„Kompetenzmessung in der Gesundheits- und Krankenpflegeausbildung“ in Anlehnung an das Teilprojekt „TEMA“

Inauguraldissertation

zur Erlangung des

Doktorgrades der Pflegewissenschaft

(Dr. rer. cur.)

an der Pflegewissenschaftlichen Fakultät der

Philosophisch-Theologischen Hochschule Vallendar

vorgelegt von:

Simone Brigitte Ries, geb. Henning

Matrikelnummer: 162234591004

Erstgutachter: Prof. Dr. Albert Brühl

Zweitgutachterin: Prof. Dr. Constanze Eylmann

vorgelegt im:

März 2020

Widmungen

Beruflich: allen Auszubildenden und Studierenden in der Gesundheits- und Krankenpflege und der Evangelischen Hochschule in Ludwigsburg

Privat: meinem Mann Bernhard Ries, meinen Eltern Brigitte und Jürgen Henning mit unserer Familiendackeldame Jule und unseren Pateneseln Tommie und Titus

Danksagung

Ganz besonders danke ich meinem Doktorvater, Herrn Prof. Dr. Albert Brühl, Inhaber des Lehrstuhls für Statistik und standardisierte Verfahren der Pflegeforschung an der Philosophisch-Theologischen Hochschule in Vallendar. Er hat mir den Weg durch den quantitativen „Dschungel“ der Erstellung einer Doktorarbeit mit sehr viel Mühe und Geduld, gepaart mit hoher Fachkompetenz, aufgezeigt. Er war absolut wertschätzend und endlos verlässlich! Dafür bin ich sehr dankbar!

Mein besonderer Dank geht ebenfalls an meine Zweitgutachterin, Frau Prof. Dr. Constanze Eylmann, der Leiterin des Bachelorstudiengangs Pflege an der Evangelischen Hochschule in Ludwigsburg, die den Blick auf den pflegerisch-pädagogischen Rahmen und die Möglichkeit auch statistische Wege gehen zu dürfen unterstützt und begleitet hat.

Auf dem Weg der Erstellung dieser Dissertation sind mir viele „fremde bzw. neue“ Menschen begegnet, die mir häufig ohne große Kenntnisse meiner Person großes Entgegenkommen geschenkt haben. Das hat vor allem den Zugang zu den meisten Gesundheits- und Krankenpflegeschulen bzw. Hochschulen enorm erleichtert. Die Projektumsetzung und Übertragung auf die Gesundheits- und Krankenpflege konnte durch den Austausch mit Mitarbeiterinnen und Mitarbeitern aus dem Ursprungsprojekt „TEMA“ erleichtert werden. Für statistischen Austausch und so manchen Hinweis fand ich Rat an der Philosophisch-Theologischen Hochschule in Vallendar, bei den Pädagogischen Hochschulen Karlsruhe und Ludwigsburg, ebenso an meiner eigenen, der Evangelischen Hochschule in Ludwigsburg. An dieser Stelle sollen noch einige Personen aus unterschiedlichsten Kontexten namentlich erwähnen werden: u.a. Sr. Magdalena Schleinschok, Frau Miriam Peters, Frau Helen Schneider, Herr Dr. Thomas Knappich, Frau Birgit Lanz, Herr Werner Kraus, Frau Jenny Kasper, Herr Marius Rebmann, Herr Prof. Dr. Norbert Collmar, Bruder Prof. Dr. Peter Schiffer, Frau Prof. Dr. Sandra Fietkau, Frau Maria Rehm-Kordesee, Frau Franziska Bach, Herr Prof. Dr. Michael Bossle, Herr Prof. Dr. Christian Rester, Frau Prof. Dr. Eveline Wittmann, Herr Dr. Ottmar Döring, Frau Julia Simon, Herr Thomas Schley, Herr Dr. Janne Kleinhans, Herr Philipp Ganßer, Herr Dr. Roman Kaspar, Herr Prof. Dr. Wolfgang Lenhard, Frau Dr. Carolin Hahnel, die Firma AfB gGmbH social & green IT in Paderborn, Herr Fabian Mundt, Frau Angelina Bartz, Frau Prof. Dr. Olga Kunina-Habenicht, Herr Dr. Markus Burkhardt, Frau Beate Kniffke-Koch, Herr Jürgen Spitzlay, Frau Stephanie Pompetzki, Frau Doris Bastian, Frau Tanja Löffler, Herr Dr. Helmut Engelhart, Herr Josef Gabel, Herr Gernot Wamser.

Großer Dank gilt allen Auszubildenden und Studierenden, die sich auf den computergestützten, beruflichen Kompetenztest eingelassen haben und so bereit waren, eine Selbsteinschätzung im Wissens- und Lernprozess zu gewinnen. Erst dadurch war der Einsatz des „Testinstrumentes“ möglich. Gleicher Dank für die Bereitschaft zur Durchführung der Kompetenzmessungen in den einzelnen Schuleinrichtungen gilt den verantwortlichen Schulleitungen und Schulteams bzw. Hochschulleitungen und Hochschulteams.

Ich wünsche mir, dass die erstellte Dissertation allen Schülerinnen und Schülern bzw. Studentinnen und Studenten in der Gesundheits- und Krankenpflege, die sich mit viel Engagement und Lernbereitschaft den hohen und stetig wachsenden Anforderungen im Pflegealltag stellen, eine Möglichkeit zur Unterstützung für die Vorbereitungen auf die praktische Gesundheits- und Krankenpflegeabschlussprüfung ist.

Für die weitere Zukunft sind neue, kompetente Pflegekräfte in der klinischen und ambulanten Praxis für eine „gesunde“ Gesundheits- und Krankenpflege unerlässlich.

Abschließend möchte ich mich bei meiner ganzen Familie für die immerwährende Unterstützung im Rahmen des gesamten Promotionsprojektes recht herzlich bedanken. Hier möchte ich im Besonderen meinen Vater, Jürgen Henning, erwähnen, der mich im „Feld“ bei allen Fahrten und computergestützten Kompetenzmessungen unterstützt und begleitet hat. Gleichermaßen ein herzliches Dankeschön meiner Mutter, Brigitte Henning, die zu Hause die Stellung hielt und über Handy mit uns auf den häufig sehr langen Autofahrten, teils mit Übernachtung, als „Fremdnavigatorin“ aufgetreten ist und für das leibliche Wohl nach der Rückkehr von den Kompetenzmessungen sorgte, meinem Bruder Ralf Henning und seiner Frau Sonja Henning für die vielen guten Worte sowie deren aller Verständnis für meine Aufgabe. Last but not least danke ich meinem besten Freund, strukturgebenden Manager, positiven Motivator, finanziellen Unterstützer und Ehemann Bernhard Ries für das Verständnis, in all den Jahren meinen wissenschaftlichen Weg in der Pflege stets unterstützend und mit großer Geduld begleitet zu haben. Er hat mir dies mit unserer kleinen Familiendackeldame Jule erst möglich gemacht!

Abstract

Hintergrund:

Rapider, paradigmatischer Wandel, erhöhter Personalbedarf und gestiegene Anforderungen verlangen im Bereich der beruflichen Bildung gerade im Sektor Pflege eine klare, transparente Form von Qualitätssicherung und Festlegung praktikabler Mindeststandards. In der groß angelegten „ASCOT“-Studie wurde im Teilprojekt „TEMA“ eine Premiere für die Altenpflegepflegeausbildung gestartet. Es wurde ein technologiebasiertes, computergestütztes Messinstrument zur potentiellen, standardisierten Einschätzung von praktischer Pflegekompetenz, bezogen auf das dritte Ausbildungsjahr, entwickelt. Eine valide Übertragung auf die Gesundheits- und Krankenpflege steht noch aus und wird vom Projektteam aus „TEMA“ grundsätzlich vorgeschlagen. Gerade mit Blick auf die zukünftige generalistische Ausbildung bzw. das Studium der Pflege ab dem Jahr 2020 ist die pflegewissenschaftlich, pädagogisch orientierte Kompetenzforschung weiter voranzutreiben.

Methode:

In toto haben sich 408 Auszubildende und Studierende der Pflege aus 20 Bildungseinrichtungen im süddeutschen Raum an der videogestützten Kompetenzmessung, angelehnt an die „TEMA“-Kriterien, beteiligt. Die statistischen Datenanalysen fanden in personenbezogener und kompetenzbezogener Diagnostik statt. Häufigkeitsermittlungen zu Testitems und Teilnehmern wurden berechnet. Zwei spezifizierte Formen der Item-Response-Theorie wurden eingesetzt, um u.a. die Raschvalidität zu prüfen. Die Feststellung eines geeigneten Strukturgleichungsmodells wurde mit der Konfirmatorischen Faktorenanalyse bestimmt.

Ergebnisse:

Den pflegebasierten Kompetenztest haben nach erfolgtem Scoring mehr als 2/3 der Teilnehmer (= 267 TN) bestanden. Der Mittelwert der erreichten Punktzahl liegt auf genannter Grundlage bei 43,75 Punkten von 96 gescorten Punkten. Die WLE-Reliabilität zeigt sich bei einem akzeptablen Wert von 0.868. Das heuristisch entwickelte Modell deckt inhaltlich die gewünschten, pflegerischen Fähigkeiten in seiner Gesamtheit ab. Eine Raschkonformität lässt sich im „Large-Scale-Assessment“ nicht nachweisen. Die Interpretation der faktoriellen Ergebnisse schließt auf die Bevorzugung eines „eindimensionalen“ Modells zur Messung der praktischen Pflegekompetenz.

Zusammenfassung und Fazit:

Das technologiebasierte, computergestützte Messinstrument stellt sich als geeignet und valide für die Einschätzung der praktischen Pflegekompetenz in der Gesundheits- und Krankenpflege dar. Es zeigt sich aber als anspruchsvoll und herausfordernd für die einzelnen Teilnehmer. Die Endergebnisse aus „TEMA“ werden bestätigt. Unter Berücksichtigung des neuen Pflegeberufegesetzes und des generalistischen Gedankens werden von der Forscherin u.a. die Ausweitung der Pflege-Settings auf „Akutversorgung“, „Pädiatrie“ usw. und eine Überarbeitung einzelner

Fragestellungen auf den interdisziplinären Gesamtkontext vorgeschlagen. Darüber hinaus sollte die Feldforschung auf den gesamten nationalen und internationalen Rahmen ausgeweitet werden, um Vergleichsstudien durchführen zu können. Ein ergänzender Einsatz könnte in der Personalentwicklung unter der Betrachtung des „Lebenslangen Lernens“ von examinierten Pflegekräften geleistet werden.

Abstract

Background:

Rapid, paradigmatic change, increased staff needs and more demanding requirements claim a clear, transparent form of quality assurance and the determining of practicable minimum standards in the area of occupational education, especially in the sector of nursing care. The "TEMA" sub-project, part of the large-scale "ASCOT" study, was a first in geriatric nursing education. A technology-based, computer-supported measuring instrument was developed for a potential, standardised estimate of practical nursing competency pertaining to the third year of study. A valid extrapolation to health care and nursing is still pending and has been categorically suggested by the "TEMA" project team. Especially in view of the prospective generalized nursing education and studies starting in 2020, health care science and educational-oriented competence development is to be further pursued.

Methods:

Overall, 408 nursing students and trainees from 20 educational institutions in Southern Germany participated in the video-supported competency assessment, based on the "TEMA" criteria. The statistical data analyses took place as a personalised and competency-based diagnostic report. Numerical frequency of test items and participants were calculated. Two specified forms of the item-response theory were used to check the Rasch validity, among other things. The establishment of a suitable structure equation model was determined with confirmatory factor analysis.

Results:

After evaluating the scoring, more than 2/3 of the participants (= 267 participants) passed the nursing-based competency test. The average number of points achieved on aforementioned basis was 43.75 points out of 96. The WLE reliability was shown to be of an acceptable value of 0.868. The heuristically developed model covers the total content of the desired nursing skills. Rasch conformity can not be proven in the large-scale-assessment. The interpretation of factorial results suggest the preference of a "one-dimensional" model to measure the practical nursing competency.

Summary and conclusion:

The technology based, computer-supported measuring instrument appears to be suitable and valid for the estimate of practical nursing competency in healthcare and nursing. However, it proves to be demanding and challenging for the individual participants. The final results from "TEMA" were confirmed. Taking into account the new nursing occupation law and the generalistic approach behind it, the researcher suggests among other things an expansion of the nursing setting to include "acute care", "paediatric care" etc. and a revision of individual questions pertaining to the overall interdisciplinary context. Beyond this, the field research should be expanded to a national and international framework in order to conduct comparative studies. A complementary use could be achieved in the field of staff development of registered nursing staff in the context of "lifelong learning".

Résumé

Informations générales:

Dans le domaine de la formation professionnelle, en particulier dans le secteur des soins, le changement rapide et paradigmatique, les besoins en personnel et les exigences accrues demandent une forme claire et transparente d'assurance qualité et la détermination des normes minimales applicables. C'est dans le cadre du projet partiel "TEMA", situé dans l'étude à grande échelle "ASCOT', qu'une nouvelle initiative a été lancée qui s'occupe de la formation professionnelle de soignant(e)s gériatriques. On y a développé un instrument informatisé pour l'évaluation potentielle et standardisée des compétences infirmières pratiques liées au cours de la troisième année de formation. L'équipe de projet «TEMA» envisage de le valider en le reportant sur les soins d'hygiène et les soins infirmiers. Le projet n'a cependant pas encore abouti. Spécialement dans la perspective de la future filière généraliste ou des études de soins infirmiers à partir de 2020, la recherche de compétence à vocation pédagogique et scientifique doit être intensifiée.

