Referierte Beiträge

Martin Baethge / Lena Arends

Die Machbarkeit eines internationalen Large-Scale-Assessment in der beruflichen Bildung: Feasibility Study VET-LSA

Eine komparative Analyse von Ausbildungsinhalten und Berufsprofilen in acht Europäischen Ländern

1. Die Besonderheit der Berufsbildung als Herausforderung für large-scale-assessments

Die Frage nach der Machbarkeit eines internationalen Large-Scale-Assesment (LSA)[1] von Ausbildungsleistungen ist in der beruflichen Bildung sehr viel schwerer zu beantworten als im Bereich der Allgemeinbildung. Die vergleichweise größeren Schwierigkeiten für ein Large-Scale-Assessment in der Berufsbildung liegen im Wesentlichen in zwei eng miteinander verknüpften Sachverhalten begründet, die u. E. die zentralen methodischen Herausforderungen für ein „Berufsbildungs-Large-Scale-Assessment" (VET-LSA)[2] abgeben: zum einen der spezifische Typus von Wissen und Kompetenz, der in der beruflichen Bildung vermittelt wird; zum anderen die enge Einbindung der Berufsbildung in die je nationalen Arbeitsmarkt-, Arbeitsteilungs- und Berufsstrukturen:

1 Der Artikel stellt das Konzept und einige Ergebnisse der Feasibility Study für ein international vergleichendes Large-Scale-Assessment in der beruflichen Bildung (unterhalb der Hochschulebene) dar, die wir in Kooperation mit Experten aus acht europäischen Ländern zwischen Juli 2008 und Mai 2009 auf der Basis von Vorarbeiten zum Konzept eines „Berufsbildungs-Large-Scale-Assessment" (vgl. Baethge, Achtenhagen & Arends, 2008; Baethge, Achtenhagen, Arends, et al., 2007; Baethge, Achtenhagen, Arends, et. al., 2006) durchgeführt haben. Die acht beteiligten Länder: Dänemark, Deutschland, Finnland, Norwegen, Österreich, Schweden, Schweiz und Slowenien. In diesem Beitrag geht es uns mehr um die Vorstellung eines Modells zur vergleichenden Analyse von Berufsprofilen und ihnen entsprechenden Ausbildungsgängen in mehreren Ländern als um die Präsentation von Ergebnissen. Diese werden im Spätherbst in der Schriftenreihe des BMBF unter dem Titel „Feasibility Study VET-LSA: A comparative analysis of occupational profiles and VET-programmes in 8 European Countries" erscheinen. Ob eine deutsche Version folgen wird, ist zurzeit noch unklar. Danken möchten wir an dieser Stelle unseren deutschen Kolleginnen und Kollegen, die (zumeist) als internationale Koordinatoren einen wesentlichen eigenständigen Beitrag zur Feasibility Study geleistet haben: F. Achtenhagen, T. Bals, A. Barke, K., Breuer, B. Geißel, S. Hillen, M. Müller, R. Nickolaus, A. Schelten, E. Winther, E. Wittemann

2 Wir verwenden im Weiteren die englische Abkürzung VET-LSA (Vocational Education and Training – Large-Scale-Assessment), da sie das Label für die internationale Studie, die auf politischer Ebene vom BMBF koordiniert wird, abgibt.

– Der erste Sachverhalt, der spezifische Wissens- und Kompetenztypus, lässt sich in Abgrenzung zu anderen Large-Scale-Assessments einsichtig machen. Wir beziehen uns dabei auf berufliche Bildung im Sekundarbereich II unterhalb der Hochschulebene. Berufsbildung ist immer handlungsbezogen und ausgerichtet auf spezifische berufliche Handlungsfelder (vgl. Baethge, Achtenhagen, Arends 2008). Demgegenüber sind andere große Large-Scale-Assessments wie PISA[3] und PIAAC[4] auf die Messung und den Vergleich von generellen Kompetenzen von Schülern in allgemeinbildenden Schulen oder von Erwachsenen im Alltagsleben konzentriert: Die im Rahmen von PISA entwickelten Kompetenzmessmethoden beziehen sich auf Lesefähigkeit, Mathematik, Naturwissenschaft und Problemlöseverhalten. PIAAC zielt auf die Messung von Kernkompetenzen Erwachsener, die in wissensbasierten Gesellschaften erforderlich sind: Lesefähigkeit, Rechenfähigkeit und Problemlöseverhalten in informationsintensiven Umwelten (z.B. Schleicher 2008, OECD 2008). Weder PISA noch PIAAC stehen vor der Aufgabe, die Kompetenzen von Jugendlichen und jungen Erwachsenen in spezifischen beruflichen Kontexten zu messen[5].
  Die Kontextgebundenheit von Handlungskompetenzen (nicht diese selbst) als Spezifikum von Berufsbildung stellt die entscheidende Herausforderung für ein LSA in der beruflichen Bildung dar, und zwar in doppelter Perspektive: sie erfordert einerseits die Konzentration des inhaltlichen Messkonzepts auf domänenspezifische Kompetenzen und versperrt sich damit andererseits einer übergreifenden Repräsentativität bei der Samplebildung, die sowohl bei PISA als auch bei PIAAC gegeben ist. Domänenspezifische handlungsorientierte Kompetenzmessung mit Hilfe psychometrischer Methoden hat es bisher bei keinem der Large-Scale-Assessments gegeben. Die Herausforderung für die Methodenentwicklung ist beträchtlich, da es sich um relativ komplexe Aufgabenformate handelt. Diese Herausforderung war nicht Gegenstand der internationalen Feasibility Study. Ihr haben sich die beiden Test-Studien gestellt, die parallel auf nationaler Ebene in Deutschland für den Industriekaufmann/-frau (Achtenhagen, Winther, im Druck) und den KFZ Mechatroniker/-in (Nickolaus, Gschwendtner & Abele, im Druck) durchgeführt worden sind und über die in diesem Band gesondert berichtet wird (vgl. Nickolaus in diesem Band). Das Problem der Samplebildung stand am Anfang der Feasibility Study. Ihre zentrale Frage, ob sich in den Berufsprofilen und Ausbildungsgängen genügend Schnittmengen in den beteiligten Ländern finden lassen, ist nur für jeweils einzelne Berufe zu beantworten.
– Die Beantwortung dieser letzten Frage verweist auf den zweiten zentralen Aspekt, die mehr oder weniger enge Einbindung der Berufsbildung in nationale Arbeitsmarkt- und Beschäftigungsstrukturen sowie Ausbildungstraditionen. Selbst wenn man sich zwischen den an einem VET-LSA teilnehmenden Ländern auf bestimmte Berufe bzw. Berufs- und Ausbildungsfelder verständigt, bedeutet dies noch nicht, dass die ausgewählten Berufe und die ihnen entsprechenden Ausbildungsgänge einen identischen Zuschnitt haben. Wo beispielsweise eine Tradition stark tayloristisch geprägter Arbeitsteilung herrscht, werden wahrscheinlich berufliche Tätigkeitsprofile und Ausbildungsgänge weniger komplex und ausdifferenziert

3 Programm for International Student Assessment.
4 Programme for the International Assessment for Adult Competencies.
5 Auf internationale Berufswettbewerbe (World Skills) wird hier nicht näher eingegangen, da es sich nicht um Large-Scale Assessments handelt.

sein als in Ländern, in denen vielleicht gerade vor dem historischen Hintergrund ausdifferenzierter Berufsbildungssysteme tayloristische Arbeitsteilung weniger tiefgreifend gewirkt hat (vgl. Lutz 1976 und 1984).[6] Hinzu kommt, dass sowohl die ökonomischen Entwicklungspfade zwischen den Ländern nach Maßgabe von Spezialisierungsmustern (ob mit hohen Anteilen industrieller Produktion oder stärkerer Dienstleistungsausrichtung) differieren als auch die Dynamik in der Gestaltung der Beschäftigungsverhältnisse von Land zu Land verschieden sein kann (vgl. Berger, Dore 1996).

Es ist die Aufgabe der Feasibility Study zu klären, wie groß die Schnittmengen in den Berufsprofilen und den Berufsbildungscurricula zwischen den beteiligten Ländern sind, ob sie ein international vergleichendes LSA ermöglichen und was bei der Analyse eines VET-LSA an nationalen Besonderheiten zu beachten ist. Die bisher verfügbaren Strukturvergleiche zwischen den europäischen Ausbildungssystemen (vgl. Descy, Nestler, Tassaring 2005; Descy, Tessaring 2002) können dies nicht leisten, da sie den Charakter deskriptiv nebeneinander gestellter Länderberichte haben.

Die im Folgenden dargestellte Feasibility Study, die die Vergleichbarkeit der Tätigkeitsprofile und Ausbildungsgänge von ausgewählten Berufen bzw. Berufsfeldern prüft, ist keine nur innerszientifische Angelegenheit. Sie steht unter hoher politischer Beobachtung und soll die politische Entscheidungsfindung in den acht beteiligten Ländern unterstützen. Die Stärke der Einbindung der Berufsbildung in die ökonomischen Entwicklungspfade variiert zwar zwischen den Ländern. Überall aber bewirkt sie eine hohe politische Aufmerksamkeit für die Ergebnisse der Feasibility Study. Die politische Sensibilität ist in Ländern, in denen – wie in Deutschland – der Zusammenhang zwischen Ausbildungssystem und wirtschaftlichen Erfolg als sehr eng angesehen wird, größer als in Ländern, wo dieses weniger stark der Fall ist. In jedem Fall aber wird der Feasibility Study auch eine nicht unerhebliche Funktion bei der politischen Legitimierung eines VET-LSA zugewiesen, was die methodischen Anforderungen an ihre Durchführung in einem noch schärferen Licht erscheinen lässt.

2. Konzept und Durchführung der Feasibility Study als Ländervergleich

1.1 Basis: gemeinsames Verständnis eines VET-LSA

Ziel der Feasibility Study ist die Prüfung, ob eine nach Berufsprofilen und Ausbildungsgängen gemeinsame Basis für ein Large-Scale-Assessment in den acht Ländern existiert, d. h. ob in einem späteren VET-LSA bei den einbezogenen Berufsbildungsgängen tatsächlich vergleichbares verglichen wird. Um die Prüfung

6 B. Lutz hat schon früh den Zusammenhang von nationalen Ausbildungstraditionen und gesellschaftlichen Arbeitsstrukturen im Vergleich zwischen Deutschland und Frankreich herausgearbeitet. Er weist nach, dass die Hierarchisierung der betrieblichen Belegschaftsstrukturen in Deutschland weniger stark ausgeprägt ist als in Frankreich und die Tätigkeitsprofile in dem für Deutschland typischen mittleren Fachkräftebereich komplexer als in Frankreich sind – aufgrund im Wesentlichen der Tradition dualer Ausbildung, die ein großes Potential an Fachkräften schafft und die in Frankreich fehlt (Lutz 1976).

konzeptualisieren zu können, bedurfte es der Verständigung über zwei zentrale Punkte[7]:

– Entsprechend der Kontextgebundenheit beruflicher Kompetenzen waren gemeinsame Berufe/Berufsfelder festzulegen, in denen Kompetenzmessungen in einem VET-LSA stattfinden sollten. Die Experten aus den Ländern einigten sich in den ersten internationalen Workshops auf folgende vier Berufe/Berufsfelder:
  - carmechatronics (in Deutschland KFZ-Mechatroniker/-in)
  - electricians industry / craft (in Deutschland Elektroniker/in für Energie- und Gebäudetechnik im Handwerk und Elektroniker/-in für Automatisierungstechnik in der Industrie)
  - business and administration (in Deutschland am ehesten Industriekaufmann/-frau)
  - social and health care (in Deutschland am ehesten repräsentiert durch einen Beruf wie Altenpfleger/in).

  Der Verständigung lag die Idee zugrunde, möglichst aus jedem der drei in allen Ländern großen Berufs- und Beschäftigungsfelder – gewerblich-technische, kaufmännische, sozial-pflegerische Berufs – einen Beruf repräsentiert zu haben (vgl. Baethge, Achtenhagen, Arends et. al., 2006). Es ist unschwer erkenntlich, dass die vier ausgewählten Berufe/Berufsfelder einen unterschiedlichen Umfang haben: die beiden (bzw. drei) gewerblich-technischen Berufe bilden jeweils ein in sich geschlossenes Berufsbild ab, während „business and administration“ und „social and health care“ relativ breite Berufsfelder repräsentieren[8]. Die Unterschiede im Umfang der Berufs-(feld)definition spiegelt die Ausbildungsrealität in den verschiedenen Ländern. Die Berufsfeldbezeichnungen „business and administration“ und social and health care“ sind Ergebnis intensiver Diskussionen in den berufsfeldspezifischen Expertengruppen.[9]
– Der zweite zentrale Punkt betrifft das Konzept für ein VET-LSA. Es wurde ein schneller Konsens erzielt, dass man nicht ein eng arbeitsplatzbezogenes, sondern ein weites Konzept von Berufsbildung, das berufliche Handlungsfähigkeit im Betrieb wie auf dem Arbeitsmarkt anstrebt und gesellschaftliche Teilhabe als

7 Die Einigung fand auf der Grundlage von Vorarbeiten für ein VET-LSA statt, die von Wissenschaftlern aus mehreren europäischen und nicht-europäischen Ländern gemeinsam erarbeitet wurden. Die Vorarbeiten umfassen mehrere internationale Workshops als Grundlage für eine Feasibility Study (Baethge, Achtenhagen, Arends, et. al., 2006) sowie eine anschließende Pre-feasibility Study mit den an der Feasibility Study beteiligten 8 Ländern (Baethge, Arends, Winther, 2008).

