einbezogen werden. Dass das Zusammenfassen von Informationen vielen Lernern selbst in der Muttersprache Schwierigkeiten bereitet, ist nicht mitbedacht worden.

Zusammenfassend kann festgestellt werden, dass der GeR für diese Stufe tatsächlich ein relativ umfassendes berufliches Sprachhandlungsvermögen vorsieht, solange es sich um „unkomplizierte Inhalte“ und „einfache Standardsprache“ handelt. Was genau unter diesen Konzepten zu verstehen ist, bleibt wiederum den Interpretationen der einzelnen Nutzer überlassen.

Dem ZD (bzw. der Kompetenzstufe B1) auf unspezifische Art und Weise berufliche Verwertbarkeit zu bescheinigen, halte ich für sehr gewagt. Nicht nur, weil es so viele verschiedene Berufswelten wie Berufe gibt, jede mit ihrem eigenen sprachlichen Anforderungsprofil, sondern auch weil es im Arbeitsleben immer wieder zu Situationen kommt, in denen präzises Erfassen oder Produzieren von mündlicher oder schriftlicher Information von großer Bedeutung sein – und Missverständnisse mitunter erheblichen (wirtschaftlichen) Schaden anrichten können.

Dass im täglichen Berufsleben solche, sprachliche Genauigkeit erfordernde Situationen von anderen, „unkomplizierten, auf einfacher Standardsprache basierenden“, immer klar getrennt werden können, um den „Deutschkönnen auf B1-Niveau“ nicht zu überfordern, wage ich zu bezweifeln. Insofern können berufsbezogene Aussagen bezüglich des Kompetenzniveaus dazu führen, dass bei Lernern und Arbeitgebern falsche Erwartungen geweckt werden.

Da nicht sichergestellt werden kann, dass potentielle Arbeitgeber die Internetseiten der Prüfungsanbieter einsehen, sollten wichtige Information (mitsamt einschlägiger Internetadressen) dort angebracht werden, wo sie am leichtesten zu finden ist: Als Anhang zur Prüfungsurkunde oder auf deren Rückseite. Angesichts der Tatsache, dass Prüfungsergebnisse immer wieder zweckentfremdet verwendet werden, bleibt den Prüfungsanbietern immerhin der Spielraum, ausdrücklich darauf hinzuweisen, welche Art von Kompetenzen das Niveau nicht bescheinigt – auch wenn dabei der Grundsatz der „positiven Formulierungen“ einmal zurückgestellt werden müsste.

4. Ausblick

Wer am Busbahnhof in Puebla ankommt, stößt auf mehrere große Schilder mit folgendem Wortlaut:

Für Ihre eigene Sicherheit benutzen und kontrollierten Taxis durch CAPU. Kaufen Sie Ihre Ticket direkt am Schalter.

Diese Sätze entsprechen den Kriterien des schriftlichen Ausdrucks B1, d.h. denkbar ist, dass sie von einem ZD-Absolventen verfasst wurden: Sind sie doch der „unkomplizierten Alltagssprache“ zuzurechnen und gerade noch verständlich, vor allem wenn man den unmittelbaren Kontext (neben einem Schalterhäuschen stehen Taxis in der Warteschlange, allesamt mit der Aufschrift „CAPU“) zu Hilfe nimmt. Reicht es nun aber in dieser Situation aus, wenn die Aussage verständlich ist? Mitnichten! Es muss in Betracht gezogen werden, dass der Busbahnhof für viele deutschsprachige Besucher der erste Anlaufpunkt in Puebla ist, also eine Art Repräsentativfunktion hat. Diese Schilder tragen zu einem ersten Eindruck von Stadt und Bewohnern bei und führen bei manch einem auch zu einem ersten Urteil. Unkomplizierte Inhalte und einfache Standardsprache? Ja! – Dennoch wird diese Kommunikationssituation von B1-Beschreibungsbarometern nicht erfasst!

Das Beispiel zeigt, wie groß das Vertrauen in Deutschlerner und deren Diplome sein kann. Hier wurde eine sicher nicht unbedeutende Summe ausgegeben, ohne daran zu denken, den sprachlichen Inhalt der Schilder überprüfen zu lassen.