Méthode:

Au total, 408 stagiaires et étudiants en soins infirmiers de 20 établissements d'enseignement du sud de l'Allemagne ont participé à une étude basée sur un système de mesure des compétences par vidéo selon les critères du "TEMA". Les analyses de données statistiques ont été réalisées dans le domaine du diagnostic relatif aux personnes et aux compétences. On a déterminé et calculé la fréquence des éléments de test et des participants. Deux formes spécifiques de la théorie "Item-Response" ont été utilisées pour examiner entre autres la validité de la méthode "Rasch". Le constat d'un modèle d'équation structurelle approprié a été utilisé avec le test de confirmation.

Résultats:

Plus de deux tiers des participants (= 267) ont réussi. La moyenne du score obtenu est de 43.75 points sur 96 points à atteindre. La fiabilité WLE montre une valeur acceptable de 0.868. Le modèle développé de manière heuristique couvre le contenu des compétences infirmières souhaitées dans son intégralité. La conformité avec "Rasch" ne peut être démontrée dans le "Large-Scale-Assessment". Les résultats factoriels suggèrent la préférence pour un modèle "unidimensionnel" pour mesurer la compétence infirmière pratique.

Résumé et conclusion:

L'instrument de mesure assisté par ordinateur et basé sur la technologie s'avère approprié et valable pour évaluer la compétence infirmière. Mais il se montre difficile et exigeant pour l'individu. Les résultats définitifs du "TEMA" sont confirmés. Tenant compte de la nouvelle loi des professions de soins et de l'approche généraliste, l'auteur propose p. ex. une extension des "settings" des soins généraux aux "urgences", à la "pédiatrie" etc. ainsi qu'une révision de différentes questions qui concernent le contexte général interdisciplinaire. En outre, le travail sur le terrain devrait

couvrir l'ensemble des programmes nationaux et régionaux pour qu’on puisse mener des études comparatives. Une utilisation complémentaire pourrait être faite dans le développement du personnel compte tenu du fait que les infirmières diplômées sont soumises à une formation permanente.

Inhaltsverzeichnis

Danksagung

Abstract in Deutsch, Englisch und Französisch

Erstgedanken 13

1 Einführung 14
1.1 Hintergrund und Problemstellung 15
1.2 Forschungsfragen einschließlich Hypothesen 20
1.3 Zielsetzung und erhoffter Erkenntnisgewinn 21

I Theoretischer Teil 22

2 Gesetzliche Vorgaben in der praktischen Gesundheits- und Krankenpflegeausbildung 22

3 Kompetenzentwicklung in der praktischen Ausbildung 24
3.1 Definitionsannäherung: Kompetenzen 24
3.2 Erwerb von praktischen Kompetenzen 27
3.3 Messung von praktischen Kompetenzen 28

4 Technologiebasierte Kompetenzmessung in der beruflichen Bildung 30
4.1 „ASCOT“-Studie 30
4.2 Teilprojekt: „TEMA“ 32
4.3 Teilprojekt: „CoSMed“ 40

5 Erkenntnisgewinn aus „TEMA“ mit Einschluss von „CoSMed“ 42

6 Abschließender Kommentar zum „Theoretischen Teil“ 43

II Empirischer Teil 1 - Datenerhebung 44

7 Methodische Vorgehensweise 44
7.1 Vorbereitungen für den Feldzugang 44
7.1.1 Ethische Reflexion 45
7.1.2 Gütekriterien der quantitativen Forschung 46
7.1.3 Bestimmung der Stichprobengröße 48

8 Rahmenbedingungen für den Feldzugang 49
8.1 Kontextuelle Rahmenbedingungen 49
8.2 Technische Rahmenbedingungen 49
8.3 Finanzielle Rahmenbedingungen 49
8.4 Förderungen und Hemmnisse im Rahmen der Projektarbeit im Feld 50

9 Untersuchungsdesign 51
9.1 Einzugsgebiet 51
9.2 Teilnehmende Hochschulen und Bildungseinrichtungen 52
9.3 Ein- und Ausschlusskriterien der Teilnehmer 54
9.4 Zeitlicher Rahmen für die Datengewinnung 55

10 Datengewinnung 56
10.1 Pretests 56
10.2 Haupterhebungsphase und Ablauf der Kompetenzmessungen 56
10.2.1 Einführungsvortrag in die Kompetenzmessung 57
10.2.2 Einverständniserklärung und Datenschutzerklärung 57
10.2.3 Erhebung der personenbezogenen Daten 57
10.2.4 Durchführung der „Technologiebasierten“ Kompetenzmessung 57
10.2.5 Beendigung der Kompetenzmessung vom TN und Forschungsteam 58

11 Abschließender Kommentar zum „Empirischen Teil 1“ - Datenerhebung 59

III Empirischer Teil 2 - Statistische Aufbereitung und Datenanalyse 60

12 Statistische Aufbereitung der gewonnenen Daten 60
12.1 Software: „Open Office Calc“ und „Microsoft EXCEL“ 60
12.2 SPSS und „Missing Values“ 61

13 Verwendete Statistikprogramme für die Datenanalyse 62
13.1 Auswertungen in SPSS 62
13.2 Auswertungen in R bzw. RStudio 62
13.3 Verwendete Packages in R 63

14 Deskriptive statistische Datenauswertung 65
14.1 Personenbezogene Datenanalyse 65
14.2 Kompetenzbezogene Datenanalyse aus „TEMA“ und „CoSMed“ 68
14.3 Weitere statistische Teilergebnisse in SPSS 74

15 Die Item-Response-Theorie 77
15.1 Rasch-Modell 80
15.2 Partial-Credit-Modell 87
15.3 Generalized-Partial-Credit-Model 91
15.4 Ergebnisse aus der IRT-Datenanalyse in R 91

16 Strukturgleichungsmodelle und Konfirmatorische Faktorenanalyse 105
16.1 Dimensionalität der Leistungsdaten 114
16.2 Ergebnisse in R innerhalb der Datenanalyse 115

17 Abschließender Kommentar zum „Empirischen Teil 2“ - Statistische Aufbereitung und Datenanalyse 122

IV Resümee 123

18 Zusammenfassung der Endergebnisse im durchgeführten Projekt einschließlich weiterer Anregungen 123

19 Impulse und Ausblick für weitere Forschungsvorhaben 127

Abschlussgedanken 132
Literaturverzeichnis 133
Abkürzungsverzeichnis 144
Abbildungsverzeichnis 146
Tabellenverzeichnis 147
Mathematisches Formelverzeichnis 148
Anhang 149
Eigenständigkeitserklärung 233

Anmerkungen: Aus Gründen einer kurzen Ausdrucksweise und Vereinfachung ist in der vorliegenden Promotionsarbeit unter der männlichen Anrede auch immer die weibliche miteingeschlossen. Eine Ausnahme bildet der Begriff Promovendin/Forscherin, da die Arbeit von einer weiblichen Person erstellt und geschrieben wurde.

Je nach Kontext wird von Auszubildenden und/oder Studierenden gesprochen. Die DUAL Studierenden an den Hochschulen sind gleichzeitig Auszubildende bzw. Schülerinnen und Schüler in den beruflichen Schulen.

Hinweise: Aufgrund des Urheberrechts im Forschungsprojekt „TEMA“ dürfen die vollständigen Fragestellungen mit Antwortoptionen nicht benannt und aufgeführt werden. Für Transparenz sorgen bei Bedarf die Benennung des Themenkomplexes und die vorliegenden Veröffentlichungen der durchgeführten Literaturanalysen. Darüber hinaus steht der ermittelte Datensatz im vorliegenden Dissertationsprojekt nur der Forscherin und dem Promotionsbetreuer zur Verfügung. Er darf aus datenschutzrechtlichen Gründen für weitere Forschungsbelange nicht weiterverwendet oder an Dritte weitergereicht werden.

Erstgedanken

„O, God. Thy sea is so great and my boat is so small."

(Altes Gebet Bretonischer Fischer; favorisiert von Präsident J.F. Kennedy am 23.05.1963 beim Gedenken an vermisste Personen auf hoher See)

„Der Anfang ist die Hälfte des Ganzen." (Aristoteles)

Formel des RASCH-Modells:

$$p(1 \mid v, i) = \frac{e^{\xi_v - \sigma_i}}{1 + e^{\xi_v - \sigma_i}}$$

$$p(0 \mid v, i) = \frac{1}{1 + e^{\xi_v - \sigma_i}}$$

1 Einführung

Die politischen, gesellschaftlichen, akademischen und ausbildungsbezogenen Veränderungen in der Gesundheits- und Krankenpflege der letzten 20 Jahre und auch in Zukunft (z.B. neues Pflegeberufegesetz - PflBG - mit vollständiger Gültigkeit ab 2020)[1] stellen den Pflegeberuf und seine Angehörigen vor weitere große Herausforderungen. Die Pflegekräfte in der heutigen stationären wie ambulanten Praxis benötigen immer vielfältigere Kompetenzen, um den täglichen Anforderungen des klinischen Alltags gewachsen zu sein. Darüber hinaus verlangen die kaufmännischen Leitungen in Klinikeinrichtungen, ambulanten Einrichtungen, Wohngruppen, Tagespflegeeinrichtungen, Palliativeinrichtungen usw. die ökonomischen Rahmenbedingungen bei jeder Pflegehandlung mit zu bedenken. Zusätzlich ist auf eine transparente Kommunikation und gleichzeitig empathischen Umgang mit dem zu behandelnden Patienten und dessen familiärem Umfeld zu achten. Hier sind noch nicht einmal die vielfältigen Aufgaben und Handlungen innerhalb des interdisziplinären ambulanten wie stationären Teams, um einige Aufgabenschwerpunkte anzusprechen, berücksichtigt. Diesem doch sehr komplexen, pflegerischen Rollenverständnis gerecht zu werden, bedeutet, multiple praktische Kompetenzen auszubauen, wenn nicht vorhanden oder nötig zu erwerben sowie dauerhaft an einem pflegerischen Lehr- und Lernprozess beteiligt zu sein (vgl. Gerlach 2017: 60ff.; vgl. Kälble 2017: 27ff.; vgl. Bensch 2018a: 316ff.; vgl. Bensch 2018b: 321).

Damit bereits in der Gesundheits- und Krankenpflegeausbildung diese Eigenschaften geweckt, entwickelt und weiter gefördert werden können, wäre es notwendig, über die gesetzlichen Rahmenbedingungen und Curricula hinaus bestimmte Mindestkriterien für die zukünftige autonome, praktische berufliche Arbeit am Bett festzulegen. Bisher überprüfen die Lehrkräfte an den Bildungseinrichtungen die praktischen Kompetenzen, je nach Ausbildungsstand, in Form von Praxisanleitungen bzw. Praxisbegleitungen und neuerdings ergänzend im schuleigenen oder klinikeigenen Skills Labs[2]. Am Ende der dreijährigen Gesundheits- und Krankenpflegeausbildung werden in einem von der Fachabteilung der jeweiligen Klinikeinrichtung zur Verfügung gestellten praktischen Kontext die geforderten Kriterien bzw. Kompetenzen durch ein praktisches Examen bzw. eine praktische Prüfung überprüft und festgestellt (vgl. Reibnitz 2008: 36ff.; vgl. Frei Blatter u. Ochsner Oberarzbacher 2008: 113ff.).

Eine auf empirischen Kriterien beruhende, zeitlich begrenzte, pädagogisch etablierte und praxistaugliche Möglichkeit für die Vorbereitung und Einschätzung der praktischen Pflegekompetenzen mit anschließender Festlegung der erreichten Pflegeleistung für das dritte Ausbildungsjahr steht bisher noch aus. Ein geeignetes Hilfsmittel zur Vorbereitung auf die praktischen bzw. alltäglichen pflegerischen Herausforderungen in Form eines standardisierten, technologiebasierten Instruments für die

[1] siehe hierzu Bundesministerium für Gesundheit: „URL: https://www.bundesgesundheitsministerium.de/service/begriffe-von-a-z/p/pflegeberufegesetz.html [Stand: 06.12.2018]“.

[2] „Skills“ werden aus dem Englischen übersetzt als „Fertigkeiten“ und „Lab“ bedeutet ebenfalls aus dem Englischen wörtlich „Labor“ (vgl. Docherty 2005).

Auszubildenden bzw. Pflege DUAL Studierenden[3] fehlt (vgl. Döring et al. 2016a: 243ff.; vgl. Gerlach 2017: 67f. u. 76ff.).

Deshalb wird in der vorliegenden Arbeit nach einem geeigneten und allgemein gültigen, berufsbildenden Messinstrument für die Gesundheits- und Krankenpflege gesucht. Dafür wird exemplarisch ein bereits entwickeltes und wissenschaftlich intensiv aufbereitetes und geprüftes Messinstrument aus dem Verbundprojekt der ASCOT[4]-Initiative des BMBF[5], als ein Meilenstein für ein mögliches „Outputinstrument“[6], genauer untersucht (vgl. Beck et al. 2016b: 10; vgl. Darmann-Finck 2011: o.S.; Deutsches Institut für Internationale Pädagogische Forschung - DIPF 2017). Das aus dem Teilprojekt „TEMA“[7] entwickelte, technologiebasierte, computergestützte Testinstrument für die Auszubildenden in der Altenpflege wird auf die Auszubildenden in der Gesundheits- und Krankenpflege übertragen.