8 Dieser Unterschied wurde auch in den internationalen Expertengruppen bei der Identifizierung gemeinsamer ISCO Kategorien deutlich: Während sich die Expertengruppen bei den gewerblich-technischen Berufen relativ eindeutig auf einen Viersteller verständigen konnten, konnte bei „business and administration“ und social and health care“ nur ein gemeinsamer Zweisteller festgelegt werden.

9 Für „business and administration“ war der Vorschlag, Ausbildungsberufe, die für eine kaufmännische Tätigkeit in der Industrie vorbereiten, einzubeziehen (in Deutschland wäre dies am ehesten für den Industriekaufmann/-frau der Fall). Für „social and health care“ sollten Berufe zwischen Sozial- und Gesundheitspflege im Fokus stehen (in Deutschland trifft dies am ehesten für den Altenpfleger/-in zu). Für „nursing“ (in Deutschland Krankenpfleger/in) trat das Problem auf, dass die Berufsausbildung in mehreren Ländern akademisiert und daher für einen Vergleich von Berufsbildung unterhalb der Hochschulebene insbesondere von den Skandinavischen Ländern ausgeschlossen wurde. Für beide großen Bereiche stellte sich heraus, dass die Ausbildungsgänge in einzelnen Ländern, anders als in Deutschland, auf jeweils ein breites Spektrum an spezifischen Berufstätigkeiten hin ausgelegt sind.

wichtige Dimension auch beruflicher Bildung versteht, einem VET-LSA zugrunde legen will. Dementsprechend sollte sich die Kompetenzmessung auf zwei Dimensionen konzentrieren:

- *Berufs- und domänenspezifische Kompetenzen:* sie beziehen sich auf die Fähigkeit, Wissen, Fertigkeiten und Erfahrungen erfolgreich auf authentische Arbeitsaufgaben in spezifisch definierten Beschäftigungssituationen anzuwenden und konkrete Arbeitsaufgaben zu bewältigen.
- *Berufsübergreifende Kompetenzen:* sie sind ebenso integraler Bestandteil jeder Berufsausbildung und sollen den einzelnen in die Lage versetzen, sich erfolgreich auf dem Arbeitsmarkt und in Arbeitsorganisationen bewegen und seine eigene Berufsbiographie planen zu können. Sie werden häufig unter dem Begriff „Schlüsselqualifikationen" subsumiert; in ihnen werden Wissen über Wirkungen von Organisationsstrukturen und Arbeitsmärkten, die Interaktionsfähigkeit in sozialen Gruppen sowie selbständiges Agieren in Arbeitssituationen und Selbstwirksamkeitsüberzeugungen zusammengefasst.[10]
- Ergänzend sollen gleichsam als Kontrollvariablen auch allgemeine Kompetenzen wie Lesefähigkeit, Mathematik mit erhoben werden, um ihre Bedeutung für berufsspezifische und –übergreifende Kompetenzen beurteilen zu können.

Institutionelle und individuelle Bedingungen von Bildungsprozessen sind von entscheidender Bedeutung für die individuelle Kompetenzentwicklung junger Erwachsener. Daher sollten zudem die wichtigsten Kontextfaktoren in das Forschungsdesign eines VET-LSA miteinbezogen werden: Institutionelle Bedingungen auf der Makro- und Meso-Ebene; organisatorische Bedingungen auf der Mikroebene sowie individuelle Hintergrundmerkmale (vgl. Baethge, Achtenhagen, Arends et. al., 2006).

Bezogen auf das Format eines VET-LSA bestand zwar Konsens darüber, dass der Königsweg eine Längsschnittstudie mit drei Messzeitpunkten (am Beginn, am Ende und drei Jahre nach Ende der Ausbildung) abgäbe. Aus forschungspragmatischen Gründen – hoher Aufwand, langer Zeitraum, Samplegröße bei erwartbarer Panelmortalität (vgl. Baethge, Achtenhagen, Arends et. al., 2006, S. 127 ff.) aber verständigte man sich auf den Kompromiss, als gemeinsamen Messzeitpunkt für alle Beteiligten das Ende der Ausbildung festzulegen und darüber hinaus den Ländern Optionen für weitere Messzeitpunkte in einem nach Kohorten differenzierten Design einzuräumen (z. B. zu Anfang der Ausbildung und nach deren Ende).

2.2 Methodischer Ansatz und Implementation der Feasibility Study

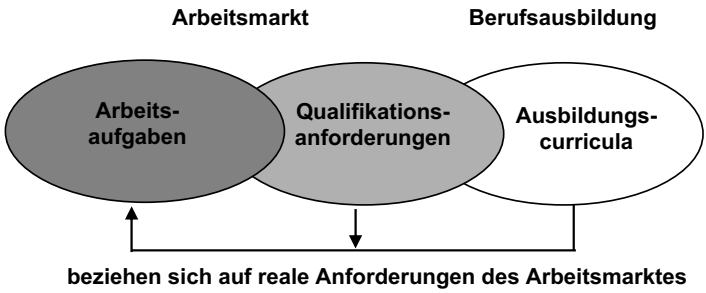
Die zentrale Frage der Feasibility Study nach der für ein VET-LSA erforderlichen gemeinsamen Basis in Berufsprofilen und Ausbildungsgängen versuchten wir

10 Die Trennschärfe der beiden Kompetenzdimensionen variiert zu einem gewissen Grade und für bestimmte Merkmalsausprägungen zwischen Berufen. So ist beispielsweise Kommunikationsfähigkeit in vielen kaufmännischen und sozialpflegerischen Berufen integraler Bestandteil der domänenspezifischen Kompetenzen (vgl. Abschnitt 4.3).Für Fremdsprachen gilt ähnliches. Dennoch sind die beiden großen beruflichen Kompetenzdimensionen analytisch und bei Messinstrumenten zu trennen, – zum einen, um Vergleiche zwischen Berufen für berufsübergreifende Kompetenzen durchführen zu können, zum anderen weil sich berufsübergreifende Kompetenzen nicht in Kommunikations- und Interaktionsfähigkeit erschöpfen.

mit Hilfe eines komplex angelegten Expertenratings zu beantworten, wie es bei vergleichbaren Large-Scale-Assessments üblich ist (z.B. PISA, PIAAC). Den Ausgangspunkt der ratings bildet der dissoziierte Zusammenhang zwischen Berufs- und Tätigkeitsprofilen auf dem Arbeitsmarkt und beruflichen Ausbildungscurricula. Der Begriff „dissoziierter Zusammenhang" bezeichnet den Sachverhalt, dass Ausbildungscurricula zwar auf Arbeitsmarktanforderungen hin entwickelt werden, aber nicht in ihnen aufgehen, vor allen Dingen nicht in der Durchführung der Ausbildung. Ausbildungsordnungen und Tätigkeitsprofile von Berufen sind nicht deckungsgleich, weil beide unterschiedlichen Logiken und Veränderungsdynamiken folgen: In der Arbeitswelt werden Berufs- und Tätigkeitsprofile durch Arbeitsaufgaben und Qualifikationsforderungen von Beschäftigungsorganisationen (Unternehmen) für ihre Mitarbeiter entsprechend ihren ökonomischen Bedürfnissen definiert und unterliegen deren Dynamik. Ausbildungscurricula werden demgegenüber mit Blick auf die individuelle Kompetenzentwicklung im Rahmen betriebsübergreifender nationaler Berufsordnungen entwickelt. Diese Differenz in den Definitionskriterien von beruflichen Tätigkeitsprofilen und Ausbildungsgängen führt immer wieder zu mismatches zwischen Arbeitsmarktanforderungen und Ausbildungsergebnissen. Diese mismatches werden umso größer, je rapider die Dynamik des ökonomischen und technologischen Wandels in der Arbeitswelt ist. Da sich in der letzten Dekade die mismatches in den meisten europäischen Ländern verstärkt zu haben scheinen, hat die Verbesserung der Abstimmung zwischen Arbeitsmarkt und Berufsbildung eine hohe Priorität in der europäischen Berufsbildungspolitik bekommen (vgl. z.B. Cedefop 2009, European Union, 2008).

Weil wir nicht wissen, wie groß die Diskrepanz zwischen beruflichen Tätigkeitsprofilen und den entsprechenden Ausbildungsgängen in den Teilnehmerländern ist, haben wir die Expertenratings in zwei voneinander unabhängig konzipierten Beurteilungsverfahren durchgeführt: eines, das berufliche Arbeitsaufgaben und Qualifikationsanforderungen auf dem Arbeitsmarkt für die vier einbezogenen Berufsfelder thematisiert, ein anderes, das für dieselben Berufsfelder die Outcomes der Berufsausbildung zur Diskussion stellt.

Abb. 1: Verbindung zwischen Ausbildungscurriculum und beruflichem Tätigkeitsprofil auf dem Arbeitmarkt

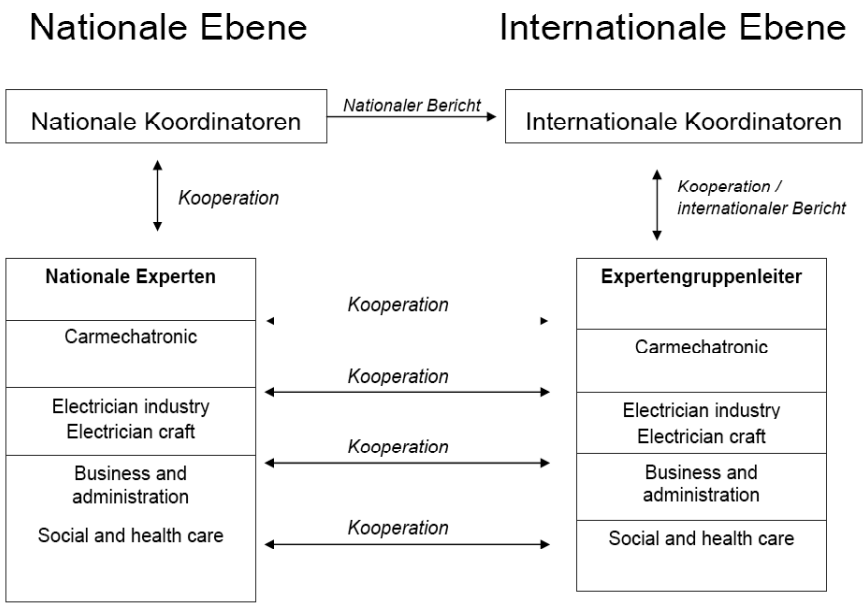


Unser für Lage-Scale Assessments übliches Vorgehen lässt sich am ehesten als qualitatives Verfahren mit quantitativen Elementen beschreiben, das in einem

Wechselspiel zwischen internationaler und nationaler Ebene durchgeführt wurde: Auf der internationalen Ebene wurden in vier Workshops von den internationalen Koordinatoren und Expertengruppenleitern sowie den nationalen Koordinatoren und den Experten für jeden Beruf/Berufsfeld (vgl. Abbildung 2) die Vergleichsstrategie erörtert und festgelegt, d. h. es wurden die Kriterien, Fragestellungen und Items für die Expertenratings in den Ländern entwickelt und später die Ergebnisse der nationalen ratings erörtert, validiert und in Empfehlungen zur Durchführung eines VET-LSA zusammengefasst. In jedem Land wurden für jedes Berufsfeld zwei Expertenworkshops durchgeführt, auf denen die auf der internationalen Ebene gemeinsam erarbeiteten Vergleichsinstrumente aus der je nationalen Perspektive erörtert, beurteilt und gegebenenfalls durch national spezifische Items ergänzt wurden.

Mit Hilfe mehrfacher Rückkopplungsschleifen zwischen nationaler und internationaler Expertenebene gelang es zum einen, die nationalen Besonderheiten in den beruflichen Tätigkeitsprofilen und Ausbildungsgängen angemessen einzubeziehen und zugleich eine intensive Diskussion über die Bedingungen eines VET-LSA in jedem Land anzustoßen; zum anderen konnten damit die ratings in der internationalen Expertendiskussion noch einmal kritisch validiert werden. Insofern begreifen wir auch die quantitativen Skalierungen im Wesentlichen als Ergebnis vorgängiger qualitativer Diskussionen und in ihnen erfolgter Festlegungen auf analytische Aspekte. Dieses Verfahren rechtfertigt auch die Größe des Experten-Sample in jedem beteiligten Land.