Am Beispiel des ZD hoffe ich deutlich gemacht zu haben, dass man sich nicht vom Glanz des neuen Prüfungskonzeptes blenden lassen sollte. Die Prüfung ist durchaus noch verbesserungsfähig, und ihre Ergebnisse spiegeln sich nicht durchweg in den Kann-Beschreibungen des GeR, ebenso wenig können letzere ein sicheres Profil von Prüfungsabsolventen liefern. Anzunehmen ist, dass diese Befunde auch für andere GeR-basierte Fremdsprachenprüfungen symptomatisch sind. Bei der gegenwärtig zu beobachtenden Expansionsdynamik solcher Prüfungen in Mexiko, die sich u.a. in neuen DaF-Prüfungszentren niederschlagen, aber auch in intensiver Überzeugungsarbeit mittels Konferenzen und Kursen zum GeR, wird überspielt, dass wir uns mit all dem eigentlich immer noch in der Versuchsphase befinden. Kritische Stimmen scheinen angesichts des vorherrschenden Trends nicht wirklich erwünscht zu sein, sollten aber keinesfalls verstummen.

Literatur

Adkins Wood, Dorothy (1990), *Elaboración de tests*. Mexiko: Trillas.

Altmayer, Claus (2004), Sprachkultur und Mehrsprachigkeit: Neuerscheinungen zur europäischen Sprachenpolitik (Teil 3). *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 9: 2, 10 S. [Online: http://spz1.spz.tu-darmstadt.de/projekt_ejournal/jg-09-2/beitrag/sprachenpolitik2.htm. Dezember 2008.]

Alyaz, Yunus & Freibichler, Hans (2005), *Software für computergestütztes Sprachenlernen auswählen*, 38 S. [Online: <http://www.fts-heidelberg.de/18-K-5-10.pdf>. Dezember 2008.]

Apeltauer, Ernst (1997), *Grundlagen des Erst- und Fremdspracherwerbs*. Fernstudieneinheit 15. München: Langenscheidt.

Heike Gruhn (2009), Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptoren-Tabellen des Gemeinsamen Europäischen Referenzrahmens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