Zunächst werden weitere Aspekte der Notwendigkeit und kritischen Perspektiven von Kompetenzmessungsverfahren in der beruflichen Ausbildung der Gesundheits- und Krankenpflege beleuchtet, und dann wird auf den aktuellen Stand von Kompetenzeinschätzungsverfahren im Ausbildungsverlauf eingegangen.

1.1 Hintergrund und Problemstellung

Die Entwicklung von Kompetenzen passiert zum einen kontinuierlich im täglichen Sein und in allen aktiven Handlungen des täglichen Lebens. Nauerth u. Walkenhorst weisen darauf hin, dass durch die Schaffung von konkreten Lern- bzw. Arbeitssituationen und der Förderung eines positiven Lernumfeldes die Entfaltung von Pflegekompetenzen unterstützt wird. Die Messung kann über objektive und subjektive Messverfahren in Form von Performanztestungen oder Kompetenztestungen erfolgen.

Auf die beiden Begriffe wird in Kapitel 3.1 noch vertieft eingegangen werden. Die Aufbereitung und inhaltliche Benennung innerhalb dieser Testverfahren geschehen u.a. über psychologische, pädagogische, soziologische Sichtweisen. Je nach individueller Kontextualisierung und persönlicher kognitiver Auffassung von der Begrifflichkeit der Kompetenz wird die Wahl des Messverfahrens festgelegt. Da es in der beruflichen Bildung bisher nur sehr wenige geeignete Möglichkeiten der Kompetenzerfassung gibt, sollten weitere empirische Entwicklungen für die Bildungseinrichtungen erfolgen. Aus empirischer Sicht müssen diese kreierten Bewertungsverfahren genau untersucht, ausprobiert, bewertet, überarbeitet und anschließend wieder evaluiert werden, so dass Lehrende und Lernende alternative Möglichkeiten der Einschätzung und Überprüfung im beruflichen Feld der Pflege geschaffen bekommen. Die kritische wissenschaftliche Betrachtung sollte auf dem Focus der

[3] Hinweis: Im weiteren Verlauf der Arbeit wird aus Gründen der Einfachheit häufig „nur“ von Auszubildenden gesprochen, die Pflege DUAL Studierenden sind hier immer miteingeschlossen, außer es wird ausdrücklich darauf hingewiesen.

[4] Technology-based **AS**sessment of Skills and **CO**mpetences in Vocational Education and **T**raining

[5] Bundesministerium für Bildung und Forschung

[6] mit der Frage: wie groß ist der Lernerfolg der Auszubildenden nach abgeschlossener und beruflich orientierter Lehrveranstaltung? (vgl. Beck et al. 2016: 10).

[7] Entwicklung und Erprobung von technologieorientierten Messinstrumenten zur Feststellung der beruflichen Handlungskompetenz in der Pflege älterer Menschen

Zuverlässigkeit und Gültigkeit liegen, mit der Zielsetzung des Erkenntnisgewinns für die Entwicklung von Kompetenzprofilen für die berufliche pflegerische Praxis. Die Pflegekräfte erhalten ein persönliches Kompetenzprofil für die Ausübung ihrer hochkomplexen bzw. komplexen praktischen Tätigkeiten im täglichen Pflegehandeln. Auch der Diskurs zwischen grundständig ausgebildeten Pflegekräften und akademischen Pflegekräften[8] könnte als Möglichkeit für die Beurteilung und Festlegung von Kompetenzprofilen genutzt werden. Lösungsansätze können für die individuelle Betreuung der zu versorgenden Patienten ermittelt werden. Somit geht es nicht nur um Kompetenzentwicklung und Kompetenzerfassung, sondern auch um Kompetenzbewertung und Kompetenzeinschätzung für den sich in der Weiterentwicklung befindenden professionellen Pflegeberuf (vgl. Nauerth u. Walkenhorst 2009: 9ff.; vgl. Gerlach 2017: 76ff.; vgl. Bundesministerium für Gesundheit; Bundesministerium für Familie, Senioren, Frauen und Jugend 2017 u. 2018; vgl. Bundesministerium für Bildung und Forschung 2015).

Kälble weist auf die kontinuierlichen Veränderungen im Pflegeberuf hin, indem er auf den immer komplexer werdenden Pflegebedarf unter Berücksichtigung des demographischen Wandels eingeht. Dadurch ist auch ein erweitertes Qualifizierungsprofil in der Pflege unter Beachtung der ökonomischen Rahmenbedingungen gefragt (vgl. Kälble 2017: 29ff.).

Der Entwicklungsstand von Kompetenzmessverfahren bzw. geeigneten Kompetenzmessinstrumenten in der Pflege auf nationaler Ebene steht noch immer am Anfang. Geeignete Messinstrumente sind häufig nicht empirisch valide genug aufgebaut und eine Verallgemeinerung lässt sich nicht umsetzen, um für die praktische Pflegeausbildung eine Allgemeingültigkeit festlegen zu können. Deshalb ist es an der Zeit, neue und geeignete Messinstrumente zu entwerfen und auf deren Gültigkeit für die reale Praxis zu untersuchen[9]. Die „Erprobung“ auf Eignung dieser entwickelten Testinstrumente darf somit nicht auf theoretischer Ebene enden, sondern muss in der pflegepädagogischen Praxis direkt in den Gesundheits- und Krankenpflegeschulen an den Auszubildenden in der Gesundheits- und Krankenpflege stattfinden. Dies bedeutet, dass im Bereich der Kompetenzdiagnostik die Messungen von praktischen Pflegekompetenzen einen hohen Stellenwert für ein ergebnisorientiertes Denken haben (vgl. Darmann-Finck, Reuschenbach 2013: 27f.; vgl. Darmann-Finck, Glissmann 2011: 198ff.).

Darmann-Finck u. Reuschenbach weisen in Ihrem Fachartikel ausdrücklich auf das TEMA-Projekt hin:

> „Die systematische, theoretisch und empirisch abgesicherte Entwicklung von Kompetenzdiagnostika befindet sich noch in den Anfängen, Fortschritte sind von dem BMBF-geförderten Verbundprojekt „Entwicklung und Erprobung von technologie-orientierten Messinstrumenten zur Feststellung der beruflichen Handlungskompetenz in der Pflege älterer Menschen“ (TEMA) zu erwarten.“ (Darmann-Finck, Reuschenbach 2013: 27).

[8] hier sind DUAL Studierende im Bereich Pflege gemeint

[9] Darmann-Finck u. Reuschenbach weisen auf die „Schlüsselrolle hinsichtlich der Verlaufskontrolle und Überprüfung der Effektivität und Effizienz von Bildungsprogrammen“ hin (Darmann-Fink, Reuschenbach 2013: 23).

Geeignete Modelle für die Schaffung neuer Testinstrumente in der Pflege, die die Wechselwirkung von vorgeschriebenen Leistungsanforderungen und erlernten Kompetenzen berücksichtigen, fehlen. Sehr gute praktische Lernergebnisse zeigen sich nicht alleine in einem hohen Kompetenzfaktor, sondern in der umfassenden, individuellen Versorgung des zu betreuenden Patienten. Somit stehen Bildung und Kompetenz in Abhängigkeit voneinander und sind bei der Entwicklung von Kompetenzmessmodellen zu berücksichtigen (vgl. Darmann-Finck, Reuschenbach 2013: 28).

Bei der methodischen Annäherung an Kompetenzmessverfahren wurden in der Vergangenheit computergestützte, technologiebasierte Messinstrumente so gut wie gar nicht entwickelt, in der Praxis versucht und anschließend statistisch evaluiert. Die Entwicklung solcher Kompetenztestungen verlangt ein hohes Expertenwissen, welches die Verknüpfung von entwickelten Testitems und anschließender Auswertung der erhobenen Testergebnisse unter Zuhilfenahme von statistischen Methoden aus dem Bereich der Item-Response-Theorie, wie beispielsweise in den PISA[10]-Testungen, verlangt (vgl. Darmann-Finck, Reuschenbach 2013: 28; vgl. OECD 2018; vgl. Hermeneit 2008: 11).

Ein weiteres Problem zeigt sich nach Abschluss der Gesundheits- und Krankenpflegeausbildung, wenn die frisch examinierten Pflegekräfte bzw. ehemaligen Auszubildenden den Anforderungen in der beruflichen pflegerischen Praxis nicht gewachsen sind und negative Rückmeldungen an die Ausbildungsstätten über deren Inkompetenz im täglichen praktischen Krankenhausalltag herangetragen werden. Als Lösung wird versucht, ein mögliches Angebot oder die Schaffung von Überprüfungsmöglichkeiten mittels temporärer, computergestützter Simulationen zu schaffen. Diese Entwicklung steht für den pflegerischen Bereich noch am Anfang.[11] (vgl. Darmann-Finck, Reuschenbach 2013: 28; vgl. Funke, Reuschenbach 2011: 604ff.).

Das Erwerben von neuem Wissen der pflegerischen Auszubildenden geschieht nicht einfach ad hoc, sondern u.a. durch diverse Lernangebote. Diese Lernangebote helfen einen Blickwinkel auf eine Fragestellung zu entwickeln und nach handhabbaren Lösungen im komplexen Pflegealltag zu suchen. Damit es aber überhaupt zum konstruktiven Weiterdenken von Auszubildenden kommen kann, müssen von den Lehrenden verschiedene bildungsentwickelnde Faktoren mitberücksichtigt werden, wie beispielsweise die einzelnen Auszubildenden in ihrem jeweiligen Gesamtkontext oder auch der Gruppendynamik in Lehr- und Lernsituationen untereinander (Klassengefüge). Kompetenzen entwickeln sich und sind nicht steuerbar sofort zu erwerben und kompetent umzusetzen, sondern die pädagogische, gestalterische Hinführung auf neue Handlungsoptionen muss bei der Erweiterung von theoretischem und praktischem Wissen mitberücksichtigt werden.

---

[10] Programme for International Student Assessment bzw. Programm zur internationalen Schülerbewertung von 15-jährigen Schülern gegen Ende ihrer Schulpflichtzeit, die auf Initiative der OECD (Organisation für wirtschaftliche Zusammenarbeit und Entwicklung) durchgeführt wurde und in regelmäßigen Abständen wiederholt wird.

[11] siehe hierzu die wissenschaftlichen Aufbereitungen bzw. Entwicklungen von Reuschenbach in der computergestützten Simulation von „NursePlan“ (vgl. Funke, Reuschenbach 2011: 605f.; vgl. Reuschenbach 2008: 223ff.).

Darüber hinaus bestand in der Vergangenheit das Problem, bei entwickelten Messinstrumenten zur Kompetenzüberprüfung die Inhalte auf Nachweisbarkeit von theoretischem oder praktischem Wissen empirisch darzustellen. Eine einfach zu verstehende Gestaltung von Items und deren Antwortmustern erwies sich als erforderlich. Zielsetzung war, das Ablegen von anstehenden Prüfungsteilen im schriftlichen, praktischen oder mündlichen Teil in der Pflege zu bewerten. Die Gefahr bestand bisher, dass sich der Blick fast ausschließlich auf die fachlich-technische Seite richtete. Es wurden Punkte, wie beispielsweise der persönlichen, teamorientierten und kommunikativen Seite, unberücksichtigt gelassen. Anzumerken ist, dass diese nur schwer messbar gemacht werden können (vgl. Darmann-Finck, Reuschenbach 2013: 28; vgl. Wittmann et al. 2015: 360ff.; vgl. Reuschenbach 2008: 17ff.).

Das Voranbringen und Verbessern von Kompetenzmessinstrumenten sollte aber wegen der zuvor angefügten Punkte nicht beendet werden, sondern, wie auch Darmann-Finck und Reuschenbach anmerken,

> „… der Absicherung von Qualitätsmindeststandards und dem Vergleich von Bildungsangeboten auf abstraktem Niveau dienen können. Darüber hinaus sind entsprechende Messinstrumente natürlich auch für die Forschung unabdingbar.“ (Darmann-Finck, Reuschenbach 2013: 28).

Als Komplettlösung zur Evaluation einzelner, abgehaltener, theoretischer und praktischer Unterrichtsszenarien für Lehrende und Studierende wird dem sicher nicht Rechnung getragen werden. Wiederum für die Gesamteinschätzung pflegerisch erworbenen Wissens unter der Berücksichtigung des temporären Faktors, wie beispielsweise nach den jeweiligen Ausbildungsabschnitten bzw. Modulabschnitten, kann es sich als durchaus sinnvoll erweisen und als Einschätzungsinstrument von erworbenen Pflegekompetenzen herangezogen werden (vgl. Funke, Reuschenbach 2011: 602ff.).