Abb. 2: Organisationsstruktur der Feasibility Study



2.3 Zur Struktur des Experten-Samples

Dem doppelten Bezug auf Ausbildung und Arbeitsmarkt entsprechend war auch das Sample der Experten in den Ländern für das rating zusammenzustellen: einbezogen wurden Ausbildungsexperten aus Unternehmen und Lehrer in beruflichen Schulen als zwei Hauptgruppen sowie als dritte Gruppe Vertreter aus dem politisch-administrativen und dem Forschungsbereich (vgl. Tabelle 1).

Auf der internationalen Ebene einigten wir uns auf relativ kleine Samples für jede Berufsgruppe[11]. In den nationalen Workshops sollten mindestens sieben Experten in jedem Berufsfeld die ratings vornehmen, im zweiten Workshop möglichst einige mehr. Bei dem ersten Workshop sollten vor allem Experten aus Betrieben und beruflichen Schulen, im zweiten darüber hinaus Repräsentanten des politisch-administrativen Umfelds einbezogen werden. Alle Experten sollten eng mit dem jeweiligen Berufsfeld verbunden sein, um die für kleine Sample erforderliche Homogenität sicherzustellen. Auf diesem Weg gelang es, relativ homogene Expertensample in allen beteiligten Ländern aufzubauen (vgl. Tabelle 1).

Tab. 1: Sample der Experten nach beruflicher Funktion und Berufsfeldern

1. nationaler Workshop	**Lehrer in Berufsschulen**	**Experten im Betrieb**	**Vertreter von beruflichen Aus- und Weiterbildungsinstituten, Prüfungs- und Bildungsausschüssen, Sozialpartner, Administration und Forschung**	Gesamt
Carmechatronic	39	21	23	83
Electrician craft	22	15	21	58
Electrician industry	19	15	24	58
Business and administration	20	18	28	66
Social and health care	34	26	17	77
Gesamt	134	95	113	342

2. nationaler Workshop	**Lehrer in Berufsschulen**	**Experten im Betrieb**	**Vertreter von beruflichen Aus- und Weiterbildungsinstituten, Prüfungs- und Bildungsausschüssen, Sozialpartner, Administration und Forschung**	Gesamt
Carmechatronic	24	20	32	76
Electrician craft	22	16	27	65
Electrician industry	18	9	27	54
Business and administration	22	14	34	70
Social and health care	20	8	56	84
Gesamt	106	67	176	349

Für die Feasibility Study reicht eine begrenzte Zahl von Experten für jedes Berufsfeld aus, Repräsentativität ist keine notwendige Bedingung[12]. Gleichwohl muss diese Prämisse empirisch verifiziert und an den Ergebnissen ausgewiesen werden: Im

11 Expertenratings mit kleinen Samples sind insbesondere in der klinischen Psychologie ein gängiges Verfahren (z.B., Craigie, Loader, Burrows, et.al., 2002; Cendales, Kirk, Moresi, et.al., 2006; Braithwaite, Luft, Bender, et.al., 2007).

12 Für die Sampleauswahl mussten z.B. Aspekte wie Betriebsgröße, Branchenzugehörigkeit, Größe der Ausbildungseinrichtung/Schule u. a. nicht in Betracht gezogen werden, weil es um Berufs- und Ausbildungsprofile insgesamt, nicht um deren betriebs- oder branchenspezifische Ausprägung geht.

Vorgriff auf die wichtigsten Ergebnisse kann konstatiert werden, dass die methodologische Prämisse für kleine Expertensamples bestätigt werden konnte: in allen vier Berufsgruppen bestand ein hoher Konsens zwischen den Experten – mit einer Ausnahme.[13]

Tabelle 1 zeigt, dass im Durchschnitt die vereinbarten Kriterien für die Samplebildung in allen Ländern realisiert werden konnten. Auch wenn die Zusammensetzung der Expertensamples nach den drei Hauptgruppen aufgrund nationaler Besonderheiten in der Ausbildungsorganisation (z. B. eher schul- als betriebsbasierte Ausbildung) zwischen einzelnen Ländern etwas variiert, kann insgesamt von einem homogenen Sample ausgegangen werden. Am ersten nationalen Expertenworkshop nahmen insgesamt 342 Experten (im Durchschnitt etwa 9 pro Workshop in jedem Land) am zweiten 349 Experten teil. Beim zweiten Workshop konnte die intendierte höhere Beteiligung der nicht direkt am Ausbildungsprozess beteiligten Experten erreicht werden (176 gegenüber 113 beim ersten Workshop). Dies war wichtig, um noch eine zweite qualitative Kontrolle der ratings vom ersten Workshop zu haben.

2.4 Die Erhebungsinstrumente

Man kann einen internationalen Vergleich von Berufsprofilen und Ausbildungsgängen nicht einfach auf viele einzelne Expertenmeinungen gründen. Man muss diese systematisieren und bis zu einem gewissen Grad standardisieren, um sie gewichten zu können. Den Prozess der Systematisierung und Standardisierung der Rating-Aspekte vollzogen wir in Form von internationalen Expertendiskussionen zum einen für die Arbeitsaufgaben und Qualifikationsanforderungen zu in der Forschung verfügbaren Item-Sets, die auf die Bedingungen der Berufsfelder in den acht beteiligten Ländern angepasst wurden; zum anderen durch das Sammeln, Erörtern und Komprimieren von Prüfungsaufgaben der Abschlussexamina in den Ländern, die zu einem einheitlich in allen Ländern zur Bewertung angewandten Satz von Prüfungsaufgaben für den komparativen Vergleich von Lernergebnissen (learning outcomes) am Ende der Ausbildung verdichtet wurden.

Berufliche Arbeitsaufgaben und Qualifikationsanforderungen auf dem Arbeitsmarkt

Für die komparative Analyse von Berufsprofilen werden international häufig Deskriptoren aus Berufsklassifikationssystemen oder Entgeltrahmenabkommen[14] eingesetzt die sich in der Praxis erfolgreich bewährt haben. Beispielsweise orientiert sich der in PIAAC eingesetzte „Job Requirement Approach (JRA)“ an dem UK Skills Survey[15] und dem amerikanischen Berufsklassifikationssystem O*NET[16] (Green, 2009). Die einzige bisher international anerkannte Systematik zu Klassifizierung von Berufen ist jedoch die „International Standard Classification of Occupations (ISCO)“. Die internationale Standardklassifikation der Berufe ist ein Klassifikationsschema für Gruppen von Berufen; eng verwandte Berufe werden zu einer Berufsgruppe zusam-

13 Sie betrifft den Bereich „business and administration“„bei den Prüfungsaufgaben am Ende der Ausbildung zum Vergleich von Lernergebnissen in der Ausbildung.
14 z.B. ERA Entgeltabkommen, IG Metall Bezirk Baden Würtemberg (2004).
15 Felstead, Gallie, Green and Zhou (2007).
16 Occupational Information Network: http://online.onetcenter.org/

mengefasst, nach den jeweiligen Aufgaben und Pflichten in einer Tätigkeit gruppiert und in eine vierstufige, mit jeweils ein- bis vierstelligen Ziffernfolgen gekennzeichnete Hierarchie eingeordnet (International Labour Organisation, 2009). Für die Feasibility Study wurde ISCO als Ausgangspunkt für die Einigung auf gemeinsame Berufsgruppen in jedem Beruf/Berufsfeld genutzt: Die internationalen Expertengruppen für die gewerblich-technischen Berufe einigten sich auf einen gemeinsamen Viersteller, die Expertengruppen für die Dienstleistungsberufe auf einen gemeinsamen Zweisteller[17]. Für einen Vergleich von konkreten Arbeitsanforderungen und Qualifikationsanforderungen ist ISCO jedoch nicht genug elaboriert.

Neben ISCO werden derzeit weitere Instrumente zum Vergleich von beruflichen Anforderungsprofilen entwickelt. Das Dictionary of Skills and Competences[18] bietet beispielsweise auf einer Onlineplattform eine umfangreiche Sammlung von Begriffen für Kompetenzen und Fertigkeiten, wie sie in Lebensläufen, Stellenanzeigen, Stellenprofilen verwendet werden. Eine fachliche Spezifizierung bei der Beschreibung von Qualifikationen wird beispielsweise mit dem VQTS-Modell verfolgt (Luomi-Messerer, Markowitsch 2006; Markowitsch, Becker, Spöttl 2006). Mit Hilfe einer „Kompetenzmatrix" sollen „Stufen der Kompetenzentwicklung oder die von einer Person zu einem bestimmten Zeitpunkt bereits erworbenen Stufen als Kompetenzprofile abgebildet werden" (Markowitsch, Luomi-Messerer 2008, S. 59). Vorschläge für Beschreibungen von arbeitsbezogenen Kompetenzen in einer Kompetenzmatrix existieren für die Bereiche Mechatronik und Elektronik/Elektrotechnik[19]. Eine internationale Validierung steht bisher noch aus.

Für die Beantwortung der Frage, ob in den vier Berufen/Berufesfeldern in den 8 beteiligten Ländern eine ausreichend große Schnittmenge für ein VET-LSA vorhanden ist, reichen die genannten Ansätze nicht aus. Für die komparative Analyse von Arbeitsaufgaben und Qualifikationsanforderungen in der Feasibility Study wurde neben ISCO ein elaborierter unabhängiger Referenzrahmen aus den USA herangezogen. Das „Occupational Information Network (O*NET)" bietet gegenwärtig das weltweit ausführlichste und wissenschaftlich am besten fundierte Deskriptoren-System zur Beschreibung von Berufen. O*NET wurde auch bei PIAAC für die Entwicklung des sogenannten „Job Requirement Approach" herangezogen (z.B. Green 2009). Da O*NET für den amerikanischen Arbeitsmarkt entwickelt ist, wurden die Deskriptoren in den internationalen Expertengruppen dem europäischen Arbeitsmarkt angepasst. Für den Zweck der Feasibility Study ist O*NET zudem besonders geeignet, weil es neben „skills" auch „knowledge" und „education" als Arbeitsanforderungen erfasst, d.h. für jeden Beruf/Berufsfeld liegen Beschreibungen für Arbeitsaufgaben und Qualifikationsanforderungen in einem unabhängigen System vor, die nach Niveaus differenziert sind[20].

Für die Feasibility Study wurde O*NET zunächst zur Einigung auf gemeinsame Berufs- und Tätigkeitsprofile („occupations") neben ISCO eingesetzt. Ausgehend von den gewählten „occupations" wurden Deskriptoren für Arbeitsaufgaben und

17 Es wurde die Version ISCO-08 in der Entwurfsphase zugrunde gelegt.

18 Dictionary of skills and competences: http://www.disco-tools.eu/index.php

19 Die Kompetenzmatrix für Mechatronik wurde in einem Workshop mit deutschen und Österreichischen Experten entwickelt, die Kompetenzmatrix für Elektronik/Elektrotechnik wurde von Experten aus den Niederlanden entwickelt.

20 Die O*NET Deskriptoren werden in regelmäßigen Abständen aktualisiert. Die letzte Aktualisierung der Datenbank wurde im Juni 2009 durchgeführt.

Qualifikationsanforderungen für Berufe die mittlere bis erhebliche Praxiserfahrung benötigen („job zone“ 3 und 4) und zwischen ein bis vier Jahre berufliche Ausbildung voraussetzen („specific vocational preparation (SVP) range“ 6 und 7)[21] ausgewählt und an die Bedingungen des Europäischen Arbeitsmarktes angepasst (z. B. wurden für alle domänenspezifischen Qualifikationsanforderungen. Beispiele zur Konkretisierung ergänzt). Die internationalen Expertengruppen einigten sich jeweils auf ein gemeinsames Set von Arbeitsaufgaben und Qualifikationsanforderungen für den jeweiligen Beruf/Berufsfeld als Ausgangbasis für die Bewertung in nationalen Workshops (ca. 11 bis 19 Items pro Set). Um nationale Besonderheiten zu berücksichtigen, war die Möglichkeit für Ergänzungen von Arbeitsaufgaben oder Qualifikationsanforderungen vorgesehen. Die Skalierung wurde von O*NET übernommen und geringfügig adaptiert[22]. Die Erhebungsinstrumente wurden für die nationalen Workshops in fast allen Ländern in die Landessprache übersetzt.