- Bauer, Ulrich (2007), *Profesionalismo y éxito en la enseñanza del alemán en México*. In: AMPAL/DAAD/UNAM (Hrsg.) (2007), *El idioma alemán en México. Medio de enlace, oportunidades y prestigio. Hacia un futuro trilingüe*. Mexiko: o.V., 29-43.
- Butzkamm, Wolfgang & Butzkamm Jürgen (2004), *Wie Kinder sprechen lernen. Kindliche Entwicklung und die Sprachlichkeit des Menschen*. Tübingen, Basel: Francke.
- Clalüna, Monika (2004), *Gerahmte Sprachen? Zum Gemeinsamen Europäischen Referenzrahmen für Sprachen. DaF-Brücke* 6, 24-27.
- Da Silva Gomes, C. Helena María & Signoret Dorcasberro, Aline (2005), *Temas sobre la adquisición de una segunda lengua*. Mexiko: Trillas.
- Davies, Alan (1997), *Demands of being professional in language testing. Language Testing* 14: 3, 328-339.
- Europarat (2001), *Gemeinsamer europäischer Referenzrahmen für Sprachen: Lernen, lehren, beurteilen*. Berlin: Langenscheidt. [Online: <http://www.goethe.de/z/50/commeuro/deindex.htm>. Oktober 2009.]
- Farhady, Hossein (1998), *Ethics in Language Testing*. Vortrag am Summer Institute on The Social Responsibility of Language Testers, Carleton University, Ottawa (Manuskript). [Online: http://www.aua.am/academics/dep/hf_publications/4_Ethics.pdf. Juni 2009.]
- Glaboniat, Manuela; Müller, Martin; Rusch, Paul; Schmitz, Helen & Wertenschlag, Lukas (2008), *Profile Deutsch. Lernzielbestimmungen, Kann-Beschreibungen und kommunikative Mittel für die Niveaustufen A1, A2, B1, B2, C1 und C2 des „Gemeinsamen europäischen Referenzrahmen für Sprachen“*. Buch und CD-ROM. München: Langenscheidt.
- Grotjahn, Rüdiger (2007), *Testen und Prüfen: Aktuelle Tendenzen. Neue Beiträge zur Germanistik* 6: 2, 19-38.
- Hamps-Lyon, Liz (1998), *Ethical test preparation practice: The case of the TOEFL. TESOL Quarterly* 32: 2, 329-337.
- Klatt, Karola (2008), *Lernen auf einem anderen Stern. Kulturaustausch* 58: 4, 81.
- Kunnan, Antony John (2005), *Language assessment from a wider context*. In: Hinkel, Eli (Hrsg.) (2005), *Handbook of Research in Second Language Teaching and Learning*. Mahwah, New Jersey: Lawrence Erlbaum, 779-794.
- McNamara, Tim (2008), *Language Assessment as Social Practice*. Vortrag vom 05. August 2008 beim IV Coloquio CELU, Universität San Martín, Buenos Aires (Manuskript). [Online: http://www.celu.edu.ar/images/stories/pdf/McNamara_conferencia.pdf. Oktober 2009.]
- Messick, Samuel (1989), *Validity*. In: Linn, Robert (Hrsg.) (1989), *Educational Measurement*. New York: Macmillan, 13-103.
- Österreichisches Sprachdiplom Deutsch (ÖSD) (2004), *Allgemeine Informationen zu den ÖSD-Prüfungen*. Wien: ÖSD. [Online: <http://www.osd.at>. Oktober 2009.]
- Rea-Dickins, Pauline (2001), *Fossilisation or evolution: The case of grammar testing*. In: Elder, Catherine et al. (Hrsg.) (2001), *Experimenting with Uncertainty. Essays in Honour of Alan Davies*. Cambridge: Cambridge University Press, 22-32.
-
- Heike Gruhn (2009), *Anmerkungen zum Zertifikat Deutsch im mexikanischen Kontext vor dem Hintergrund der Deskriptoren-Tabellen des Gemeinsamen Europäischen Referenzrahmens. Zeitschrift für Interkulturellen Fremdsprachenunterricht* 14: 2, 55-75. Abrufbar unter <http://zif.spz.tu-darmstadt.de/jg-14-2/beitrag/Gruhn.pdf>.

- Schifko, Manfred (2001), Prüfungen, Zertifikate, Abschlüsse als Planungskategorien für den Unterricht. In: Helbig, Gerhard; Götze, Lutz; Henrici, Gert & Krumm, Hans-Jürgen (Hrsg.) (2001), *Deutsch als Fremdsprache. Ein internationales Handbuch* (2. Bd.). Berlin, New York: de Gruyter, 827-834.
- Schmidt, Maria Gabriela (2007), EU-Sprachenpolitik, die Rolle der deutschen Sprache und der Gemeinsame Europäische Referenzrahmen. Konzept von „oben“ – Praxis von „unten“. Mit einem Blick nach Japan. *Neue Beiträge zur Germanistik* 6: 2, 116-132.
- Schmundt, Hilmar (2008), Die Kunst des Stammelns. *Spiegel Online* [Online: <http://www.spiegel.de/spiegel/0,1518,544420,00.html>. März 2008.]
- Selinker, Larry (1972), Interlanguage. *IRAL* 10: 2, 209-231.
- Shohamy, Elana (2001), The social responsibility of the language testers. In: Cooper, Robert L.; Shohamy, Elana & Walters, Joel (Hrsg.) (2001), *New Perspectives and Issues in Educational Language Policy. In Honour of Bernard Dov Spolsky*. Amsterdam, Philadelphia: John Benjamins, 113-130.
- Völzing, Paul-Ludwig (1995), Wieviel Grammatik brauchen Lehrer und Schüler? Grammatik im fremdsprachlichen Unterricht. *Info DaF* 22: 5, 509-527.