Ein weiterer Blickwinkel bezüglich des Kompetenzgedankens zeigte sich für die Forscherin bereits im Vorfeld des Projektstarts in den durchgeführten leitfadengestützten, episodischen Interviews einschließlich beinhalteter Struktur-Lege-Technik in der erstellten Masterarbeit zum Thema „Lehrergesundheit an Gesundheits- und Krankenpflegeschulen“ und den zuvor durchgeführten Forschungspraktika. Lehrkräfte beklagen des Öfteren die starken Veränderungen der Schülerpersönlichkeit in Richtung Funktionalismus und reduzierter empathisch-kommunikativer Kompetenz im Pflegealltag, allgemein gesprochen, eine große Kluft zwischen hochkompetenten und nicht geeigneten Auszubildenden in der Pflege. Durch das Fehlen von klaren gesetzlichen Regelungen, Mindeststandards und regelmäßigen schriftlichen, praktischen und mündlichen Überprüfungen wird die Entwicklung eines individuellen Schüler-Kompetenzprofils erschwert. Eine Behebung von Defiziten in bestimmten Kompetenzbereichen kann von den Lehrenden und Praxisanleitern an den Ausbildungsstätten aufgrund der engen zeitlichen Ressourcen nur schwierig an die täglichen praktischen Pflegeanforderungen angeglichen werden (vgl. Ries 2014b: 62, 68 u. 105-107; vgl. Ries 2014a: 4).

Eine Orientierung, Prägung und das Erleben einer sozial- und individualorientierten Welt wird den Lehrenden durch die ökonomischen, politischen und gesellschaftlichen Rahmenbedingungen erschwert mit gleichzeitig zu vermittelnder starker Theorielastigkeit in der Ausbildung in der Gesundheits- und Krankenpflege. Andererseits sind Kompetenzentwicklungen bzw. Kompetenzveränderungen bei den Auszubildenden nur schwer steuerbar, entwickelbar und veränderbar, da meist die zeitlichen Ressourcen für den komplexen Lernstoff im Schulalltag fehlen. Die Festlegung eines Mindeststandards für die praktische Pflegeausbildung ist wünschenswert. Eine Unterstützung und gleichzeitig Hilfsmittel können geeignete Messinstrumente für die Überprüfung, Bewertung und Einschätzung von erlernten Kompetenzen sein. Bisher stehen den Lernenden sowie auch den Lehrenden an Gesundheits- und Krankenpflegeschulen diese nicht in geeignetem Umfang zur Verfügung (vgl. auch Ries 2014b: 68, 77 u. 75).

Sinnvoll ist bei der gesamten Auseinandersetzung mit der Entwicklung, Verwendung und empirischen Evaluation eines Kompetenzmessinstrumentes, sprich der gesamten komplexen Kompetenzdiagnostik, dass der Focus nicht nur aus dem pflegerischen Kontext geleitet wird, sondern auf Interdisziplinarität mit anderen angrenzenden Professionen beruht, die ihre Erkenntnisse und Erfahrungen im Bereich der Kompetenzermittlung mit in die wissenschaftliche Perspektivenbetrachtung einfließen lassen (vgl. Darmann-Finck, Grissmann 2011: 202).

Als weitere Erschwernis zeigt sich die Diversität der einzelnen Schulcurricula, falls auf Länderebene keine allgemeingültige Entwicklung erfolgt ist. Die daraus resultierenden individuellen, schulinternen Festlegungen und Vorstellungen von zu entwickelnden praktischen Pflegekompetenzbereichen bei Auszubildenden in der Gesundheits- und Krankenpflege erschweren den Entwurf eines für alle Schuleinrichtungen gültigen und zur Verfügung stehenden standardisierten, technologiebasierten Messinstruments. In der curricularen Entwicklung bzw. Überarbeitung fließen essentielle pflegerische Handlungskompetenzen mit ein. Die angeführten Aspekte sind gleichermaßen in der Entstehung von zu entwerfenden Hochschulcurricula auf Bundesebene zu beobachten, wie beispielsweise im Bereich der Pflegewissenschaft. Gleichzeitig wird inhaltlich nach festgelegten Schlüsselkompetenzen gesucht, die die immer rasanteren Veränderungen und notwendigen Anpassungen im täglichen pflegerischen Praxisalltag für die zukünftige examinierte Pflegekraft vorbereiten und erleichtern. Ein Problem der Vermittlung der genannten Schlüsselkompetenzen liegt in einer häufig vorhandenen, reduzierten Grundlagenkompetenz der Auszubildenden in der Gesundheits- und Krankenpflege (vgl. Wittmann et al. 2014; vgl. Schön 2011: 79ff.; vgl. Dietze 2011: 133 u. 144; vgl. Hülsken-Gießler et al. 2010: 216ff.; vgl. Döring et al. 2014: 36f.; vgl. Evers 2011: 49ff.; Döring, Möllers, Schöpf 2014: 34ff.).

In Anbetracht der hier angeführten Punkte bietet sich ein weiterer intensiver, wissenschaftlicher Blick auf das bereits zuvor mehrfach erwähnte Messinstrument des BMBF aus der ASCOT-Studie im Teilprojekt „TEMA“ an. Es nimmt inhaltlichen Bezug auf die erwähnten Aspekte und versucht, diese in einem zuvor entwickelten technologiebasierten, computergestützten Messinstrument in den Berufsschulen für

Altenpflege einzusetzen und auf deren Eignung zu prüfen, anschließend folgt die statistische Auswertung der ermittelten Ergebnisse mit Prüfung der Modellpassung und der Eignungsfeststellung des Frage-Antwort-Verhaltens der einzelnen Items (vgl. u.a. Wittmann, Kaspar, Döring 2017: 185ff.; vgl. Kaspar et al. 2016: 185ff.; vgl. Kaspar, Hartig 2016: 105ff.; vgl. Simon et al. 2015; vgl. Wittmann et al. 2014: 53ff.; Wittmann et al. 2015).

Könnte dieses Messinstrument ein Hilfsmittel auch für die Gesundheits- und Krankenpflege sein oder darüber hinaus für die zukünftige Generalistik in der Pflege?

Die weitere Analyse und Aufbereitung des sehr komplexen TEMA-Projekts erfolgt ab Kapitel 4 in der wissenschaftlichen Untersuchung.

Zunächst wird für die weiteren Betrachtungen die Fragestellung einschließlich der Hypothesenformulierung festgelegt.

1.2 Forschungsfragen einschließlich Hypothesen

Aus den in Kapitel 1.1 erläuterten Aspekten ergeben sich aus Sicht der Forscherin mehrere empirische Fragestellungen, die im weiteren Verlauf der Arbeit einer wissenschaftlichen Beantwortung überführt werden.

Eine der resultierenden, zentralen Forschungsfragen im Bereich der weiterführenden Analyse der Kompetenztestentwicklung, Kompetenztestmessung im Feld und anschließenden Kompetenztestungsevaluation unter Bezugnahme auf das Ursprungsprojekt „TEMA“ lautet:

Ist der bereits wissenschaftlich entwickelte, technologiebasierte, computergestützte Pflegekompetenztest aus dem Setting der Altenpflege übertragbar in die Gesundheits- und Krankenpflege?[12]

- Weitere Subfragestellungen:

Zeigen sich vergleichbare Gesamtergebnisse in der Gesundheits- und Krankenpflege wie in der Altenpflege?

Welche Hintergründe und Verknüpfungspunkte sind im Kompetenztest zu erkennen?

Wie ist das technologische, computergestützte Testinstrument aus „TEMA“ in Bezug auf Praktikabilität (Handhabung des Messinstrumentes) einzustufen?

Wird über die Messung der entwickelten, einzelnen Items eine Feststellung der praktischen Kompetenz[13] in der Pflege überhaupt möglich?

[12] Bisher liegt nur eine Relevanzeinschätzung vor (vgl. Döring et al. 2016b: 128).
[13] „berufliche Handlungskompetenz“ gemäß Döring et al. 2016b: 117

- Weitere statistische Fragestellungen:

Welche Häufigkeiten sind in den „TEMA“ und „CosMed“ technologiebasierten, computergestützten Testergebnissen zu erkennen?

Ist die Item-Response-Theorie ein geeignetes Auswertungsverfahren für den entwickelten bzw. im Feld eingesetzten Kompetenztest?

Liegt Raschvalidität vor?

H_0-Hypothese: Es liegt Raschvalidität vor.
H_1-Hypothese: Es liegt keine Raschvalidität vor.

Welches Messmodell ist im Rahmen der Konfirmatorischen Faktorenanalyse[14] zu bevorzugen?

Bezüglich der Konfirmatorischen Faktorenanalyse:

H_0-Hypothese: Das ermittelte Modell und die vorhandenen Daten passen zusammen.

H_1-Hypothese: Das ermittelte Modell und die vorhandenen Daten passen nicht zusammen.

1.3 Zielsetzung und erhoffter Erkenntnisgewinn

Mit den Fragestellungen ist beabsichtigt,

- eine evidenzbasierte Übertragung des computergestützten Kompetenztests (Messinstruments) in die Gesundheits- und Krankenpflege zu erreichen,
- ein möglichst breites Teilnehmerfeld von Auszubildenden aus der Gesundheits- und Krankenpflege zu erreichen und gleichzeitig ein breites Bild von Schuleinrichtungen aus der Gesundheits- und Krankenpflege zu erhalten,
- das Testinstrument kennen zu lernen und auf seine statistische Gültigkeit in Bezug auf Modellgüte und Frage-Antwort-Verhalten sowie Itemhomogenität zu prüfen,
- die zentralen statistischen Ergebnisse aus den durchgeführten Kompetenzmessungen der Altenpflege mit denen der Gesundheits- und Krankenpflege in Beziehung zu setzen und zu vergleichen,
- die Anwendung und Eignung des „TEMA“-Testinstrumentes für die Schüler und die Studierenden in der Gesundheits- und Krankenpflege festzustellen,
- die Kompetenzdiagnostik und den Einsatz von unterstützenden Hilfsmitteln zur beruflichen Handlungskompetenzmessung im Berufsfeld Pflege mit voranzutreiben sowie über deren Existenz aufzuklären (vgl. auch Döring et al. 2016b: 117f. u. 127f.).

[14] Abkürzung „CFA“, weitere Erläuterungen und Vertiefungen folgen im Empirischen Teil II - statistische Aufbereitung und Datenanalyse

I Theoretischer Teil

Im folgenden theoretischen Hintergrund wird zunächst auf die gesetzlichen kompetenzrelevanten Aspekte innerhalb der Gesundheits- und Krankenpflegeausbildung eingegangen. Anschließend wird die Begrifflichkeit Kompetenz dargestellt und eine ausführliche Betrachtung des TEMA-Projektes vorgenommen. Zum Ende des theoretischen Teils I wird ein Fazit gezogen und eine Überführung auf das neu anstehende Projekt vollzogen.

2 Gesetzliche Vorgaben in der praktischen Gesundheits- und Krankenpflegeausbildung

Zu Beginn des Forschungsvorhabens im Jahr 2016 war das Krankenpflegegesetz[15] einschließlich der Ausbildungs- und Prüfungsverordnung für die Berufe in der Krankenpflege[16] von 2003 noch in vollem Umfang gültig. Die Untersuchungen innerhalb des gesamten nun folgenden Kompetenzprojekts einschließlich der Eignung des vom Projekt TEMA entwickelten Messinstrumentes bezüglich der konstruierten Items, Antwortoptionen, Befragungen in den Hochschul- und Schuleinrichtungen, Auswertung der erhobenen Daten usw. fand entsprechend der zu diesem Zeitpunkt gültigen Gesetzgebung in der Gesundheits- und Krankenpflege statt.

Das Krankenpflegegesetz von 2003 zeigt im § 1 Berufsbezeichnung den Kompetenzbegriff in Zusammenhang mit Abschnitt (1) den Begriff der „vermittelten erweiterten Kompetenzen“ auf. Die praktische Ausbildung gliedert sich in einen „allgemeinen“ Bereich und einen „differenzierten“ Bereich. In der Ausbildungs- und Prüfungsverordnung § 15 Praktischer Teil der Prüfung wird auf die Rahmenorganisation des Prüfungsablaufs eingegangen. Es ist die Versorgung einer Patientengruppe aus einem der drei Differenzierungsbereiche der Inneren Medizin, Chirurgie oder Psychiatrie vollständig zu übernehmen. Bei den gewählten Patienten sind häufig Pflegeprobleme aus dem Bereich der Gerontologie mit zu berücksichtigen, wie u.a. beispielsweise verschiedenste Formen der Demenz oder der Osteoporose. Gesamtheitlich betrachtet liegt in den klinischen sowie ambulanten Kontexten vermehrt Multimorbidität[17] vor. Es wird in § 15 (1) darauf hingewiesen, dass der Nachweis über die erworbenen praktischen Kompetenzen im gesamten Ausbildungszeitraum zu erbringen ist. Es wird hier nicht vertiefend auf spezielle Kompetenzen eingegangen, sondern auf das Krankenpflegegesetz § 3 (1) verwiesen. Hier wird innerhalb der Ausbildungszielbenennung konkretisiert von „fachlichen, personalen, sozialen und methodischen Kompetenzen“, wie z.B. zur Bewältigung von Krankheit gesprochen. Was unter den einzelnen Kompetenzbereichen zu verstehen ist, wird nicht genauer erläutert. Es wird weiter auf die Aufgabenbereiche eingegangen, in denen die Kompetenzen benötigt werden (vgl. Bundesministerium der Justiz und Verbraucherschutz, Bundesamt für Justiz 2003a u. b - KrPflG u. KrPflAPrV Anlage 1 zu § 1 Abs. 1, A Theoretischer und praktischer Unterricht und B Praktische Ausbildung).