Die O*NET Deskriptoren wurden häufig als zu abstrakt kritisiert. Dieser Kritik können drei Argumenten entgegengehalten werden:

- Aus theoretischer Perspektive muss berücksichtigt werden, dass die Formulierungen der O*NET Deskriptoren aktuelle Beschreibungen für Arbeitsaufgaben und Qualifikationsanforderungen von Jobs auf dem amerikanischen Arbeitsmarkt bereitstellen und damit gewährleistet ist, dass globale Trends in der Arbeitswelt widergespiegelt werden. Die Zunahme von wissensintensiven Tätigkeiten die die kognitiven Anforderungen an Auszubildende/Arbeitskräfte erhöhen, ist bspw. ein globales Phänomen (z. B. Bell 1972, Polanyi 1966, Nonaka, Takeuchi 1997, Child, Heavens 2001).
- Zweitens, um die theoretische Sicht nicht zu sehr in den Fokus zu stellen, wurden einzelne Deskriptoren an die Bedingungen des Europäischen Arbeitsmarktes angepasst.
- Drittens, um nationale Besonderheiten zu berücksichtigen, wurden besonders wichtige Arbeitsaufgaben und Qualifikationsanforderungen der Länder, die nicht in O*NET enthalten sind, mit aufgenommen.

Prüfungsaufgaben in zentralen Inhaltsbereichen am Ende der Ausbildung

Über Prüfungsaufgaben am Ende der Ausbildung sollten die Schnittmengen in den Ausbildungsgängen, die zu qualifizierten Ausbildungsabschlüssen in den vier Berufsfeldern führen, geklärt werden. Dieses Verfahren schien mehr Treffsicherheit zu versprechen als eine vergleichende Analyse von Lehrplänen und Ausbildungsordnungen, die in anderen Large Scale Assessments häufig herangezogen werden (z. B. Curriculumanalyse in TIMSS, vgl. Schmidt, Raizen, Britton, Bianchi, Wollte 1997). In der Berufsausbildung sind diese aber in den einzelnen Ländern unterschiedlich ausführlich gestaltet, und man weiß von ihnen auch nicht genau, wie viel

21 Die “job zone” gibt an, wie Personen in den Job einsteigen, d.h., wie viel Berufserfahrung, Ausbildung und Einarbeitung am Arbeitsplatz notwenig ist, um die Arbeitsaufgaben erfolgreich zu bewältigen. Die „specific vocational preparation (SVP) range” beschreibt hingegen, wie viel Zeit benötigt wird, um sich die benötigen Fähigkeiten und Fertigkeiten in der Ausbildung anzueignen.

22 Die Arbeitsaufgaben wurden anhand einer 5-Punkte Skala hinsichtlich “relevance” (1=not relevant; 5=extremely relevant) und “frequency” (1=several times a year; 5=several times daily), die Qualifikationsanforderungen hinsichtlich “relevance” bewertet.

der jeweils offiziell definierten Inhalte in welcher Breite und Tiefe in der Ausbildung tatsächlich vermittelt wird. Ein Versuch im Rahmen einer Pre-Feasibility Study (vgl. Baethge, Arends, Winther 2008) hatte gezeigt, dass wir über die Sammlung und den Vergleich offizieller Dokumente, wie z.B. Rahmenlehrpläne, keine auch nur annähernd ausreichende Informationsbasis zu den Ausbildungsinhalten erhalten. Das Hauptproblem bei einer derartigen Analyse ist, dass der Grad der Abstraktheit der Beschreibungen von „learning outcomes" in den einzelnen Ländern sehr variiert. Sie aber bilden den unverzichtbaren Bezug für die Kompetenzmessung.

Für die Wahl der nationalen Ausbildungsprogramme wurde die International Standard Classification of Education ISCED (UNESCO, 2006) herangezogen mit dem Fokus auf den ISCED-Stufen 3B und 3C[23], die den Zugang zum Arbeitsmarkt sowie zu weiterführenden Bildungsgängen (ISCED 5B oder anderen ISCED 3 und 4 Programmen) eröffnen.[24] Für die Ausbildungsgänge auf ISCED-3-Niveau wurden für alle vier Berufsfelder von den internationalen Expertengruppen zentrale Inhaltsbereiche am Ende der Ausbildung („core areas") definiert als Grundlage für die Auswahl von nationalen Prüfungsaufgaben, die wesentlich konkreter Einblick in die Inhalte und das Niveau der Ausbildung geben als abstrakte Beschreibungen in offiziellen Dokumenten.

Als Kernbereiche für die Ausbildung der Berufsbereiche wurden festgelegt für:

Carmechatronics	Motormanagement und Abgasverhalten, Bremssystem, Fahrwerk, Kraftübertragung, Komfort- und Sicherheitssystem
Electrician craft	Installations- und Steuerungstechnik, Antriebstechnik, Beleuchtungstechnik, Gebäudeautomation, Messtechnik, Fehlersuche
Electrician industry	Steuerungstechnik, Antriebstechnik, Messtechnik, Wartung, Fehlersuche
Business and administration:	Einkauf, Sales & Marketing, Kundenservice, Finanzierung/Rechnungswesen (Grundkenntnisse), Lagerhaltung (Grundkenntnisse, Organisation/Administration)
Social and health care:	Patienten-/Klientenversorgung (Assistenz im Versorgungsprozess), Kommunikation und Beziehungsaufbau, inter- und intradisziplinäre Kooperation, Gesundheit und Sicherheit, Administration und rechtliche Rahmenbedingungen)

Auf die Kernbereiche bezogen wurden dann von den nationalen Experten in jedem Beruf/Berufsfeld Prüfungsaufgaben der jeweiligen Ausbildungsgänge gesammelt und von den Expertengruppenleitern zu einem Satz von 15-20 Aufgaben komprimiert.

23 In Schweden und Finnland, wo der Abschluss der beruflichen Bildung auch den Zugang zur Tertiärbildung eröffnet, sind die beruflichen Ausbildungsgänge in ISCED 3A klassifiziert. Da in der ISCED-Klassifikation der OECD ISCED 3A auf allgemeinbildende Programme festgelegt ist, es hier aber um den beruflichen Teil der Ausbildung geht, haben wir sie in 3A/B eingeordnet.

24 Die mögliche Alternative, den Europäischen Qualifikationsrahmen ((EQF) heranzuziehen, wurde geprüft, aber negiert, weil die Zuordnung von Ausbildungsgängen zu EQF-Stufen in einzelnen Ländern noch nicht abgeschlossen ist.

Die Komprimierung folgte folgenden Kriterien:
– Wird die Aufgabe für das Tätigkeitsprofil und die Qualifikationsanforderungen des Berufs auf dem Arbeitmarkt als wichtig angesehen?
– Repräsentiert die Aufgabe einen der in der internationalen Expertengruppe festgelegten zentralen Inhaltsbereiche?
– Ist sie unabhängig von nationalen Spezifika (z. B. nur national geltenden Rechtsbestimmungen)?
– Ist die Länge der Aufgabe für das rating handhabbar?
– Sind in der Aufgabensammlung alle beteiligten Länder vertreten?

Da der endgültigen Satz von Prüfungsaufgaben von allen Ländern abgestimmt wurde, kann man davon ausgehen, dass er die wesentlichen Ausbildungsinhalte in allen nationalen Programmen repräsentiert.

Der so gemeinsam entwickelte Satz von Prüfungsaufgaben wurde – ohne Angabe, aus welchem Land jede Prüfungsaufgabe stammt – auf der Basis kurzer Zusammenfassungen einheitlich in allen Ländern im zweiten nationalen Workshop von den Berufsexperten nach zwei Kriterien bewertet: 1) Die Aufgabe repräsentiert zentrale Lernergebnisse am Ende der Ausbildung (von „stimme gar nicht zu“ bis „stimme voll zu“) und 2) Beurteilen Sie die Komplexitätsstufe der Prüfungsaufgabe (1 = geringe Komplexität; 4 = hohe Komplexität)

2.5. Qualitative und quantitative Ergebnisanalyse

Qualitative Analysis

Die qualitative Analyse ist in den Prozess der Feasibility Study integriert. Sie besteht aus einer Vielzahl von Experten-Diskussionen auf nationaler und internationaler Ebene, die protokolliert und ausgewertet wurden. Diese Gruppendiskussionen standen am Anfang der Feasibility Study und fanden ihren Niederschlag auf internationaler Ebene in der Entwicklung der Items des „research tool 1“ (Modifizierung und Erweiterung der O’NET Descriptoren für Arbeitsaufgaben und Qualifikationsanforderungen) durch die internationale Expertengruppe jedes Berufsfeldes.

Die Gruppendiskussionen auf der nationalen Ebene, die der Erörterung und der Beurteilung der modifizierten O’NET Items aus der je nationalen Expertenperspektive galten, gingen in die ratings und in Ergänzungen von Items für die Arbeitsaufgaben und Qualifikationsanforderungen ein. Sie wurden in den nationalen Berichten dokumentiert und gingen über sie in den internationalen Bericht ein. Des Weiteren dienten die nationalen Gruppendiskussionen auf dem ersten Workshop der Auswahl von Prüfungsaufgaben der Abschlussprüfungen für den internationalen Abstimmungsprozess.

Auf der internationalen Ebene fanden Diskussionen in den domänenspezifischen Expertengruppen, denen ein Experte pro Beruf/Berufsfeld aus jedem Land angehörte, statt. Die hier eingebrachten nationalen Aspekte und Resultate der nationalen Workshop-Diskussionen wurden in einem neuen Abstimmungs- und Bewertungsprozess zu neuen Instrumenten bzw. – im letzten internationalen Workshop – zu einer gemeinsamen Interpretation der Ergebnisse amalgamiert.

Es sind diese mehrfachen Rückkoppelungsschleifen qualitativer Expertendiskussionen, die den Einsatz quantitativer Instrumente überhaupt erst sinnvoll und möglich machten. Insofern muss man die quantitativen Ergebnisse als Ausfluss vorgängiger qualitativer Aushandlungsprozesse zu Arbeitsaufgaben und Qualifikationsanforderungen für den Arbeitsmarkt sowie Prüfungsaufgaben für Lernergebnisse am Ende der Ausbildung betrachten. Bei einem zufällig ausgewählten Sample von Experten und unvorbereiteten ratings wäre der Übereinstimmungsgrad vermutlich niedriger gewesen. Aber Expertenrating heißt, Sachverstand bestmöglich zu aktivieren und nicht ein beliebiges Meinungsbild zu erstellen.

Quantitative Analyse

Zum Zweck der Vergleichbarkeit von ratings in den Ländern sind quantitative Analysen unumgänglich. Die Probleme, statistische Prüfverfahren auf kleine Probandenzahlen anzuwenden, sind allgemein bekannt. Angesichts der kleinen Samples in jedem Land und des qualitativen Charakters der Studie kamen nur non-parametrische Methoden für die quantitative Datenanalyse in Betracht (vgl. Bortz, Lienert 2008; Pett 1997).

Bei den Schätzdaten wurden für jedes Land die (arithmetischen) Mittelwerte[25] und die durchschnittliche absolute Abweichung für jedes Item berechnet. Analog verfuhren wir beim Vergleich der Länder: Neben dem Ländermittelwert bei jedem Item wurde der Mittelwert der Ländermittelwerte und die mittlere absolute Abweichung der Ländermittelwerte vom Mittelwert aller Ländermittelwerte berechnet, um so den Ländern ein gleiches Gewicht zu verleihen, unabhängig von der Zahl der Experten, die sich jeweils am rating beteiligten und in den Mittelwert des Landes eingingen.[26]

Um den Grad der Übereinstimmung der Ergebnisse testen zu können, wurden die Mittelwerte der Länder in eine ordinale Rangskala transformiert. Die Übereinstimmung der ratings zwischen den einzelnen Ländern wurde mit Kendall's Konkordanz-Koeffizient W getestet (vgl. Howell 2009). Ergänzend zu Kendall's W berichten wir auch die äquivalente „durchschnittliche Spearman's Korrelation". Von beiden Tests wird in der Regel eine hoch signifikante Übereinstimmung der ratings zwischen den Ländern ausgewiesen (vgl. Abschnitt 4, Tabelle 3)

3. Ähnlichkeiten und Differenzen in den nationalen Ausbildungssystemen

Um ein VET-LSA durchführen zu können, bedarf es in allen beteiligten Ländern einer Reihe struktureller Ähnlichkeiten. Da der Sinn eines VET-LSA darin liegt, über Kompetenzmessung institutionelle Stärken und Schwächen von Ausbildungssystemen zu identifizieren, bedarf es ebenfalls eines bestimmten Maßes an Unterschieden in der institutionellen Organisation der Ausbildung.

25 Den Mittelwert haben wir dem Median wegen der Struktur der Daten vorgezogen. Bei konsistent hohen Schätzwerten würde der Median zu gleichen Werten führen und so zu einem Verlust an Informationen.

26 So nahmen beispielsweise beim rating der Prüfungsaufgaben zum Ausbildungsabschluss im Feld „business and administration" in Slowenien 19, in Österreich 10 Experten teil.