Anmerkungen

- ¹ In diesem Text werden die Begriffe “Deskriptorentabellen”, “Kann-Beschreibungen” und “Skalen” synonym verwendet.
- ² Laut Altmayer (2004: 3) kamen die Spezialisten, die auf der 22. *Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts* zum Ansatz des GeR befragt wurden, praktisch einhellig zu dem Ergebnis, dass “Eine eigene fremdsprachenlerntheoretische Position [...] nicht erkennbar [sei], sie sei allenfalls inkohärent, eklektisch, vage oder unverbindlich“.
- ³ Weitere Trägerinstitutionen des ZD sind die Universität Fribourg (CH) und Weiterbildungs-Testsysteme / TELC. Da in Mexiko aber nur das GI und das ÖSD als ZD-Anbieter operieren, beschränken sich die Ausführungen der vorliegenden Arbeit auf diese zwei Institutionen.
- ⁴ Siehe dazu Glaboniat, Manuela; Müller, Martin; Rusch, Paul; Schmitz, Helen & Wertenschlag, Lukas (2008), *Profile Deutsch. Lernzielbestimmungen, Kann-Beschreibungen und kommunikative Mittel für die Niveaustufen A1, A2, B1, B2, C1 und C2 des „Gemeinsamen europäischen Referenzrahmen für Sprachen“*. Buch und CD-ROM. München: Langenscheidt.
- ⁵ Z.B. eine detaillierte Analyse der Bewertungsparameter beider Anbieter (Anzeichen zu Unterschieden bestehen), sowie der konkreten Handhabung vor Ort (Wie und in welchem Ausmaß greifen die Prüfer in das mündliche Prüfungsgespräch ein? Kennen die Prüfer die Prüflinge? u.a.). Darüber hinaus wäre auch zu untersuchen, welches fachliche Ausbildungsprofil die Prüfer beider Anbieter haben, außerdem müsste der Schwierigkeitsgrad der Aufgaben der einzelnen Prüfungssätze verglichen werden.
- ⁶ Auf diese Unterscheidung werde ich nachfolgend aber nicht weiter eingehen, da im hiesigen Zusammenhang ausschlaggebend ist, dass die Gesamtheit aller Deskriptoren die Niveaustufe beschreiben soll.
- ⁷ *Profile Deutsch* (2.0, Info zu Kann-Beschreibungen [CD-ROM]).