[15] Abkürzung „KrPflG“
[16] Abkürzung „KrPflAPrV“
[17] Vielfacherkrankungen, wie z.B. Hypertonie, Diabetes mellitus, Hypercholesterinämie, usw.

Weitere Eingrenzungen oder Benennungen zum Kompetenzerwerb werden in der Gesetzgebung von 2003 nicht gefunden, allerdings ist auf eine gravierende gesetzliche Veränderung in diesem Zusammenhang noch hinzuweisen.

Das Besondere im Rahmen der Gesetzgebung in der Bundesrepublik Deutschland ist, dass in der Zeit der Erstellung dieser Arbeit (Oktober 2016 bis Juli 2019) ein neues Gesetz in der Pflege verabschiedet wurde. Das neue Pflegeberufegesetz bzw. Pflegeberufereformgesetz wurde im Juli 2017 verabschiedet, die Ausbildungs- und Prüfungsverordnung folgte im Oktober 2018. Beide sind ab dem Jahr 2020 vollständig an den gesamten Schul- und Hochschuleinrichtungen umzusetzen.

Als Anmerkung sei hier erlaubt, dass bei der Durchsicht der zukünftigen Gesetzgebung der Begriff „Kompetenz“ durchgängig zu finden ist. Somit lässt sich vermuten, dass das nun aufzuzeigende Projekt nicht der Einmaligkeit unterliegt, sondern die Perspektiven eines zukunftsfähigen, gemeinsam gestalteten Pflegeberufs mit technologiebasierten, computergestützten Hilfsmitteln bereichert und unterstützt.

Die Begründungen dieser Annahme zeigen sich beispielsweise innerhalb der neuen Gesetzgebung darin, dass das neue Pflegeberufegesetz die drei Ausbildungsberufe der Gesundheits- und Krankenpflege, Gesundheits- und Kinderkrankenpflege und Altenpflege verbindet. Es entsteht gemeinsame Kompetenz innerhalb eines Berufsfeldes, das sich über den gesamten Lebenslauf eines Menschen erstreckt. Die neuen Gesetzesbeschlüsse bzw. Verordnungen sehen u.a. Kompetenzbereiche vor, die beispielsweise durch hochschuleigene Curricula ergänzt werden können. Themenbereiche bzw. Lernfelder werden durch den Begriff Kompetenzbereiche ersetzt und inhaltlich darauf ausgerichtet. Im PflBG wird im § 5, § 17, § 37 und § 39 und in der KrPflAPrV im § 1, § 2, § 3, § 6, § 9, § 13, § 14, § 30, § 31, § 32, § 35, § 36 und § 37 über den komplexen Bereich von Kompetenzen in der gesamten Pflegeausbildung gesprochen, wie z.B. von Kompetenzvermittlung und Kompetenzentwicklung, sozialen und personalen Kompetenzen und in der praktischen Prüfung von diversen Kompetenzbereichen. Somit erhält die Begrifflichkeit von Kompetenzen in der Pflege eine neue, aktuelle und zentrale Bedeutung in der zukünftigen theoretischen, mündlichen sowie praktischen Pflegeausbildung. Es entsteht eine völlig neue Form von Pflegequalifikation durch verschiedenste Kompetenzfassetten (vgl. auch BfG, BMFSFJ 2017 u. 2018 zu § 9 Absatz 1 Satz 2 Anlage 2, zu § 35 Absatz 2, § 36 Absatz 1 und § 37 Absatz 1 Anlage 5 u. zu § 1 Absatz 1 Nummer 1, § 25 Anlage 6; BfG 2018; vgl. Kerngruppe Curriculum 2006: 61 u. 82ff.; vgl. Darmann-Finck, Reuschenbach 2013: 23ff.; vgl. Knigge-Demal, Eylmann 2006: 8ff.; Hundenborn 2016: 46).

Die Bedeutung in gesetzlicher und zusätzlich noch curricularer Hinsicht für die Entwicklung eines standardisierten Kompetenzmessinstrumentes wird in Kapitel 4 noch näher erläutert werden.

3 Kompetenzentwicklung in der praktischen Ausbildung

Die Entwicklung von Kompetenzen für den beruflichen Pflegealltag entsteht nach Schlegel durch Wissen und den Erwerb von Fähigkeiten für die Bewältigung von Alltagssituationen. Zu berücksichtigen ist, dass z.B. akute Notfälle im klinischen Alltag einer anderen Handlungskompetenz bedürfen als der „klassische“ stationäre internistische Patient mit Diabetes Mellitus Typ II. Der Auszubildende benötigt deshalb innerhalb seiner Ausbildung die Möglichkeit, verschiedene Kontexte kennen zu lernen und die notwendigen Pflegehandlungen einschließlich der Fähigkeit von kommunikativer und psychologischer Bewältigung in den Grundzügen zu erwerben (vgl. Schlegel 2018: VII; vgl. Bachmann u. Schlegel 2018: 349; vgl. Holoch 2007: 383f.; vgl. Brater 2016: 209f.).

Zur Annäherung und Vorstellung über die Begrifflichkeit von Kompetenz soll zunächst in Kapitel 3.1 eine erweiterte Betrachtung stattfinden.

3.1 Definitionsannäherung: Kompetenzen

Bei der literarischen Betrachtung des Kompetenzbegriffes ist zunächst festzustellen, dass sich keine eindeutige Definition von Kompetenz festlegen lässt. Es zeigen sich sehr divergente Bezüge der Begrifflichkeit und darüber hinaus. Innerhalb dieses Kapitels sollen deshalb einige Facetten der Begrifflichkeit von Kompetenz aufgegriffen und mögliche Sichtweisen betrachtet werden.

Ganz allgemein ist unter „Kompetenz“ Vermögen, Fähigkeit, Sachverstand, Zuständigkeit oder Befugnis zu verstehen (vgl. Dudenredaktion 2015; vgl. Brockhaus 1999).

Im pflegewissenschaftlichen Kontext spricht Bartholomeyczik von dem Begriff der „professionellen Kompetenz in der Pflege“ und meint damit:

> „…die komplexen Fähigkeiten, die eine Professionsangehörige besitzen muss, um ihren Beruf bestmöglich ausüben zu können.“ (Bartholomeyczik 2001a: 285).

Darüber hinaus weist sie, bezogen auf die notwendigen pflegerischen Kompetenzen, auf die Entwicklung von professionsbezogenen Eigenschaften mit ausreichendem, distanziertem Analyseverhalten hin. Es gilt ein Gleichgewicht in der Betreuung der anvertrauten Patienten zu erzielen, nicht nur bezogen auf fachlich-technische Pflege, sondern u.a. auch auf der psychologischen und psycho-sozialen Ebene, wie beispielsweise der professionelle Umgang mit emotionalem Verhalten von zu Pflegenden. Dies alles verlangt eine Form von Kompetenz. Gleichzeitig ist darüber nachzudenken, welche Kompetenzen in der Pflege verfolgt werden sollten bzw. sich anzueignen sind, um von einer „kompetenten“ Pflegekraft bzw. „kompetenter“ Pflege sprechen zu können (vgl. Bartholomeyczik 2001a: 286f.). Hier weist sie auf die hohe Bedeutung von einer umfassenden, individuellen Prozesssteuerung gepaart mit hohem Fachwissen in der Pflege hin, die eine systemische, logische und verständliche Struktur aufweist (vgl. Bartholomeyczik 2001b: 344ff.). Dafür sind „Schwerpunktsetzungen“ in diversen Pflegekompetenzbereichen notwendig.

Dies lässt Bartholomeyczik zu dem Schluss kommen, dass „kompetente“ Pflege nach folgendem Konstrukt entstehen kann:

> „Zu einer kompetenten Pflege gehören neben der direkten Pflege unter anderem auch die Beratung und Anleitung der PatientInnen und ihrer Angehörigen, Case und Trajekt[18]-Management, die Umsetzung wissenschaftlicher Forschung sowie soziale Kompetenz im weitesten Sinne.“ (Bartholomeyczik 2001c: 412).

Im pflegewissenschaftlichen Sinn ist dies u.a. auch so zu verstehen, dass Pflegekräfte immer wieder einen individuellen Perspektivenwechsel eingehen können, um sich als Pflegekraft in den zu Pflegenden hineindenken zu können, dabei aber trotzdem professionell für den Patienten entscheiden.

Benner bringt dagegen nicht nur den Begriff der Kompetenz in die Definition mit ein, sondern stellt eine Verbindung zu dem Begriff der Performanz her:

> „Bedeutsam für die Patientenergebnisse ist somit nicht die pflegerische „Kompetenz“ sondern die „Performanz“ („Leistung“). „Kompetenz“ bezeichnet nur die Möglichkeit, etwas zu leisten „potential to perform“ nicht die tatsächliche Leistung. … Somit gilt es „Kompetenz in Performance zu übersetzen“.“ (Benner 2012: 38-39).

Sie betont die Breite des praktischen und pflegerischen Handelns, das von vielfältigen Kompetenzen geprägt ist, wie beispielsweise Pflegetechnik, Kommunikation, usw. Gleichzeitig wird auf die Schwierigkeit der Messung von Kompetenzen aufgrund der Vielfältigkeit hingewiesen. Die bisherigen Messinstrumente zeigen hier nur eine reduzierte empirische Güte (vgl. Benner 2012: 38f.)[19].

Im Gegensatz zu Benner erklärt Weidner nicht den Kompetenzbegriff, sondern greift das Verständnis der „Konstitutiven Kompetenzen“ nach Raven (ursprünglich entwickelt für professionelles Handeln im medizinischen Kontext unter Berücksichtigung des Verständnisses nach Oevermann mit dem Focus der „therapeutischen Kompetenz des Fallverstehens“) auf. Er überträgt das Kompetenzkonstrukt auf notwendiges professionelles Pflegehandeln im Praxisalltag. Dieses besteht u.a. aus pflegewissenschaftlicher Begründungskompetenz sowie Entscheidungs- und Handlungskompetenz (vgl. Weidner 2011: 124ff.; vgl. Raven 1989: 38-41 u. 62-65).

Sauter u. Staudt wiederum greifen nicht nur den Begriff der Kompetenz auf, sondern verbinden diesen mit der Einschätzung von Kompetenz:

> „Kompetenzen, die Fähigkeit, Problemstellungen selbstorganisiert und kreativ zu lösen, schlagen sich immer in Handlungen nieder. Sie sind keine Persönlichkeitseigenschaften. Kompetenzmessungen schließen deshalb vom aktuellen Handeln auf vorhandene Handlungsfähigkeiten, nicht auf verborgene Persönlichkeitsmerkmale.“ (Sauter, Staudt 2016: VII).

---

[18] „Überfahrt“, in der Pflege als „Überleitungsmanagement“ oder „Überleitungspflege“ bezeichnet

[19] Anmerkung: Bereits 1970 arbeitet Chomsky im Bereich der Linguistik heraus, dass zwischen den Begriffen der „Sprachkompetenz“ und der „Sprachperformance“ zu unterscheiden ist. Erst dann existierte im linguistischen Lernprozess ein Verständnis von sprachlicher Kompetenz! (vgl. Chomsky 1970: 13-14; vgl. Chomsky 2017: 120-121).

Nohl greift aus pädagogischer Sichtweise die Komplexität des Kompetenzbegriffes bezugnehmend auf den Deutschen Bildungsrat auf.

> „Kompetenz umschreibt damit die Auswirkungen des Lernens auf die Person.“ (Nohl 2009: 74).

Ein Hinweis auf eine mögliche Einordnung des hochkomplexen Kompetenzbegriffes zeigt sich in der Möglichkeit von der Festlegung von vier „Definitionsclustern“. Diese umfassen: Bildung, Handlung, Denken und Innovation bezogen auf die eigene Handlungsfähigkeit (vgl. Erpenbeck, Grote, Sauter 2017: XXff.).

Entsprechend dem psychologischen Verständnis der Kompetenzbegrifflichkeit zeigt Bandura anhand des „Lernens am Modell“, das in vier Phasen eingeteilt ist, dass in der ersten Phase, der „Aneignungsphase“, der Kompetenzerwerb durch u.a. Lernen von anderen Personen in Form von Nachahmung von Verhaltensmustern (z.B. Lehrenden) und späterer kognitiver Abspeicherung dieser Lernprozesse bzw. Handlungen, wie sich Kompetenz entwickelt, auf (vgl. Bandura, Verres, Kober 1979: 31ff.; vgl. Bandura, Kober 1976: 48ff.).

Schlegel spricht im Rahmen der Lerntheorie nach Bandura und seiner Definition und Vorstellung von Kompetenz:

> „Nach der Definition von Bandura ist Kompetenz, einfach ausgedrückt, die Fähigkeit, eine Handlung einem Kontext anzupassen.“ (Schlegel im Interview mit Bachmann 2018: 349).

In der ASCOT-Studie, die im Kapitel 4.1 noch näher erläutert wird, findet sich Kompetenz bezugnehmend auf PISA in der Begrifflichkeit der „komplexen Handlungskompetenz“ und der „Erhebung von berufsfachlichen Kompetenzen“ wieder. Diese besteht aus drei Facetten: Kognition, Motivation und Emotion. Auch hier wird die Diversität der Begrifflichkeit erwähnt und deshalb eingegrenzt auf den Bereich der „kognitionspsychologischen“ Sichtweise. Durch dieses Verständnis sollen Resultate beim Auszubildenden erzielt werden, die die Bandbreite des zu erlernenden Berufes, wie z.B. des Pflegeberufs, abbilden (vgl. Döring et al. 2016a: 245; vgl. Baethge, Seeber 2016: 16).