– Die strukturellen Ähnlichkeiten betreffen zu vorderst die Inhalte der Ausbildungscurricula, weil es wenig sinnvoll ist, Kompetenzmessung in Inhaltsbereichen vorzunehmen, die in der Ausbildung nicht vorgekommen sind. Sie betreffen weiterhin die durchschnittlichen kognitiven Zugangsvoraussetzungen zur Ausbildung, vorausgesetzte Schulabschlüsse sowie das Gewicht der Berufsausbildung in den Sekundarstufe 2 insgesamt.
– Erwartbare institutionelle Differenzen, die als Kontextvariable in einer Hauptstudie VET-LSA gründlich auszuleuchten sind, da sie neben individuell-biographischen Merkmalen als Erklärungsmomente für unterschiedliche Kompetenzmessergebnisse zwischen Ausbildungspopulationen herangezogen werden sollen, lassen sich auf die Art der Steuerung und der Organisation von Ausbildung, ihre Finanzierung, auf Formen der Qualitätssicherung sowie auf die Verbindungen zu anderen Bildungsgängen und zum Arbeitmarkt beziehen. Zu den genannten Aspekten wurde von den nationalen Koordinatoren der beteiligten Länder im „national report" ein erstes kurzes Kapitel zu strukturellen und institutionellen Merkmalen der Berufsausbildung im jeweiligen Land geschrieben.

Ähnlichkeiten[27]

Eine erste wichtige Gemeinsamkeit aller Länder besteht darin, dass es in keinem Land nur ein einziges Berufsbildungssystem gibt, wenn man auf die Form der Institutionalisierung schaut. In den meisten beteiligten Ländern finden wir duale Ausbildungssysteme und schulische Varianten. Die größte institutionale Einheitlichkeit weisen Schweden (fast ausschließlich Schulorganisation, duale Ausbildung wird derzeit in einer Pilotstudie getestet) und Norwegen auf (vgl. Tabelle 2). Alle für die Feasibility Study ausgewählten Ausbildungsprogramme des Sekundarbereichs (upper secondary education) dauern zwischen drei und vier Jahren (Ausnahme: ein Ausbildungsgang in Slowenien).[28] Der Zugang zur Ausbildung setzt in der Regel den (erfolgreichen) Abschluss der Pflichtschulzeit voraus (mindestens 9 Jahre Schulzeit und Erreichung ISCED Level 2). Die Teilnahmequote an beruflichen Ausbildungsgängen (bezogen auf alle Schüler in der Sekundarbildung) liegt in allen Ländern bei 50 % oder darüber, am niedrigsten in Dänemark (47 %), am höchsten in Österreich (72 %).[29]

Was für die spätere Durchführung eines VET-LSA besonders wichtig erscheint: Alle vier Berufsfelder (carmechatronic, electrician craft/industry, business and administration, social and health care) sind in der Beschäftigtenstruktur der Länder stark repräsentiert und die ihnen entsprechenden Ausbildungsgänge weisen eine hinreichend große Zahl an Auszubildenden auf.[30]

27 Über die Gemeinsamkeiten in den Ausbildungsgängen wird im vierten Abschnitt dieses Artikels berichtet.

28 Wieweit die offiziellen Angaben der Realität aller Ausbildungsgänge in den Ländern entsprechen, konnte nicht geprüft werden, ist für die Hauptuntersuchung aber im Auge zu behalten.

29 Zahlen von 2005 nach CEDEFOP 2007.

30 Ausnahme eventuell electricians (Industry) in Norwegen.

Tab. 2: Strukturen der beruflichen Erstausbildung in den beteiligten Ländern

	Österreich	Dänemark	Finnland	Deutschland	Norwegen	Slovenia	Sweden	Switzerland
Struktur der beruflichen Erstausbildung auf ISCED Stufe 3*	Ausbildung im dualen System (1/3 der Auszubildenden) vollzeitschulische Berufsbildung (1/3 der Auszubildenden) Berufsfachschule (10%)	Ausbildung im dualen System Berufliche Grundbildung in allgemeinbildenden Schulen (5 Möglichkeiten: landwirtschaftlich, kaufmännisch, Sozial und Gesundheit, technisch, kombinierte Colleges)	Berufliche Ausbildungsstätten (kombiniert mit Praxisanteilen am Arbeitsplatz): 82% aller Auszubildenden kompetenzbasierte Berufsbildung (hauptsächlich für Erwachsene mit Berufserfahrung, ca.18%) Alternative: apprenticeship	Ausbildung im dualen System (ca.70% der Ausbildungsanfänger in 2006) vollzeitschulische Berufsbildung, insbesondere für Dienstleistungsberufe (ca. 30%)	9 Ausbildungsprogramme hauptsächliches Modell, 2 Jahre Schule, gefolgt von 2 Jahren Ausbildung	Verkürzte Berufsausbildung (2 Jahre) vollzeitschulische Berufsbildung (3 Jahre) vocational technical education (2 Jahre) technical secondary education (4 years)	13 vollzeitschulische berufliche Ausbildungsprogramme im allgemeinbildenden Schulsystem vollzeitschulische Berufsbildung (ca. 6.000 Auszubildende**)	Ausbildung im dualen System (ca. 84% in 2007) vollzeitschulische Berufsbildung (16% in 2007)
Ausbildungsdauer (durchschnittliche Jahre)	**3.5 - 4**	**3-4**	**3**	**3-3.5**	**4-4.5**	**2-4**	**3**	**3-4**

* alle Ausbildungsprogramme wurde ISCED Stufe 3 zugeordnet: zwei auf Stufe 3A/B, drei auf Stufe 3B und drei auf Stufe 3C. Da alle Finnischen und Schwedischen Ausbildungsprogramme Möglichkeiten für den Hochschulzugang bieten werden diese Programme in der Feaibility Study auf ISCED Stufe 3A/B eingruppiert.
** Pilotinitiative im Zeitraum 2008-2010 - noch nicht Bestandteil des Bildungssystems. Es ist vorgesehen, dieses neue Ausbildungssystem in das reguläre Bildungssystem ab 2011 zu integrieren.

Institutionelle Differenzen

Für die Analyse von *Governancestrukturen* existiert eine breite wissenschaftliche Diskussion darüber, nach welchem Modell man sie differenzieren bzw. typisieren sollte. Wir haben nach ausführlicher Diskussion der Möglichkeiten darauf verzichtet, sie entlang theoretisch hoch aggregierter Typologien in Anschluss an Exping-Andersen (1990) oder Greinert (2005) zu differenzieren, da diese leicht dazu verführen, eine „Vogelperspektive auf breite Charakteristika einer sozialen oder historischen Situation" (Arts, Gelissen 2002, S. 139) zu geben und die differenzierte Realität zu verfehlen. Wir haben die Governancestrukturen auf einem Kontinuum zu klassifizieren versucht entlang dem Einfluss, den auf der einen Seite der Staat, auf der anderen Seite die Sozialpartner als Hauptakteure in Berufsbildung und Arbeitsmarkt auf die wichtigen Entscheidungen in der Berufsbildung haben. Auf dem einen Pol mit Haupteinfluss des Staates und eher geringerem Einfluss der Sozialpartner steht Schweden. Der andere Pol wird von Ländern mit einem korporatistischen Governancesystem gebildet: Berufsbildung in Österreich, Dänemark, Deutschland und der Schweiz wird bei der Entwicklung von Standards und bei Kontrolle der Ausbildung stark von den Sozialpartnern beeinflusst. Finnland, Norwegen und Slowenien sind irgendwo zwischen diesen Polen zu verorten. Neben den Einflussanteilen ist bei den Governancestrukturen auch darauf zu achten, ob Entscheidungen eher zentral oder dezentral organisiert werden. Hier werden Differenzen vor allem zwischen den traditionellen dualen Systemen der deutschsprachigen Länder und Dänemark und Finnland mit stärker dezentralen Steuerungsmomenten deutlich.

Die *Formen der Finanzierung* differieren in den Ländern nicht nur nach den Anteilen staatlicher oder privater Kostenaufbringung für die Berufsausbildung. Innerhalb der privat (wirtschaftlich)en Finanzierung kann der Unterschied, ob die Kosten einzelwirtschaftlich oder über Fondlösungen aufgebracht werden, für die Ausbildungsorganisation wichtig sein. Ähnlich verhält es sich bei staatlichen Finanzierungssystemen mit der Frage, ob die Entscheidungshoheit über die Verausgabung der Ressourcen eher bei zentralen oder lokalen Autoritäten liegt.

Die *Ausbildungsform* nach Maßgabe der Anteile von schulischer und betrieblicher Unterweisung variieren nicht nur zwischen den Ländern, sondern auch innerhalb eines Landes nach Ausbildungsbereichen zum Teil erheblich. Dies macht es schwer, das Verhältnis zwischen schul- und betriebsbasierten Ausbildungsanteilen in jedem Land eindeutig zu bestimmen. Als große Linie schält sich heraus, dass in den Länder mit dominant dualen Systemen (Dänemark, Deutschland, Österreich, Schweiz) der Anteil betriebsintegrierter Ausbildung zwischen 55 % und 80 % variiert, während in eher schulisch organisierten Ausbildungssystemen dieser Anteil deutlich niedriger liegt. (In Schweden bei 15 %, in Slowenien bei maximal 24 %).

Der Frage der *Qualitätssicherung in der Berufsbildung* kommt in fast allen Länder-Berichten ein hoher Stellenwert zu. Sie markiert auch eines der Hauptthemen der europäischen Berufsbildungspolitik und –forschung (vgl. Bulgarelli 2008; Lipinska, Schmid, Tessaring 2007; European Union 2008). Sie betrifft alle Maßnahmen, die in Ländern zur Verbesserung der Kompetenzen von Lehrern und Ausbildern, zur besseren Kooperation zwischen unterschiedlichen Ausbildungsträgern oder zur Kontrolle der Prozesse und zur Erhöhung der Transparenz über Ausbildungsleistungen von Ausbildungsanbietern ergriffen werden. In den Länder-Berichten spielt die Qualität des Personals eine geringe Rolle im Vergleich mit der Entwicklung von

Qualitätsstandards für die Evaluation von Ausbildung. Hier verläuft eine deutliche Grenzziehung zwischen Ländern, die stärker auf zentrale Standards setzen (wie die traditionellen dualen Systeme), und solchen, die den Akzent auf die dezentrale Implementation und Kontrolle von Standards legen (wie z. B. Dänemark und Finnland).

Wir wissen wenig über die Wirkung, die der *Zugang zu Hochschulbildung* und zum Arbeitmarkt auf die Motivation der Auszubildenden und auch auf die Qualität der Ausbildung hat. Soweit die Länder-Berichte diese Aspekte aufgreifen, zeigen sich extreme Differenzen. Bei der Durchlässigkeit zur Hochschule stehen Schweden und Finnland mit 76 % eines Altersjahrgangs, die ein Studium beginnen, an der Spitze, Deutschland mit 35 % am Ende der acht beteiligten Länder. Letzteres wird von Experten nicht zuletzt darauf zurückgeführt, dass die Durchlässigkeit von der Berufsbildung zur Hochschule in Deutschland nur zwischen ein und zwei Prozent der Studienanfänger liegt. Anders verhält es sich beim Übergang in den Arbeitmarkt: Hier haben die vier Länder mit dualen Systemen weniger Friktionen als Schweden und Finnland, auch wenn sich die Abstände (z.B. in den Jugendarbeitslosigkeitsquoten) im letzen Jahrzehnt zwischen den Ländergruppen deutlich verringert haben.

Die hier nur angedeuteten institutionellen Regulierungsdifferenzen von Ausbildung zwischen den Ländern machen deutlich, wie viel Kontextvariable in einer international vergleichenden Hauptstudie VET-LSA zu beachten und zu kontrollieren sein werden, wenn man Kompetenzunterschiede nicht nur messen, sondern von ihnen her auch Rückfragen an das institutionelle Ausbildungsumfeld stellen will. Man wird sinnvoller Weise zwischen Faktoren, die direkt auf die Ausbildungsqualität wirken dürften, weil sie prozessnah sind, wie die Kompetenz der Ausbilder, Standards und Normen für die Prozessgestaltung und ihre (Qualitäts-)Kontrolle, und indirekt wirkenden Faktoren der Governancestrukturen, der Finanzierung und des Zusammenhangs zu höherer Bildung und Arbeitsmarkt unterscheiden. Für letztere wird man zunächst mögliche Wirkungsketten theoretisch konstruieren müssen, um gehaltvolle Hypothesen und Indikatoren zu entwickeln.

4. Schnittmengen in Berufsprofilen und Ausbildungsgängen

Der Intention dieses Artikels entsprechend, Konzept und Durchführung der Feasibility Study VET-LSA zur Diskussion zu stellen, kann es im Folgenden nicht darum gehen, die Ergebnisse der Studie ausführlich darzustellen, sondern sie nur so weit aufzugreifen, wie es zum Verständnis und zur Prüfung der Validität der eingesetzten Verfahren und Methoden nötig ist.[31] Wir werden zunächst in sehr komprimierter (und für den Leser anstrengender) Form das zentrale und für die Weiterarbeit am VET-LSA entscheidende Resultat der Experten-Ratings präsentieren: die hohe Übereinstimmung in den Beurteilungen bei fast allen geprüften Merkmalen, die signalisiert, dass die Schnittmengen in den Berufsprofilen und den ihnen entsprechenden Ausbildungsgängen zwischen den acht Ländern groß sind. Danach werden wir an einem Beispiel („electrician craft“) den Umgang mit den Daten einsichtig machen.