- ⁸ Das ist u.a. der Fall bei den monologischen Kompetenzen des Lernalers (mündlicher Ausdruck 4.4.1.1.); seinen audiovisuellen Kompetenzen (Verstehen von Fernsehsendungen und Filmen 4.4.2.3.) sowie seiner Fähigkeit, Konferenzen und Vorträge zu verstehen (Hörverstehen 4.4.2.1.).
- ⁹ Sinngemäß: Wer viel umfassen möchte, hält wenig wirklich fest.
- ¹⁰ Der GeR sieht nicht die klassischen vier, sondern acht Fertigungsbereiche vor, die dort allerdings "Aktivitäten" genannt werden: Produktion, Rezeption, Interaktion, Sprachmittlung, jeweils mündlich und schriftlich, wobei die Sprachmittlung nicht in Deskriptoren-Form aufgeschlüsselt ist.
- ¹¹ In diesem Unterkapitel am Beispiel der Vorgaben des Goethe-Instituts, da diese im Internet öffentlich zugänglich sind, die des ÖSD sind m.E. ähnlich. Vgl. http://www.goethe.de/lrn/prf/pro/ZD_Prueferbl_UES04.pdf, Übungssatz 04, Prüferblätter, S. 10 u. 17. Siehe ebenso: Prüfungsordnung ZD.
- ¹² Bei anderen Deskriptoren zur Sprachsystematik mag es sich analog verhalten, ich habe die Grammatik hier nur als Beispiel ausgewählt und nicht alle einschlägigen Deskriptoren untersucht.
- ¹³ Die Adjektivdeklinationsform im Dativ ist auch folgerichtig in der funktionalen Grammatik der Niveau-Stufe bei Profile Deutsch nicht zu finden. Es soll im Übrigen hier nicht unterstellt werden, dass *Profile Deutsch* diese oder andere grammatikalische Elemente verbindlich vorschreiben wolle. Das Werk (CD-ROM Version 2.0, Grammatikübersicht) erhebt „nicht den Anspruch zu benennen, welche Strukturen ein Lerner auf einem bestimmten Niveau fehlerfrei anwenden kann oder können soll“.
- ¹⁴ In einer vom DAAD in Auftrag gegebenen Studie wurde festgestellt, dass rund 50% der ausländischen Studenten in Deutschland ihr Studium abbrechen. Einer der Gründe wird in unzureichenden Sprachkenntnissen gesehen, und das trotz formaler Sprachnachweise (vgl. Klatt 2008: 81).
- ¹⁵ Bauer (2007: 39) spricht von 750 bis 1000 Stunden (gemeint sind wohl Unterrichtseinheiten von 45 Minuten) und hofft auf eine Professionalisierung des Lehrpersonals, um auch in Mexiko die Zielvorgabe B1 eines Tages in 450 Stunden erreichen zu können.
- ¹⁶ Untersucht werden müsste in einer weiteren Arbeit, welche Ansprüche die B2-Prüfungen stellen und ob hier Sprachsystematik stärker ins Gewicht fällt. Anzunehmen ist natürlich, dass es verschiedene Gründe gibt, warum Lerner nach jahrelanger Anstrengung bei Erreichen des ZD beschließen, den Lernweg nicht fortzusetzen bzw. dabei scheitern.
- ¹⁷ Genaue Daten bezüglich des Ausbildungsprofils der Lehrkräfte und der Anzahl an Prüfungsabsolventen auf den verschiedenen Stufen einschließlich Zeitaufwand liegen leider nicht vor. Auch im 2007 erschienenen Band AMPAL/DAAD/UNAM (Hg.), *El idioma alemán en México. Medio de enlace, oportunidades y prestigio. Hacia un futuro trilingüe* sind sie nicht enthalten. Das liegt wohl u.a. daran, dass diese Art von Daten (vor allem was die Lehrkräfte betrifft) nicht gerne von den einzelnen Institutionen nach außen getragen werden. Als Beleg kann hier also lediglich ein weiteres – aus subjektiver Bestandsaufnahme gewonnenes – Zitat angeführt werden:

„Warum werden denn immer noch einfach muttersprachliche Sprecher der zu unterrichtenden Sprachen angeheuert, wenn diese Muttersprachler doch nie gelernt haben, ihre Sprache zu unterrichten? Die Behauptung, dass eine Person zum Unterrichten einer Sprache qualifiziert sei, nur weil sie sie spricht, ist ein Ammenmärchen, das sich nicht nur hartnäckig hält, sondern darüber hinaus auch völlig falsch ist. Und auch wenn man diese Dummheit immer wieder wiederholt, wird sie dadurch doch nicht wahrer. [...] In Notsituationen und auf Grund des enormen Mangels an Sprachlehrern in Mexiko [gemeint sind wohl

qualifizierte Sprachlehrer, H.G.], greift man von Zeit zu Zeit auf nicht ausgebildete Lehrkräfte zurück, was sich langfristig als vollkommen unverantwortlich herausstellt und zu der äußerst unbefriedigenden Lage des Fremdsprachenunterrichts in diesem Land beiträgt“ (Bauer 2007: 29f., Übers. H.G.).

- ¹⁸ Die Abbrecherquoten bei Deutschlernern im Erwachsenenbereich sind allgemein sehr hoch, viele geben bereits während der ersten Kurse wieder auf. Auch von den Deutschlernern, die sämtliche Grundstufenkurse abschließen, legen nicht alle die ZD-Prüfung ab.
- ¹⁹ Im spanischen Original wird der aktive Agens “cooperar” verwendet, den ich mir hier erlaubt habe, mit “unterstützen” zu übersetzen.
- ²⁰ Auch wenn es dabei in Mexiko hauptsächlich um A2-Prüfungen und nicht um das ZD geht.