Aus den hier aufgeführten Begriffsfacetten lässt sich schlussfolgern, dass der Begriff von „Kompetenz“ zum heutigen Zeitpunkt nicht eindeutig definiert und eingegrenzt werden kann. Er beinhaltet zusammenfassend betrachtet nicht nur fachliche Eigenschaften, sondern darüber hinaus auch soziale und persönliche Fähigkeiten, mit aktuell auftretenden Situationen umgehen und als Lösung zur Bewältigung der Lage die notwendigen Handlungen daran anzupassen zu können. In jedem Individuum sind zum einen fortdauernde, ausgeprägte und zum anderen veränderbare oder auch verschiebbare Kompetenzen vorhanden. Wichtig innerhalb der Kompetenzdebatte ist, dass die Charakteristik der eigenen Persönlichkeit nicht mit dem Begriff der Kompetenz in Verbindung steht. Zusammenfassend und bezogen auf die Auszubildenden in der Gesundheits- und Krankenpflege heißt das, dass im Rahmen von praktischen Examina in der Gesundheits- und Krankenpflege von den Prüflingen ein breites erworbenes Kompetenzkonstrukt verlangt und umfassend abgefragt wird.

3.2 Erwerb von praktischen Kompetenzen

Wie bereits in den vorherigen Kapiteln aufgeführt, bedarf es für die praktische und berufliche Eignung in den unterschiedlichsten Berufsfeldern (ca. 300 Ausbildungsberufe) einer hohen praktischen Fachkompetenz, um den täglichen Anforderungen gerecht zu werden. Dies verlangt, wie bereits bei Vergleichsstudien in allgemeinbildenden Schulen aufgezeigt (z.B. PISA), auch in all den verschiedenen beruflichen Ausbildungssettings nach einer standardisierten und rahmenorientierten Erfassung bzw. Einschätzung von erworbenen praktischen, beruflichen Kompetenzen, die als Vorbereitung für die zukünftige, berufliche Arbeit dienen können (vgl. Beck, Landenberger, Oser 2016: 9ff.; vgl. Sauter, Staudt 2016: 3f.; vgl. Winther 2010: 5ff. u. 12f.).

Um sich den beruflichen Inhalten zu nähern, verlangt es nach ausreichenden vorhandenen „Grundlagenkompetenzen“, wie beispielsweise Lesen und Verstehen eines vorgelegten Textes, Bildmaterial mit Untertext erfassen und im beruflichen Praxisalltag umsetzen können. Prämisse und zu berücksichtigen, ist der Faktor „Zeit“ (vgl. Baethge, Seeber 2016: 22f.).

Gerade in akuten, fragilen Situationen im klinischen und ambulanten Kontext der Gesundheits- und Krankenpflege verlangt es nach einer schnellen, strukturierten und umfassenden Betrachtung des zu betreuenden Patienten. Das pflegerische Vorgehen fordert dabei eine ausführliche Krankenbeobachtung, Ermittlung von Parametern (wie z.B. Pulskontrolle), Sichtung der Patientenakte, Absprachen mit dem behandelnden Arzt, um nur einige wenige zu benennen. Diese Anforderungen benötigen eine Vielfalt an Kompetenzen für die Pflegekraft, um für die professionelle Betreuung des Patienten geeignet zu sein. Eine konkrete und einheitliche Festlegung der notwendigen Kriterien bzw. Kompetenzen und darüber hinaus ein geeignetes Messinstrument für Auszubildende in der Gesundheits- und Krankenpflege in Deutschland gibt es zum momentanen Zeitpunkt nicht. Eine didaktische Neugestaltung unter Zuhilfenahme von geeigneten pädagogischen und technologieorientierten Hilfsmitteln wäre durchaus notwendig und bereichernd (vgl. Darmann-Finck, Glissmann 2011: 195ff.; vgl. Darmann-Finck, Reuschenbach 2013: 23ff.; vgl. Minks 2011: 21, 28, 30).

3.3 Messung von praktischen Kompetenzen

Die Ermittlung der verschiedenen Kompetenzprofile erfolgt u.a. in beruflichen sowie schulortspezifischen Settings. Dort werden Einschätzungen vor allem zu allgemeinen und beruflichen Basiskompetenzen, Lernfähigkeit, soziale und methodische Kompetenzen, Verhaltenskompetenz, Selbstkompetenz, Selbst- und Fremdeinschätzungskompetenz und vernetztes Denken[20] vorgenommen (vgl. dazu auch die gesamten Projekte in Erpenbeck, Rosenstiel, Grote, Sauter 2017).

Damit Kompetenzen ermittelt werden können, ist die Festlegung eines Kompetenzprofils, bezeichnet auch als „Kompetenzmodell", notwendig. Dieses beschreibt das theoretische „Soll-Profil". Die durchgeführte Messung der Soll-Kompetenzen charakterisiert nach der Analyse wiederum das „Ist-Profil" bzw. deren Abweichung vom Teilnehmer. Die Kompetenzen können u.a. in schriftlichen Testungen als quantitative Messungen oder in simulierten Fallsituationen innerhalb eines Skills Lab ermittelt werden. Aufgrund der gemessenen Abweichung wird ein Entwicklungsplan oder auch Meilensteinplan der defizitären bzw. zu fördernden Kompetenzen für den Auszubildenden erstellt. Das kann gemeinsam durch Lehrende und Lernende erfolgen. Darauf folgt die kontinuierliche Leistungsüberprüfung der beruflichen Arbeitsergebnisse und anschließender erneuter Überprüfung des „neuen Kompetenzprofils" bzw. veränderten „Ist-Zustandes" und der erhofften Annäherung zum fixierten „Soll-Zustand" im festgelegten Kompetenzprofil. Gesamtheitlich betrachtet entsteht ein individueller Kompetenzzyklus für jeden Auszubildenden (vgl. Sauter, Staudt 2016: 4f., 7f.).

Zur Vorbereitung und Selbsteinschätzung auf die praktische Prüfung und das zukünftige Praxisfeld in der Gesundheits- und Krankenpflege wäre es sinnvoll, eine Möglichkeit für den Auszubildenden, den Lehrenden und die gesamte Institution Schule zu schaffen, die eine Option zur individuellen Einschätzung bzw. Messung von praktischen Pflegekompetenzen ermöglicht. Eine Bewertung des „Ist"-Zustandes wird erreicht. Bei der Entwicklung eines solchen praxisorientierten Messinstrumentes dürfen die gesetzlichen Vorgaben (wie bereits in Kapitel 2 erwähnt) und die curricularen Vorgaben, wie z.B. vorläufiger Landeslehrplan Baden-Württemberg für die Ausbildung[21], oder auch (falls vorhanden) interne Schulrichtlinien bzw. Schulcurricula nicht außer Acht gelassen werden. Die Schaffung von realitätsnahen Szenarien soll dem Testteilnehmer die Vorstellung einer zwar simulierten, aber doch möglichst „realen" beruflichen Pflegewelt bei der Lösung bzw. Beantwortung der Fragestellungen innerhalb des eingesetzten Messinstrumentes erleichtern. Aufgrund der gedachten Entwicklung von Fallbeispielen und optischen Simulationen wäre der Einsatz eines technologiegestützten Messinstrumentes zu bevorzugen (vgl. Sauter, Staudt 2016: 7ff.; vgl. Knigge-Demal, Eylmann 2006: 15f.).

[20] die genannten Kompetenzbeispiele stammen aus den Projekten: PIAAC (Programme for the International Assessment of Adult Competencies); Kompetenzrad und Kompetenzmatrix nach North; KKR (Kasseler-Kompetenz-Raster); CAPTain (Computer Aided Personnel Test answers inevitalbe); smk[72] (Einschätzungsbogen zu sozialen und methodischen Kompetenzen); hamet 2 (Handwerklich-motorischer Eignungstest)

[21] zur „Gesundheits- und Krankenpflegerin" oder zum „Gesundheits- und Krankenpfleger" und zur „Gesundheits- und Kinderkrankenpflegerin" oder zum „Gesundheits- und Kinderkrankenpfleger"

Nach den bis zu diesem Zeitpunkt erhobenen Recherchen zeigt sich, dass das TEMA-Messinstrument einschließlich der gesamten ASCOT-Studie diese Eigenschaften und Möglichkeiten bietet. Es offeriert die Möglichkeit der Einschätzung von praktischer Pflegekompetenz in den Pflegeberufen. Das nächste Kapitel zeigt die Begründung für den weiteren Einsatz des bereits vorliegenden Kompetenzmessinstruments auf.

4 Technologiebasierte Kompetenzmessung in der beruflichen Bildung

Da bisher keine geeigneten Messinstrumente für die Messung von Kompetenzen in der beruflichen Bildung vorhanden waren, hat das „BMBF“ dieses Thema aufgegriffen (siehe dazu auch Kapitel 1). Unter dem Motto „Neue Wege der Kompetenzmessung - Impulse für die berufliche Bildung“[22] wurde ein Großprojekt ins Leben gerufen. Dieses hat sich unter dem globalen Begriff „ASCOT“[23] auf der Grundlage formiert, kompetentes Praxisverständnis wird mit erfahrender, qualifizierter Wissenschaft gepaart. Zunächst soll das Ursprungsprojekt vorgestellt werden und die Fokussierung auf das Teilprojekt „TEMA“ und seine geplante Weiterführung in der vorliegenden empirischen Studie erläutert werden.

4.1 „ASCOT“-Studie

Die Dauer des Gesamtprojektes war über drei Jahre angelegt, d.h. von Ende 2011 bis Ende 2014. Zielsetzung ist, berufliche Handlungskompetenzen zu ermitteln und Auszubildende, im letzten Ausbildungsabschnitt, auf ihre zukünftigen Aufgaben in ihrem Arbeitsleben vorzubereiten bzw. die Bewältigung von anstehenden Aufgaben einzuschätzen. In der Umsetzung der großangelegten, berufsorientierten Offensive steht gesamtinhaltlich die Entwicklung und Erprobung neuer Messinstrumente im Mittelpunkt, speziell im Bereich der beruflichen Bildung mit dem Focus auf den deutschen Arbeitsmarkt.

> „In den Projekten werden Kompetenzmodelle und Messinstrumente entwickelt und erprobt, die es ermöglichen, Aussagen über Leistungsniveaus im Zusammenhang mit Ausbildungsqualität zu treffen.“ (BMBF 2014: 7).

In den teils computergestützten Messverfahren werden videogestützte Simulationen aus dem beruflichen Alltag gezeigt, ein anderer Teil geschieht per Paper-Pencil-Verfahren[24]. Die Entwicklung der Messinstrumente basierte auf dem Bezug von gesetzlichen Vorgaben zum jeweiligen Berufsbild, curricularen Vorgaben, schulortorientierten Lehrplänen usw. Es besteht die Möglichkeit, die einzelnen entwickelten Messinstrumente auf weitere Ausbildungsberufe nach Anpassung zu übertragen oder nach bereits langjähriger Tätigkeit im erlernten Berufsfeld unter dem Blickwinkel des „Lebenslangen Lernens“ mit einzusetzen sowie bedeutende, praxisrelevante Inhalte zu reanimieren. Darüber hinaus ist es möglich, wie bei PISA, auch auf internationaler Ebene berufliche Vergleichsstudien zu entwickeln und durchzuführen. Basierend auf dieser Grundlage könnte das in der Europäischen Union die Möglichkeit bieten, einheitliche Kompetenzstandards für diverse Ausbildungsberufe festzulegen und zu analogisieren (vgl. BMBF 2014: 9 u. 26; vgl. Beck, Landenberger, Oser 2016: 12; vgl. BMBF 2012: 2-5).

[22] Hinweis: vgl. ASCOT-Studie: “URL: https://www.ascot-vet.net/index.php [Stand: 22.02.19]”.
[23] Hinweis: vgl. Broschüre des BMBF: „URL: https://www.bmbf.de/upload_filestore/pub/ASCOT.pdf [Stand: 22.02.19]“.
[24] „Papier- und Bleistifttestung“ - traditionelles und schriftliches Datenerhebungsverfahren

Die ASCOT-Initiative wurde bei drei renommierten Berufssparten durchgeführt: Focus kaufmännischer Bereich, Focus technischer Bereich und Focus gesundheitlicher Bereich (vgl. Beck, Landenberger, Oser 2016: 10). Für jeden Teil wählte man jeweils zwei bedeutende Berufsbilder aus, im Gesundheitssektor den Medizinischen Fachangestellten und den Altenpfleger (vgl. BMBF 2014: 8; vgl. BMBF 2012: 5; vgl. Beck, Landenberger, Oser 2016: 10).

Bei der Entwicklung und Testung der einzelnen Messinstrumente wurde nicht ausschließlich das ausbildungsbezogene Berufsbild betrachtet, sondern ganzheitlich auf weitere, für das jeweilige Berufsbild notwendige, komplexe Kompetenzprofil erweitert. Diese ergänzenden Messinstrumente bzw. Messungen wurden im Bereich der Basiskompetenzen sowie der sozialen und interagierenden Kompetenzen zusätzlich entworfen bzw. vorgenommen (vgl. Nickolaus, Walker 2016: 12f.).