31 Alle weiteren detaillierten Analysen der einzelnen Berufsfelder sind weiteren Arbeiten insbesondere der internationalen Koordinatoren – die für die berufsspezifischen Kapitel verantwortlich zeichnen – vorbehalten.

Schließlich werden wir die ratings für die Prüfungsaufgaben am Ende der Ausbildung für zwei Berufsfelder darstellen. Ihnen kommt, neben den Arbeitsaufgaben und Qualifikationsanforderungen, für die spätere Itementwicklung ein besonders hoher Stellenwert zu, da sie inhaltliche Übereinstimmungsbereiche abbilden, von denen her Testitems zur Messung domänenspezifischer Kompetenzen entwickelt werden können. Für alle Analysefelder – Arbeitsaufgaben und Qualifikationsanforderungen der Berufstätigkeiten sowie Prüfungsaufgaben stellvertretend für zentrale Inhaltsbereiche am Ende der Ausbildung– gilt, dass wir nirgends volle Übereinstimmung erwartet haben und benötigen. Dies war auch bei den Untersuchungsbereichen im Schul-PISA nicht der Fall. In gleicher Weise genügt es für ein Berufsbildungs-Large-Scale-Assessment, wenn hinreichend große Schnittmengen insbesondere in den Kernbereichen von Ausbildung und Berufen nachgewiesen werden können, und das ist der Fall.

4.1 Das Gesamtbild der Experten-Einschätzungen

In Tabelle 3 werden die Übereinstimmungsmaße (Kendall's W) für alle in den Expertenratings beurteilten Merkmale für alle 4 bzw. 5 Berufe/Berufsfelder aufgelistet und die Signifikanz der Ergebnisse ausgewiesen. Kendall's W measure for concordance misst die Übereinstimmung zwischen allen beteiligten Ländern (vgl. Tabelle 3, Spalte 4 „N"). Die beurteilten Merkmale sind bei jedem Beruf in der Reihenfolge der Tabelle (Spalte 3) aufgeführten Arbeitsaufgaben (Relevanz), Arbeitsaufgaben (Häufigkeit in der täglichen Arbeit), Qualifikationsanforderungen (Relevanz), Prüfungsaufgaben (Wichtigkeit für den Ausbildungsabschluss) und Prüfungsaufgaben (Komplexität) aufgeführt.

Wie nicht anders zu erwarten, unterscheiden sich die Übereinstimmungsgrade zwischen den Berufen und auch innerhalb der Berufe nach Merkmalen. Entscheidend ist, dass bis auf zwei Merkmale – Prüfungsaufgaben bei „business and administration" und „electrician industry" – alle anderen ratings Übereinstimmung auf einem hohen Signifikanzniveau (zumeist unter 1 ‰) ergeben. Extrem hohe Werte finden wir bei den „carmechatronics". Doch auch beide Elektroniker-Berufe und „social and health care" zeigen sich große Schnittmengen für ein VET-LSA, was bei letztem Berufsfeld nicht erwartet werden konnte, weil es sehr breit ist. Beim zweiten breiten Berufsfeld, „bussiness and administration", sind die ratings für die Prüfungsaufgaben nicht ganz so eindeutig, was eher auf die Auswahl der Prüfungsaufgaben als auf die bestehenden gemeinsamen Kernbereiche für Ausbildungen in diesem Bereich zurückzuführen ist[32].

32 Dies wurde von der internationalen Expertengruppe bestätigt.

Tab. 3: Kendalls'W measure for concordance by occupational fields

Vocational Area	Workshop	Set of Items	N*	Kendall's W	Chi2	df	p	Average Spearman correlation	Omitted items[1]
Business and administration	1	Occ Task Relevance	6	0,630	30,248	8	0,000	0,56	
	1	Occ Task Frequency	6	0,634	26,625	7	0,000	0,56	item 4
	1	Qualification requirement Relevance Qualification	6	0,591	42,567	12	0,000	0,51	
	2	Eval Task Major Educational Objective	6	0,204	15,898	13	n.s.	0,04	
	2	Eval Task Complexity	6	0,692	53,960	13	0,000	0,63	
Carmechatronics	1	Occ Task Relevance	7	0,701	63,809	13	0,000	0,65	
	1	Occ Task Frequency	7	0,852	77,495	13	0,000	0,83	
	1	Qualification requirement Relevance Qualification	7	0,864	48,394	8	0,000	0,84	
	2	Eval Task Major Educational Objective	7	0,411	43,145	15	0,000	0,31	
	2	Eval Task Complexity	7	0,517	54,289	15	0,000	0,44	
Electricians (craft)	1	Occ Task Relevance	7	0,450	34,648	11	0,000	0,36	
	1	Occ Task Frequency	7	0,431	30,154	10	0,001	0,34	item 7
	1	Qualification requirement Relevance Qualification	7	0,570	31,903	8	0,000	0,50	item 7
	2	Eval Task Major Educational Objective	6	0,318	22,918	12	0,028	0,18	item 7
	2	Eval Task Complexity	6	0,788	51,990	11	0,000	0,75	items 6, 7
Electricains (industry)	1	Occ Task Relevance	6	0,615	36,881	10	0,000	0,54	
	1	Occ Task Frequency	6	0,574	34,416	10	0,000	0,49	
	1	Qualification requirement Relevance Qualification	6	0,505	24,259	8	0,002	0,41	
	2	Eval Task Major Educational Objective	6	0,301	18,080	10	0,054	0,16	
	2	Eval Task Complexity	6	0,655	39,274	10	0,000	0,59	
Social and health care	1	Occ Task Relevance	7	0,479	60,356	18	0,000	0,39	
	1	Occ Task Frequency	7	0,743	62,399	12	0,000	0,70	items 11, 15, 16, 17, 18, 19
	1	Qualification requirement Relevance Qualification	7	0,476	23,325	7	0,001	0,39	
	2	Eval Task Major Educational Objective	7	0,278	29,193	15	0,015	0,16	items 14
	2	Eval Task Complexity	7	0,530	51,951	14	0,000	0,45	items 14, 16
			* N = number of countries						

1 In case of missing country means for individual items the item was onitted from the analysis rather than the country.

Ohne hier ins Detail der ratings zu gehen, lässt sich festhalten, dass die Expertenratings bei den Berufsprofilen (Arbeitsaufgaben und Qualifikationsanforderungen) durchweg und über alle Länder ein sehr einheitliches Bild zeichnen, das diese Berufe/Berufsfelder für ein VET-LSA empfiehlt[33]. Man kann angesichts dieser Ergebnisse sogar fragen, ob nicht die hohen Übereinstimmungsgrade eher erklärungsbedürftig sind, als die ebenfalls noch bleibenden Differenzen, die am Beispiel des „electrician craft" im nächsten Abschnitt sichtbar gemacht werden.

Für die hohen Übereinstimmungen lassen sich ein methodisches und ein mehr theoretisches Argument ins Feld führen. Auf das methodische hatten wir bereits hingewiesen: Die ratings basieren auf gründlichen Expertendiskussionen, weil es nicht um ein Zufallsbild ging. Insofern werten wir die Ergebnisse auch als Bestätigung der Untersuchungsmethode. Das theoretische Argument fußt auf der seit über einem Jahrzehnt in den Sozialwissenschaften geführten Diskussion über die Frage, ob eher Konvergenz oder Divergenz die Hauptrichtung bei der Organisation der Erwerbsarbeit zwischen den Ländern markiert (vgl. Berger, Dore 1996). Angesichts fortschreitender Globalisierung der Märkte und der Ubiquität der neuen Technologien in modernen Gesellschaften scheinen konvergierende Tendenzen in der Prägung von Berufsprofilen ein zunehmendes Gewicht zu erhalten, auch wenn damit länder- oder andere regionenspezifische Merkmale nicht völlig verschwinden. Auch dies ist bei einem VET-LSA im Auge zu behalten.

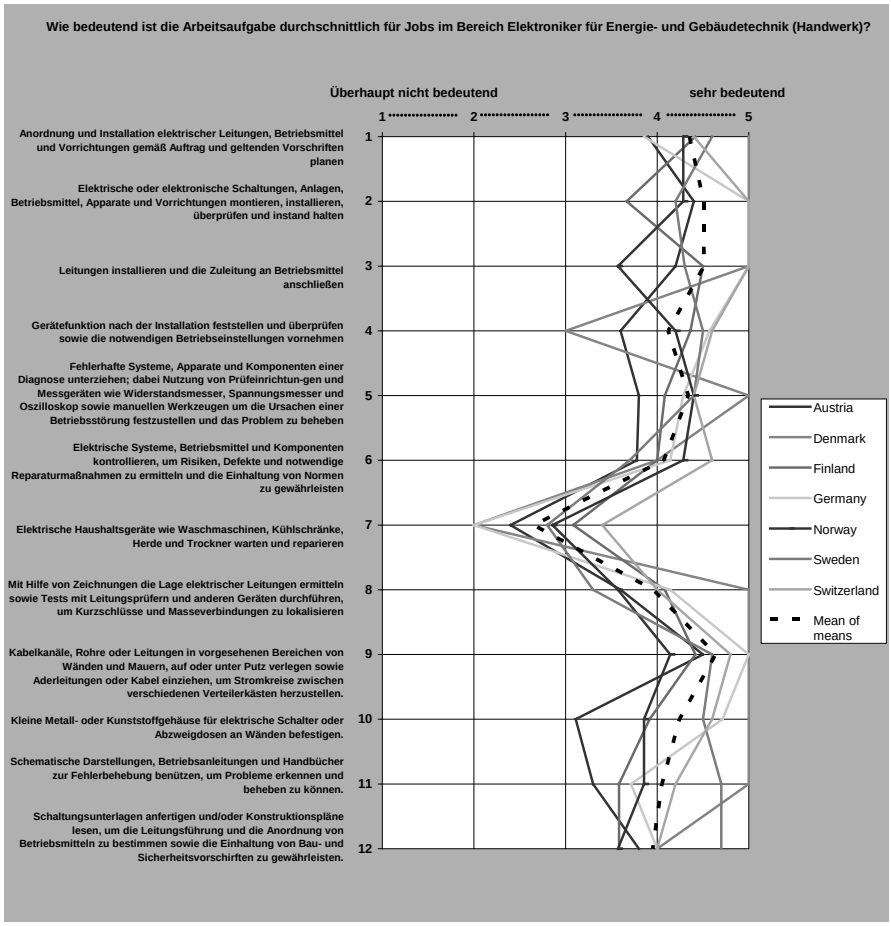
4.2 Resultate der Expertenratings am Beispiel „Eletricians craft"

Die nachfolgenden Abbildungen zeigen die Ergebnisse der Expertenratings für den Bereich „Electricians craft" im Ländervergleich (Österreich, Dänemark, Finnland, Deutschland, Norwegen, Schweden, Schweiz). Für alle Itemsets (Arbeitsaufgaben, Qualfikationsanforderungen, Prüfungsaufgaben) konnte ein ausreichend hoher Grad an Übereinstimmung erreicht werden. Kendall's W bestätigt signifikante Ergebnisse für alle drei Bereiche.

Die gemeinsamen Arbeitsaufgaben wurden von den Experten in den Ländern sehr einheitlich geratet; wie Abbildung 3 zeigt, verläuft das Profil für die einzelnen Arbeitsaufgaben in allen Ländern ähnlich. Mit Ausnahme von Aufgabe 7 (Elektrische Haushaltsgeräte wie Waschmaschinen, Kühlschränke, Herde und Trockner warten und reparieren) wurden alle Arbeitsaufgaben als „eher bedeutend" oder „sehr bedeutend" eingeschätzt, keine der Aufgaben wurde als „überhaupt nicht bedeutend" gewertet. Die Ergebnisse zeigen, dass insbesondere Arbeitsaufgaben in den Bereichen Installationstechnologie, Fehlersuche und Instandhaltung in allen Ländern eine zentrale Rolle für Jobs im Bereich „electrician craft" spielen (Kendalls W=.45, p=.000). Das gemeinsame Profil an Arbeitsaufgaben für „electrician creaft" in allen Ländern ist durch Installation diverser elektronischer Anlagen charakterisiert. Dies betrifft insbesondere Arbeitsplanung, Installation elektrischer Leitungen (z. B. elektronische Schaltungen) sowie Fehlersuche und Problembehebung mit Hilfe technischer Unterlagen (z. B. Schaltpläne) unter Berücksichtigung nationaler Regeln, Normen und Sicherheitsbestimmungen.

33 Weitere Eingrenzungen, insbesondere der noch breiten Berufsfelder, sollten im Rahmen weiterer Vorbereitungen für ein VET-LSA gemeinsam mit den Ländern vorgenommen werden

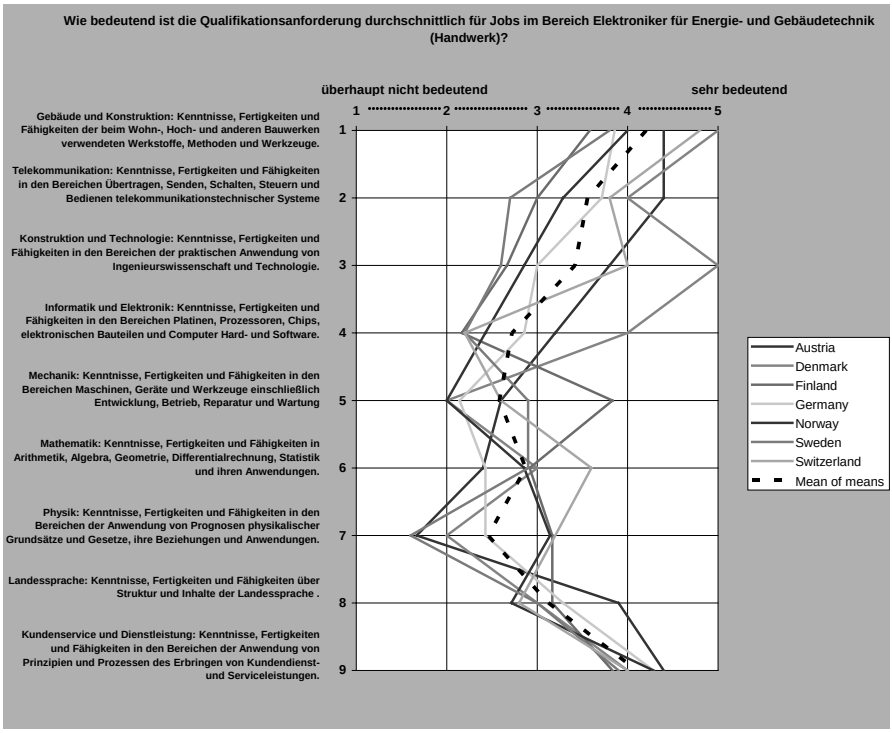
Abb. 3: Ergebnisse der Expertenratings für Arbeitsaufgaben im Bereich „Electricians craft“[1)]



1) Übersetztes Erhebungsinstrument für Deutschland, Österreich, Schweiz.

Bei den Qualifikationsanforderungen waren sich die Experten einig, dass insbesondere Gebäude und Konstruktion, Konstruktion und Technologie, Landessprache sowie Kundenservice und Dienstleistungen von hoher Bedeutung für Jobs im Bereich „electricians craft“ sind. Kendalls W (W=.57, p=.000) bestätigt auch hier signifikante Ergebnisse. Grundsätzlich wurde für die domänenspezifischen Qualifikationsanforderungen höhere Übereinstimmung als für die allgemeinen Qualifikationsanforderungen erreicht. Dies mag dadurch begründet sein, dass allgemeine Qualifikationsanforderungen, wie Mathematik oder Physik, immer in domänenspezifischen Qualifikationsanforderungen enthalten und zur erfolgreichen Bewältigung von Arbeitsaufgaben unabdingbar sind. Das gemeinsame Profil an Qualifikationsanforderungen für Jobs im Bereich „Electrician craft“ ist insbesondere durch Kenntnisse, Fertigkeiten und

Abb. 4: Ergebnisse der Expertenratings für Qualifikationsanforderungen im Bereich „Electricians craft“[1)]



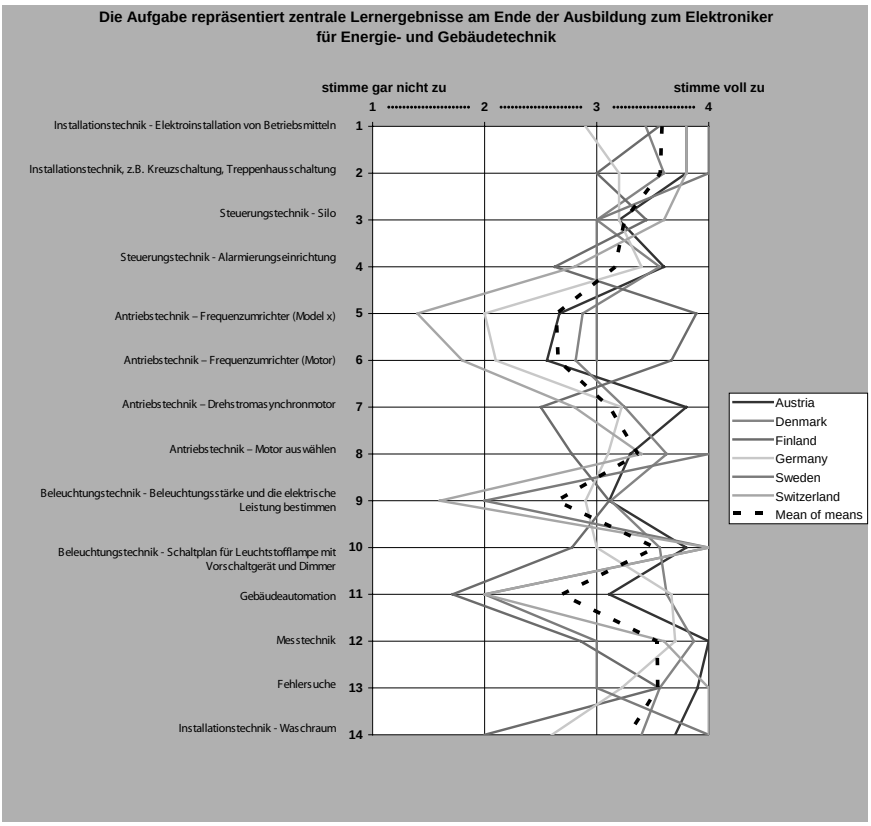
1) Übersetztes Erhebungsinstrument für Deutschland, Österreich, Schweiz.

Fähigkeiten im Bereich Gebäude und Konstruktion charakterisiert. Diese umfassen z.B. Installationstechnologie, Arbeitsorganisation, Funktionsprüfungen, Fehlersuche, Telekommunikation und allgemeine Kenntnisse, Fertigkeiten und Fähigkeiten der Landessprache sowie Kundenservice und Dienstleistungen.

4.3 Prüfungsaufgaben am Ende der Ausbildung

Abb. 5 zeigt Anlage und Skalierung der Prüfungsaufgaben am Ende der Ausbildung für den electrician craft. Der Vergleich der ratings für die Prüfungsaufgaben zeigt, dass auch hier weitestgehend Einigkeit in den Ländern besteht (Kendals W=.32, p=.028). Die etwas geringere Übereinstimmung der ratings im Vergleich zu den Ergebnissen für Arbeitsaufgaben und Qualifikationsanforderungen ist eher der gewählten Methode als den gemeinsamen Schnittmengen an sich zuzuschreiben. Der Einsatz von Prüfungsaufgaben als Instrument für den Vergleich von Lernergebnissen in zentralen Inhaltsbereichen am Ende der Ausbildung führte in den nationalen Workshops bisweilen zu Diskussionen über den Zuschnitt der Prüfungsaufgaben und

Abb. 5: Ergebnisse der Expertenratings für Prüfungsaufgaben im Bereich „Electricians craft“[1)]



1) Übersetztes Erhebungsinstrument für Deutschland, Österreich, Schweiz.

weniger über die generelle Vergleichbarkeit gemeinsamer Inhalte repräsentiert über unterschiedliche Prüfungsaufgaben. Dieses Missverständnis führte nach Auskunft nationaler Experten in Einzelfällen zu Verzerrungen in den ratings.

Über die Prüfungsaufgaben sollte die inhaltliche Struktur der beruflichen Ausbildungsgänge auf ihre Vergleichbarkeit hin geprüft werden (vgl. Abschnitt 2.4). Da sie auch am ehesten Anknüpfungspunkte für die Testentwicklung eines VET-LSA, dessen Hauptmesszeitpunkt kurz vor Ausbildungsende liegen soll, bietet, gehen wir auf sie etwas ausführlicher, d. h. ergänzend zum electrician craft noch für zwei Berufe, einen gewerblich-technischen (Carmechatronisc) und einen Dienstleistungsberuf (Social and Health Care) ein. Die Tabellen 4 und 5 sollten auch in der Kurzversion der Prüfungsaufgaben einsichtig machen, dass es sich bei Ihnen überwiegend um Aufgaben handelt, die eine Lösung praktischer Probleme fordern.[34] Die Tabellen sind so aufgebaut, dass zunächst die Kernbereiche aufgelistet sind und dann die ihnen

34 Beim Experten-Rating wurde jede Aufgabe mit einem mindestens halbseitigen Text ausdifferenziert (im gewerblich-technischen Bereich ergänzt ggf. um technische Angaben und Diagramme).

jeweils zugeordneten Prüfungsaufgaben mit dem Durchschnitt aller Expertenratings aller Länder nach den beiden Merkmalen Wichtigkeit für Abschlussprüfungen und Komplexität der Aufgabe. Um eine endgültige Zuordnung des ratings zum positiven oder negativen Pol zu erreichen, haben wir mit einer 4stufigen Skala gearbeitet.

Bei „carmechatronics" liegen alle Bewertungen der Wichtigkeit wie auch Komplexität der Prüfungsaufgabe auf dem oder oberhalb des arithmetischen Mittels (2,5), die Mehrheit tendiert zu sehr wichtig (ab 3,5). Insbesondere die Aufgaben aus den Kernbereichen Motormanagement, Bremssystem und Kraftübertragung werden hoch eingeschätzt. Innerhalb der zentralen Inhaltsbereiche kann man die Aufgaben danach typisieren, ob sie mehr dem Bereich Diagnose/Fehlersuche oder Reparatur oder Instandhaltung zuzurechnen sind. Insbesondere die Diagnose-/Fehlersuche-Aufgaben weisen hohe Komplexitätsniveaus auf. Alle fünf Kernbereiche repräsentieren wichtige Zielsetzungen in allen nationalen Ausbildungsprogrammen und sollten, neben den Arbeitsaufgaben und Qualifikationsanforderungen, für die Testentwicklung in einem VET-LSA berücksichtigt werden.

Tab. 4: Prüfungsaufgaben für „carmechatronics" Expertenratings

Zentrale Inhaltsbereiche am Ende der Ausbildung:
1. Motormanagement
2. Bremssystem
3. Fahrwerk
4. Kraftübertragung
5. Komfort- und Sicherheitssysteme

Inhaltsbereich	Aufgabentyp	Repräsentiert zentrale Inhaltsbereich am Ende der Ausbildung MWC*	Komplexität der Aufgabe MWC
1.1	Reparatur Einspritzanlage (Benzin) und Abgastest	3,7	3,4
1.2a	Reparatur Dieseneinspritzung	3,6	3,5
1.2b	Fehlersuche Abgaswerte (Diesel)	3,7	3,2
1.3a	Zahnriemenwechsel (praktisch)	3,7	2,7
1.3b	Zahnriemenwechsel (theoretisch)	3,5	2,5
2.1	Inspektion Bremsen und ABS	3,7	2,9
2.2	Fehlersuche Bremssystem (ABS/ESP)	3,7	3,4
3.1	Achsenvermessung	3,6	2,9
3.2	Federung: Diagnose/Fehlersuche und Austausch (theoretisch)	3,1	2,8
4.1	Schaltgetriebe zerlegen, zusammenbauen, Fehlerprotokoll	2,8	3,0
4.2	Reparatur Kupplung	3,5	2,6
4.3	Reparatur Antriebswelle	3,4	2,5
5.1	Reparatur Klimaanlage	3,1	3,1
5.2a	Fehlersuche Klimaanlage	3,1	3,5
5.2b	Wartung Klimaanlage	3,4	2,8
5.3	Programmierung Komfortfunktionen	3,1	2,9

* Mittelwert aller Länder

Tab. 5: Prüfungsaufgaben für „social and health care“ Expertenratings

Zentrale Inhaltsbereiche
1. Patienten/Klienten-versorgung, Assistenz im Versorgungsprozess
2. Kommunikation und Beziehungsaufbau
3. Inter- und intradisziplinäre Kooperation
4. Gesundheit und Sicherheit
5. Administration und rechtliche Rahmenbedingungen

Inhalts-bereich	Aufgabentyp	Repräsentiert zentrale Inhalts-bereich am Ende der Ausbildung	Komplexität der Aufgabe
		MWC	MWC
1.1	Komplexer Fall einer älteren Frau wartend auf einen Pflegeheimplatz (Frida Petersen)	2,8	3,5
1.2	Planung, Durchführung, Dokumentation und Evaluation eines Altenpflegeprozesses (Mutter hat sich verändert)	3,1	3,12
1.3	Beurteilung Notfall einer alten Witwe nach Koronararterien-Vorfall (Heim)	3,1	2,9
1.4	Beschreibung Zustand eines bewusstlosen Mannes (Erik)	3,1	2,6
1.5	Kultur- und handlungstheoretische Ansätze (Psychiatrie)	2,6	3,5
2.1	Kultur- und handlungstheoretische Ansätze: Kommunikation mit Migranten (Serhart)	2,8	3,2
2.2	Behandlung Neuzugänge Altenheim (Herr Tebbe)	3,1	3,0
2.3	Dialog mit hörgestörten Alten (Frau Johansson)	3,2	2,1
2.4	Unterstützende Kommunikation bei sprechgestörten Patienten	3,0	2,5
3.1	Arbeitsplanungs- und Zusammenarbeitskonzepte auf Krankenstation (Sörensen)	3,1	3,5
3.2	Schule für Schülerinnen und Schüler mit besonderem Förderbedarf	2,7	3,0
4.1	Abteilung für ältere Menschen: Unterstützung Querschnittsgelähmter	3,4	2,6
4.2	Verwendung persönlicher Schutzausrüstungen bei Operation	2,0	1,7
5.1	Ambulante Pflege nach Operation (Mutter verändert)	3,0	2,9
5.2	Staatliche Unterstützung nach Schlaganfall (A.+E. Svensson)	3,1	2,7
5.3	Qualität und Effizienz (Begriffsklärungen)	2,8	2,6
5.4	Datenschutz und berufliche Diskretion	3,4	2,4

Bei dem zweiten Berufsfeld („Social and Health Care“) hat überrascht, dass die Übereinstimmung bei den Prüfungsaufgaben zwischen den Ländern recht hoch ist.