Es entstanden nach Einsatz und Erprobung im Berufsfeld sowie statistischer Evaluation der einzelnen Messinstrumente in der beruflichen Praxis Werkzeuge auf empirisch fundierter und valider Forschungsbasis, damit sogenannte „Kompetenzniveaus“ ausgewiesen werden können (vgl. Baethge, Seeber 2016: 19ff). Damit dies erreicht werden konnte, wurden die erfassten Werte der Teilnehmer mit dem in der merkmalsorientierten Psychologie fast schon standardisiert eingesetzten, statistischen Verfahren in Anlehnung an Rost durchgeführt, der sogenannten Item-Response-Theorie[25], mit dem zentralen Kern der Erfassung von Itemschwierigkeit zur Personenfähigkeit. Bei der Entscheidung für ein konstruiertes Messmodell geht es um die methodische Verwendung von Parameterschätzungen und Modellgeltung. Unterschiedlichste Merkmalsausprägungen können so erfasst werden (vgl. Rost 2004: 89; vgl. Eid, Schmidt 2014: 20ff.; vgl. Baethge, Seeber 2016: 21ff.).

In „ASCOT“ wird darüber hinaus auch von dem Wunsch der verstärkten „Durchlässigkeit“ gesprochen (vgl. Weiß 2016: 5), d.h. es sollen verwandte Berufsbilder bei profilverwandten Kompetenzbereichen eine gegenseitige Anerkennung möglich werden lassen. Dies könnte zukünftig als ein standardisiertes und praktikables Zulassungskriterium für eine hochschulische Laufbahn hilfreich sein (vgl. Weiß 2016: 6).

Die Formierung der beiden Gesundheitsberufe war in den Projekten „TEMA“ für den Altenpflegeberuf und „CoSMed[26]“ für den Beruf des Medizinischen Fachangestellten bzw. „Arzthelfer“ zu finden.

Damit die Übertragung in die Gesundheits- und Krankenpflege erreicht wird, war es zunächst notwendig, das Ursprungsprojekt „TEMA“ mit seinem gesamten Projektablauf nachzuvollziehen und aufzubereiten. Bei der Begutachtung des Testinstrumentes aus „TEMA“ stellte sich bei den verwendeten Testitems heraus, dass eine zusätzliche Integration von spezifischen Testitems aus dem Teilprojekt „CoSMed“ stattgefunden hat. Es werden deshalb beide Teilprojekte, bezogen auf den vorliegenden Forschungsauftrag, vorgestellt.

[25] Abkürzung: IRT
[26] Competence Measurement based on Simulations and adaptive Testing in Medical Settings

4.2 Teilprojekt: „TEMA“

Im „TEMA“- Projekt waren mehrere Hürden zu überwinden, da die Entwicklung eines geeigneten technologiegestützten, computerbasierten Messinstrumentes für eine wissenschaftlich fundierte und praxisorientierte Pflege eine Premiere darstellte.[27] Es fehlte bisher für die Auszubildenden, wie auch die Lehrkräfte in der Berufssparte „Pflege“, an einer empirisch aufbereiteten Grundlage zur Einschätzung und eventuell notwendigen Anpassung von beruflichen Pflegehandlungen im täglichen Praxisalltag. Eine bereits vorliegende und nutzbare Methodenentwicklung ist nicht vorhanden. Eine große Herausforderung stellte sich für die Forschergruppe[28] durch die multifaktorielle Kompetenzkomplexität innerhalb des Pflegeberufes.

> „… Neben dem hohen fachlichen Standard sind Pflegeberufe stärker als andere Berufe durch innere (emotionale, selbstreflexive, identitätsbezogene) und spezifische äußere Bezüge (z.B. Zeitdruck, institutioneller Rahmen, Konflikthaftigkeit der Situationen, Erfordernis aktiver Beziehungsgestaltung) geprägt. Dies gilt es bei der Konstruktion von Kompetenzmodellen und Testaufgaben zur Abbildung beruflicher Handlungskompetenz zu berücksichtigen.“ (Döring et al. 2016a: 244).

Die Entwicklung eines „berufsfachlichen pflegerischen Kompetenzmodells“ wurde beispielhaft auf den Bereich der Altenpflege fixiert, aber mit der Option der Anwendbarkeit, neben den videogestützten, simulierten Handlungssituationen, in der Gesundheits- und Krankenpflege und der Gesundheits- und Kinderkrankenpflege (vgl. Döring et al. 2016a: 244f.; vgl. Döring et al. 2016b: 127ff.).

Das „domänenspezifische“ Kompetenzmodell entstand in mehreren Analyse-Phasen. Zusammengefasst gesagt, durch

1. eine ausführliche Literaturanalyse zu folgenden Themen:
   - diverse nationale wie internationale Leistungsstandards,
   - bereits vorhandene Kompetenzmodelle,
   - gesetzliche Vorschriften in der Altenpflege wie auch länderspezifische curriculare und lehrplanerische Vorgaben aller Bundesländer,
   - vertiefte, eingrenzende und exemplarische Betrachtung, aufgrund der hohen Divergenz, mit einer Vergleichserstellung von Lernbereichen und Lernfeldern bei den Bundesländern Bayern, Nordrhein-Westfahlen, Brandenburg, Sachsen,

[27] Die Zielsetzung war die Entwicklung eines Kompetenzmodells zur Messung von beruflicher Handlungskompetenz im pflegerischen Alltag, umgesetzt in einem computergestützten Testverfahren.

[28] bestehend aus Forschungsinstitut Betriebliche Bildung (f-bb) gGmbH; Otto-Friedrich-Universität Bamberg, Sozial- und wirtschaftswissenschaftliche Fakultät; Deutsches Institut für Internationale Pädagogische Forschung (DIPF); Fachhochschule Bielefeld, Lehreinheit Pflege und Gesundheit vertreten u.a. durch die Personen: Dr. Ottmar Döring, Prof. Dr. Eveline Wittmann, Prof. Dr. Johannes Hartig, Prof. Dr. Ulrike Weyland, Prof. Dr. Annette Nauerth, Dr. Roman Kaspar, Dr. Michaela Möllers, Simone Reuschenbach, Julia Simon, Iberé Worofka.

2. leitfadengestützte, halbstandardisierte Experteninterviews mit 11 Interviewpartnern[29] mit den Ergebnissen zu „Wissen und Können“ von Pflegekräften und kontextgebundene Emotionen in der Pflege und der möglichen Übertragbarkeit auf die Gesundheit- und Krankenpflege und Gesundheits- und Kinderkrankenpflege, bewohner- und personalbezogene Empathie in der Pflege,

3. ein Expertenrating,

4. eine Expertenrunde (vgl. Döring et al. 2016a: 245f.; vgl. Döring et al. 2014: 18ff.; vgl. ASCOT-Forschergruppe 2014).

Daraus resultierte das pflegerische, handlungsorientierte und „domänenspezifische“ Kompetenzmodell mit seinen drei übergeordneten „Kompetenzbereichen“ und sechs „Teilkompetenzen“. Die Operationalisierung betraf in „TEMA“ ausschließlich einen der Kompetenzbereiche, den Bereich der unmittelbar bewohner-/klientenbezogenen Kompetenz mit ihren entwickelten Teilkompetenzen diagnostisch-reflexiv (DR)[30], praktisch-technisch (PT)[31] und interaktiv-kommunikativ (IK)[32]. Der Focus richtet sich innerhalb des Kompetenzbereiches hauptsächlich auf den zu betreuenden Bewohner und sein privates Umfeld. In den simulierten Videosequenzen wird die Pflegekraft auf ihre Reaktionen, wie Verhalten im emotionalen Kompetenzbereich einschließlich der daraus folgenden Pflegehandlungen, überprüft und bewertet. Die weiteren Kompetenzbereiche und ihre Teilkompetenzen sind in der folgenden aufgeführten Abbildung ersichtlich und werden nicht weiter vertieft. Bei Wittmann, Kaspar, Döring wird von dem Begriff es „Heuristischen Kompetenzmodells“ gesprochen (vgl. Wittmann, Kaspar, Döring 2017: 189).

Das lässt den Schluss zu, dass einerseits die zeitlichen Ressourcen für die zwei weiteren Kompetenzbereiche in ihrer Tiefe nicht erfasst werden können, andererseits die Priorität im Rahmen dieser Ersterfassung auf den zentralen Pflegeaspekt der Bewohner-Pflegekraft-Beziehung gelegt werden soll. In den gesichteten Unterlagen zum TEMA-Projekt ist dazu keine eindeutige, offensichtliche Begründung zu erkennen (vgl. Döring et al. 2016a: 246).

---

[29] Die Pflegeexperten waren aus den Bereichen der Pflegewissenschaft, Pflegepädagogik, Pflegemanagement und Pflegepraxis (vgl. Döring et al. 2014: 20f.; vgl. Wittmann, Kaspar, Döring 2017: 188f.).

[30] „Kompetenz zur begründeten, Folgen berücksichtigenden Entscheidungsfindung und -veränderung durch umfassende Einschätzung des Zustandes der zu pflegenden Person sowie der Umgebungsbedingungen.“ (Döring et al. 2016a: 246).

[31] „Kompetenz zu Interventionen bezogen auf zu pflegende Personen und ihre unmittelbare Umgebung mittels pflegerischer und medizinischer Techniken, Methoden und Hilfsmittel im Hinblick auf den umfassenden Pflegebedarf. (Letzterer basiert unter anderem auf der objektiven und subjektiven Pflegebedürftigkeit).“ (Döring et al. 2016a: 246-247).

[32] „Kompetenz zur Gestaltung der Interaktion und Kommunikation zwischen Pflegepersonen, Gepflegten und Bezugspersonen bezüglich des umfassenden Bedarfs der zu pflegenden Person.“ (Döring et al. 2016a: 247).

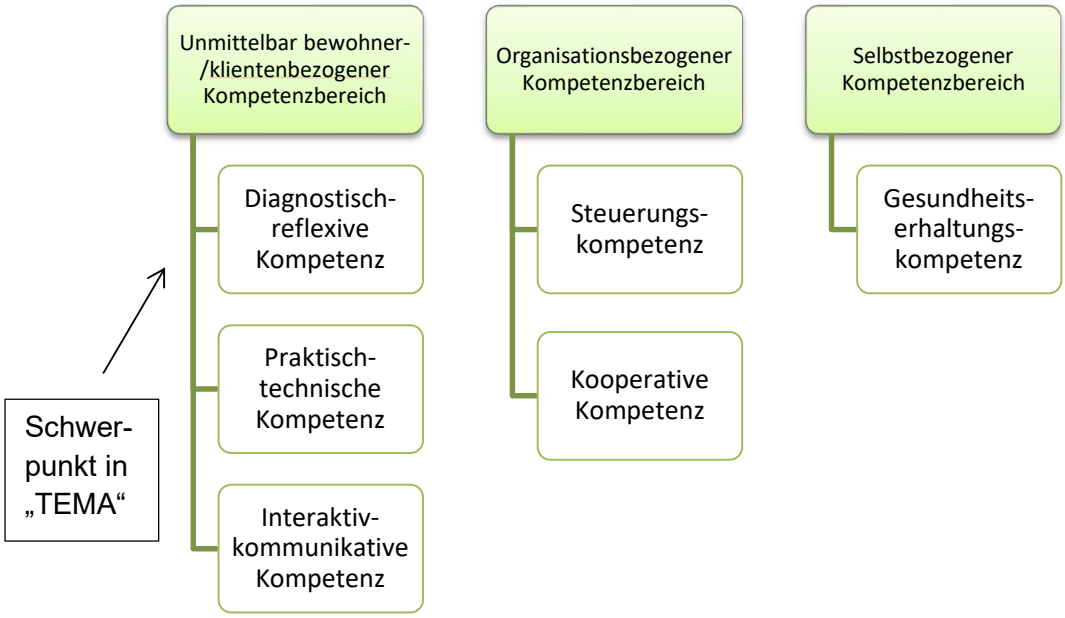


Abb. 1: Gesamtbetrachtung des „heuristischen“ Kompetenzmodells mit Bezug auf die Pflegepraxis (entnommen aus: Wittmann, Kaspar, Döring 2017: 189 leicht modifiziert von Ries 2019)

Damit eine hohe Zuverlässigkeit bzw. Reliabilität innerhalb der Kompetenztestung nachgewiesen werden konnte, wurden zu den 3 Teilkompetenzen im unmittelbar bewohner- und klientenbezogenen Kompetenzbereich jeweils über 20 Items in unterschiedlichen „Pflege“/„Handlungs“-Settings mit jeweils einem dafür typisch vorkommenden Krankheitsgeschehen entwickelt: Wohngruppe (Demenz), ambulante Pflege (Apoplex und Vorliegen eines Diabetes Mellitus Typ II) und stationäre Pflege (Sterben und Tod). Es wurden 12 videogestützte, pflegerische Handlungssituationen[33] entwickelt, die den Alltag dieser Settings aufzeigen und ein Einfinden in die Situation der dargestellten und simulierten Pflegekraft erleichtern sollten. Anschließend wird um eine Lösung oder Unterstützung im Pflegehandeln vom TN gebeten. Als Grundlagen dafür wurden das „pflegedidaktische Konzept“ nach Hundenborn und Knigge-Demal und das „methoden- und mediendidaktische Konzept“ der Fallstudie nach Kaiser herangezogen. Im Mittelpunkt steht die Verknüpfung einer auftretenden Pflegesituation und der daraus resultierenden Pflegehandlungen mit Bezug auf den gesamten Pflegeprozess. Hintergründe werden hier aber nicht weiter beleuchtet (vgl. Knigge-Demal, Eylmann, Hundenborn 2013; vgl. Hundenborn 2016: 46; vgl. Kaiser 1985: 440ff.; vgl. Kerngruppe Curriculum 2006: 88ff.; vgl. Döring et al. 2016a: 248).