Dies drückt sich auch darin aus, dass von 17 Prüfungsaufgaben 16 bei der Wichtigkeit für den Ausbildungsabschluss über dem arithmetischen Mittel liegen, bei der Einschätzung der Komplexität sind es Aufgaben des Kernbereichs „Patienten-/Klientenversorgung, Assistenz im Versorgungsprozess", bei denen es sich nicht um unmittelbar medizinische Eingriffe, sondern um soziale Interventionen in schwierigen Situationen handelt; ferner Aufgaben in „Kommunikation und Beziehungsaufbau". In der Regel finden solche Aufgaben die höchste Wichtigkeitsattribution, die zu ihrer Lösung die Reflektion auf komplexe Fallbeschreibungen erfordert. Hierher gehören auch die als sehr komplex eingestuften Aufgaben der inter- und intradisziplinären Kooperation (vgl. Tabelle 5, Item 3.1 und 3.2).

6. Fazit

Die Hauptintention des Artikels war die Vorstellung eines methodischen Modells zur komparativen Analyse von Berufsprofilen und Ausbildungsgängen, die die unabdingbare Voraussetzung für ein VET-LSA abgibt. Nicht allein die Ergebnisse scheinen die Tragfähigkeit des Methoden-Konzepts zu bestätigen. Auch die zwischenzeitlich geführten internationalen Diskussionen auf politischer und wissenschaftlicher Ebene haben in der Feasibility Study Ansatzpunkte für die weitere Arbeit an einem VET-LSA wie auch die Übertragbarkeit des Verfahrens auf weitere Länder, die sich an einem VET-LSA beteiligen möchten, ausgemacht. Hier war insbesondere das positive Votum der beiden externen (außerhalb der beteiligten Länder) wissenschaftlichen Evaluatorinnen der Feasibility Study – L. Unwin (UK) und A. Bouder (Frankreich) wegweisend. Alle bisherigen Rückmeldungen entziehen das Methoden-Modell nicht der weiteren wissenschaftlichen Auseinandersetzung, die die dringend auszubauende komparative Forschung in der Berufsbildung nur befruchten kann.

Bezogen auf das inhaltliche Ziel der Feasibility Study kam die internationale Expertengruppe sowohl in Bezug auf die institutionellen Bedingungen der vier ausgewählten Berufsbildungsgänge als auch auf die inhaltlichen Schnittmengen bei den Berufsprofilen und Ausbildungsprogrammen im abschließenden Workshop einstimmig zu dem Ergebnis, dass die Voraussetzungen für ein VET-LSA durch die Feasibility Study nachgewiesen seien und alle vier Berufsfelder einbezogen werden sollten; bei „business and administration" und „social and health care" mit zusätzlichen Spezifizierungen: im ersten Fall erscheint eine weitere Eingrenzung auf berufliche und Ausbildungsschwerpunkte erforderlich, im zweiten Fall wird eine Fokussierung auf Berufe empfohlen, die im Überlappungsbereich von Sozial- und Gesundheitspflege liegen (in Deutschland z. B. Altenpflege).

Die Ergebnisse der Feasibility Study sollen nach Votum der internationalen Steering Group mit Regierungsvertretern aus den acht Ländern in die weitere Arbeit an einem VET-LSA eingehen. Dies betrifft die vier Berufe/Berufsfelder und auch die Items für Arbeitsaufgaben, Qualifikationsanforderungen und die zentrale Inhaltsbereiche am Ende der Ausbildung (abgebildet durch Prüfungsaufgaben): z.B. sollten die Items, die bei den ratings die Schwelle von 3 bei der Relevanzeinschätzung überschritten haben, als Ausgangspunkt für die Entwicklung von Testaufgaben der Kompetenzmessung Berücksichtigung finden. Bei der Entwicklung von Testaufgaben für domänspezifische Kompetenzen sollte – wie in der Feasibility Study betont – praktischen Aspekten ein hohes Gewicht beigemessen werden.

Die nächsten Schritte für ein VET-LSA sind Aufgabe der Politik. Sie bestehen darin zu klären, welche Länder sich in welchen Berufen/Berufsfeldern beteiligen wollen, welche Ressourcen zur Verfügung gestellt werden können und wie ein internationaler „Call for Tender" gestaltet sein soll.

Literatur

Achtenhagen, F., Winther, E. (in press). Konstruktvalidität von Simulationsaufgaben: Computergestützte Messung berufsfachlicher Kompetenz – am Beispiel der Ausbildung von Industriekaufleuten. Bonn: BMBF.

Arts, W., Gelissen, J. (2002): Three worlds of welfare capitalism or more? A State-of-the-Art Report. In: Journal of European Social Policy 2002; 12, p. 137-158.

Baethge M., Arends, L., Winther, E. (2008): First international workshop Feasibility Study VET-LSA, Discussion paper, workshop 3.-4. July, 2008 in Bonn. SOFI Göttingen.

Baethge, M., Achtenhagen, F., Arends, L. (2008): How to Compare the Performance of VET Systems in Skill Formation. In: K.-U. Mayer, H. Solga (Hrsg.). *Skill Formation – Interdisciplinary and Cross-National Perspectives.* New York: Cambridge University Press.

Baethge, M., Achtenhagen, F., Arends, L., Babic, E., Baethge-Kinsky, V., Weber, S. (2006): Berufsbildungs-PISA Machbarkeitsstudie. Stuttgart: Steiner.

Baethge, M., Achtenhagen, F., Arends, L., Babic, E., Baethge-Kinsky, V., Weber, S. (2006): PISA-VET. A Feasibility-Study. Stuttgart: Steiner.

Baethge, M., Achtenhagen, F., Arends, L., Winther, E., Nickolaus, R., Gschwendtner, T., van Onna, M., Béguin, A., Sibberns, H., Watermann, R. (2007): Concept for an International Large Scale Assessment of Vocational Education and Training (VET-LSA). Working paper for the international workshop, 29.-30. October, 2007 in Bonn. SOFI Goettingen.

Bell, D. (1973): The coming of Post-Industrial Society, New York.

Berger, S., Dore, R.P. (Ed.) (1996): National diversity and global capitalism, Ithaca, New York.

Bortz, J., Lienert, G. A. (2008): Kurzgefasste Statistik für die klinische Forschung: Leitfaden für die verteilungsfreie Analyse kleiner Stichproben. Heidelberg: Springer.

Braithwaite, J., Luft, S., Bender, W., Callen, J., Westbrook, J.I., Westbrook, M.T., Mallock, N.A., Iedema, R., Hindle, D., Jochelson, T. (2007): The hierarchy of work pursuits of public health managers. Health Service Management Research, 20; 2.

Bulgarelli, A. (Ed.): Initial vocational education and training (IVET) In Europe-Review.

CEDEFOP (2009): Continuity, consolidation and change: towards a European era of

Child, J., Heavens S.J. (2001): The Social Constitution of Organizations and is Implications for Organizational Learning. In: Dierkes, M., Antal, A.B., Child, J., Nonaka, I: (Hg.),

Craigie, M., Loader, B., Burrows, R., Muncer, S. (2002): Reliability of Health Information on the Internet: An Examination of Experts' Ratings. Journal of Medical Internet Research, 4;1.

Descy, P., Nestler, K.; Tessaring, M. (2005): Internationally comparable statistics on education, training and skills: current state and prospects. In: Vocational Training No 36 (European Journal).

Descy, P., Tessaring, M. (2002): Kompetenz für die Zukunft – Ausbildung und Lernen in Europa, Luxemburg.

Esping-Andersen, G. (1990): The three worlds of welfare capitalism. Cambridge: Polity Press, 1990.

European Union (2008). The Bordeaux Communiqué on enhanced European cooperation in vocational education and training. Download at: http://www.eua.be/fileadmin/user_upload/files/Newsletter_new/Bordeaux_Communique_EN.pdf

Felstead, A., D. Gallie, F. Green, Y. Zhou (2007): Skills At Work, 1986 to 2006. University of Oxford, SKOPE.

Green, F. (2009): Final Report on the Validation of the Results of the Job Requirements Approach (JRA) Pilot Survey. OECD

Greinert, W.-D. (2005): Mass Vocational Education and Training in Europe. Classical models of the 19th century and training in England, France and Germany during the first half of the 20th. CEDEFOP Panorama, 118. Luxembourg: European Communities.
IG-Metall Bezirk Baden-Württemberg (2004). ERA-Wissen Handbuch I. Lohr
International Labour Organisation (2009): Download at: http://www.ilo.org/public/english/bureau/stat/isco/index.htm
Luomi-Messerer, K., Markowitsch, J. (Hrsg., 2006): VQTS model. A proposal for a structured description of work-related competences and their acquisition. Vienna: 3s research laboratory.
Lutz , B. (1976): Bildungssystem und Beschäftigungsstrukturen in Deutschland und Frankreich. In: Mendius, H.G. u. a. Betrieb-Arbeit-Qualifikation, S. 83. ff.
Lutz, B. (1984): Der kurze Traum immerwährender Prosperität, Frankfurt/New York.
Markowitsch, J., Luomi-Messerer, K. (2008). Entstehung und Interpretation der Deskriptoren des Europäischen Qualifikationsrahmens. Europäische Zeitschrift für Berufsbildung, 42/43. CEDEFOP..
Markowitsch, J., Becker, M., Spöttl, G. (2006): Zur Problematik eines European Credit Transfer System in Vocational Education and Training (ECVET). In: Grollmann, P.; Rauner, F.; Spöttl, G.; (Hrsg.): Europäisierung Beruflicher Bildung – eine Gestaltungsaufgabe. Hamburg: LIT, S. 173-198.
Nickolaus, R., Gschwendtner, T., Abele, S. (in press). Die Validität von Simulationsaufgaben am Beispiel der Diagnosekompetenz von Kfz- Mechatronikern. Vorstudie zur Validität von Simulationsaufgaben im Rahmen eines VET-LSA. Bonn: BMBF.
Nonaka, I., Takeuchi, H. (1997): Die Organisation des Wissens, Frankfurt/New York.
OECD (2004): PIAAC draft strategy paper – Policy objectives, strategic options and cost implications. http//www.oecd.org/dataoecd3/3/34463133.pdf.
OECD (2008): PIAAC.
Pett, M.A. (1997): Nonparametric Statistics in Health Care Research: Statistics for Small Samples and Unusual Distributions. Thousand Oaks, et.al.: Sage.
Polanyi, M. (1966); The Tacit Dimension, New York.
Schleicher, A. (2008): PIAAC: a new strategy for assessing adult competencies. International Review of Education. Springer.
Schmidt, W.H., Raizen, S.A., Britton, E.D., Bianchi, L.J., Wolfe, R.G. (1997): Many visions, many aims: A cross-national investigation of curricular intentions in school science. Dordrecht, The Netherlands: Kluwer Academy Publishers.
UNESCO (2006): ISCED 1997 International Strandard Classification of Education. Reedition. Online: http://www.uis.unesco.org/TEMPLATE/pdf/isced/ISCED_A.pdf

Anschrift der Autoren: Prof. Dr. Martin Baethge, SOFI Soziologisches Forschungsinstitut e.V., An der Georg-August-Universität Göttingen, Friedländer Weg 31, 37985 Göttingen — Dr. Lena Arends, International Bureau of the Federal Ministry of Education and Research at the Project Management Agency c/o German Aerospace Center (DLR), Heinrich-Konen-Str. 1, 53227 Bonn