Die Summe der entwickelten Testaufgaben bzw. Testitems für die einzelnen Kompetenzbereiche war nicht darauf angelegt, in allen drei Bereichen gleich hoch zu sein, sondern war so aufgebaut, dass die vorhandenen Items den Kompetenzbereich valide erklären sollten. Es wurde zur leichteren Handhabung der Auswertung der TN-Daten generell ein geschlossenes Konstrukt zur Beantwortung der Fragestellungen gewählt. Nach mehrfachen Pretests und folgenden Anpassungen der

[33] auch als „Videostimuli“ in „TEMA“ bezeichnet

entwickelten Testaufgaben ergaben sich insgesamt 77 objektive Items[34] für das „TEMA“- Messinstrument[35]. Die Anzahl der stabilen und validen Testitems sind aus der folgenden aufgeführten Tabelle zu entnehmen (Tab. 1) (vgl. Döring et al. 2016a: 248f.; vgl. Döring et al. 2016b: 122ff.).

Tab. 1: Verteilung der entwickelten Items einschließlich entwickelter Videovignetten im unmittelbar bewohner- und klientenorientierten Kompetenzbereich (aus: Döring et al. 2016: 249 leicht modifiziert von Ries 2019)

Handlungsbereiche	Anzahl der Videostimuli	Diagnostisch - reflexiv	Praktisch - technisch	Interaktiv - kommunikativ	Summe
Pflege und Betreuung dementiell erkrankter älterer Menschen (Setting 1 - Wohngruppe)	5 Videostimuli	10	9	7	26
Pfleg und Betreuung chronisch erkrankter Menschen (Setting 2 - Ambulante Pflege)	4 Videostimuli	12	7	8	27
Pflege und Betreuung älterer Menschen am Lebensende (Setting 3 - Stationäre Pflege)	3 Videostimuli	12	5	7	24
Summen der entwickelten Testitems		34	21	22	77

Zu den entwickelten Videosequenzen[36] und anschließenden schriftlich aufgeführten einzelnen Items (u.a. mit ergänzendem Bildmaterial oder Aufzeigen von Image-Maps[37]) wäre noch hinzuzufügen, dass diese für den TN so entwickelt wurden, dass sie in der visuellen Erfassung eine klare Strukturierung erkennen lassen und nicht mit unnötigen Informationen zu den einzelnen Handlungssituationen einschließlich Aufgabenstellungen angereichert sind. Das erreichte das Forscherteam aus „TEMA“ durch die Schaffung einer klaren Vorgehensweise durch den gesamten Kompetenztest, wie z.B. Buttons auf jeder Testseite zum Weiter- bzw. Zurückklicken im Testverlauf. Es waren keine hohen EDV-Kenntnisse notwendig, da (fast) ausschließlich das Bedienen der Maus notwendig war für den Teilnehmer (Bedienerfreundlichkeit!). Durch die einfache Handhabung, konnte das computergestützte Testinstrument von jedem TN bedient werden (vgl. Döring et al. 2016a: 249f.).

[34] die einzelnen Items erhalten in ihrer numerischen Auflistung den Buchstaben A und anschließend eine Zahl mit angehängt (z.B. A1). Diese werden den jeweiligen Teilkompetenzbereichen (DR, PT und IK) zugeordnet.

[35] bestehend aus 63 dichotomen Items und 14 polytomen Items; in der Mehrzahl in Form von Multiple-Choice-Aufgaben angelegt, d.h. es ist ein geschlossenes Antwortformat zu finden (vgl. Wittmann, Kaspar, Döring 2017: 195).

[36] wurde durch Laienschauspieler dargestellt nach der vorausgegangenen Entwicklung von Drehbüchern unter Bezug auf die entwickelten Testitems (vgl. Döring et al. 2014).

[37] anklickbare Bildtafeln zum Auswählen von bestimmten Teilbereichen

Der Ablauf der beruflichen Kompetenzmessung im „TEMA“-Projekt:

1. Durchführung des Logins mit einem speziellen Passwort und Code in den Kompetenztest
2. Einführung in den Test -> Handhabung erlernen und vertraut machen mit der vorhandenen Testumgebung durch einen kurzen Übungstest aus einem fachfremden Fachgebiet, z.B. Hauptstadterfragung
3. Beginn des eigentlichen Kompetenztests: bestehend aus den zuvor genannten drei Settings bzw. bestehend aus drei Testabschnitten, die alle bearbeitet werden sollten, und anschließend den ausgewählten Items aus dem Projekt „CoSMed“ (siehe auch Kapitel 4.3).

Der Kompetenztest beinhaltet sogenannte „Videovignetten“, die alle drei Kompetenzbereiche in dem jeweils vorgestellten Setting vereinen. Das bedeutet, dass z.B. im Setting Wohngruppe die DR, IK und PT Kompetenzen abgebildet werden. Es gibt „Intro-Videos“ in bildgebender Form und sprachlichen Erläuterungen, ergänzt durch einen Sprecher, „Fallvideos“ bzw. die bereits erwähnten „Videovignetten“ ohne erneutes Abspielen der Szenen (= reale Pflegesituationen können auch nicht zurückgedreht werden!), Bildausschnitte aus der bewohnerbezogenen Pflegedokumentation, Momentaufnahmen von z.B. Wundverhältnissen usw. und „Passbilder“ des Bewohners, der gerade im Mittelpunkt der pflegerischen Betreuung und Versorgung steht, diese sollen als Orientierungshilfe[38] dem TN dienen, da mehrere Bewohner in den Videovignetten vorkommen (vgl. Döring, Möllers, Schöpf 2014: 40f.; vgl. Döring et al. 2016a: 250-251).

Zur besseren visuellen Vorstellung des technologiebasierten, computergestützten, videobasierten Kompetenzmessinstrumentes soll in der folgenden Abbildung (Abb. 2) eine Übersicht der pflegehandelnden Settings gezeigt werden.

Abb. 2: Pflegesettings in „TEMA“ (entnommen aus: Döring, Möllers, Schöpf 2014: 38)

[38] Vermeidung einer Verwechslung des zu betreuenden Bewohners

Darüber hinaus sollen exemplarisch eine Einführungsseite in ein Pflegesetting „Setting-Startseite“ und eine Aufgabenstellung aus dem „TEMA“ - Messinstrument gezeigt werden, die den TN in der Kompetenztestung präsentiert wurden.

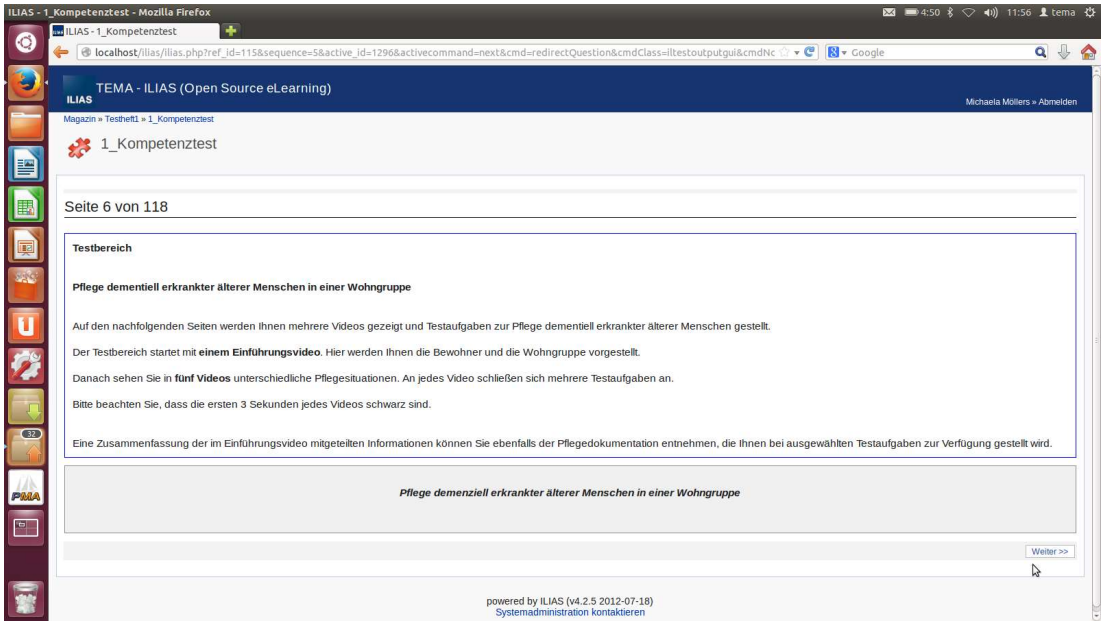


Abb. 3: „Startseite“ in ein Pflegesetting (Quelle: Screenshot „TEMA“-Messinstrument)

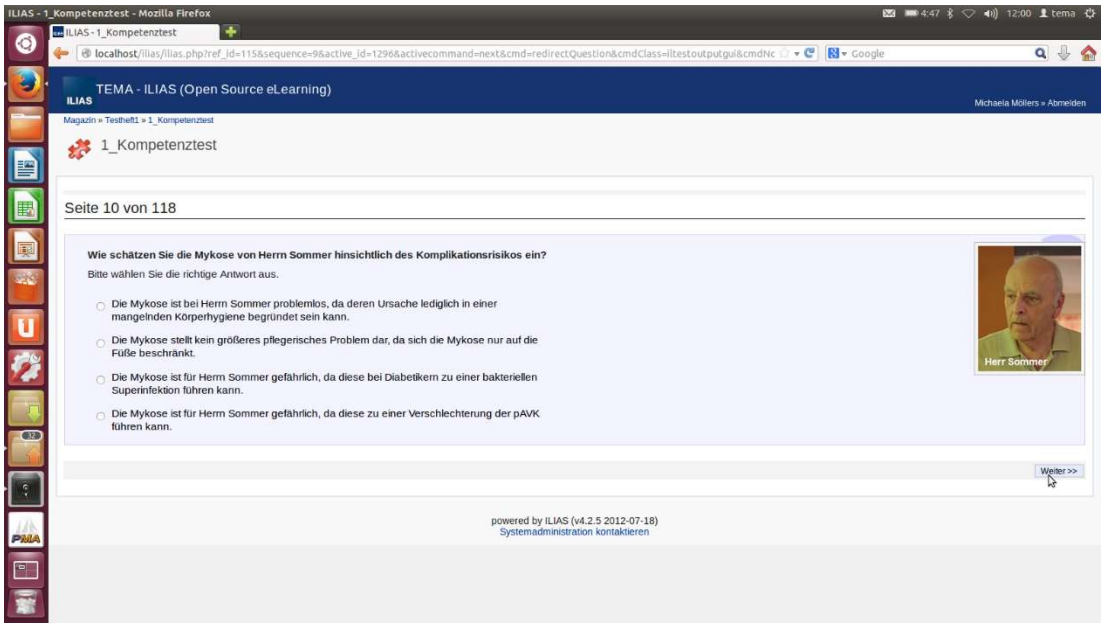


Abb. 4: „Beispieltestblatt“ aus dem pflegerischen und handlungsorientierten Kompetenztest (Quelle: Screenshot „TEMA“-Messinstrument)

Die Anwendung des pflegerischen und handlungsorientieren Kompetenztests fand im Ursprungsprojekt „TEMA“ in der Hauptuntersuchung in zwei Bundesländern, in Bayern (12 Schulen) und Nordrhein-Westfalen (11 Schulen), mit einem N von 402 (freiwilligen) Auszubildenden zum Ende der dreijährigen Ausbildung in der Alten-

pflege, statt. Die Teilnahme war für alle Auszubildenden nicht bindend. Ergänzt wurden Kontextmerkmale und Lesekompetenz mit den entwickelten Hilfsmitteln aus zwei weiteren Teilprojekten aus der ASCOT-Studie: „SiKoFak“[39] und „MaK-adapt“[40]. Aufgrund der knappen zeitlichen Ressourcen durch die bevorstehenden schriftlichen, praktischen und mündlichen Prüfungen konnten nicht in allen Schuleinrichtungen die drei Untersuchungsbereiche getestet werden (vgl. Döring et al. 2016a: 252; vgl. Döring et al. 2016b: 124f.).

Im Ergebnis standen dem Forscherteam aus dem Teilprojekt „TEMA“ in „TEMA“/Pflege 402, in „SiKoFak“/Kontext[41] 395, in „MaK-adapt“/Lesen 310 gewonnene Datensätze zur Verfügung. Die Evaluation der einzelnen Testaufgaben in „TEMA“ fand je nach Aufgabenstamm (= Videostimulus, Aufgabentext) und Aufgabenbeantwortung (= Anwortformat, Text der Antwortalternativen, Scoring[42]) in einem empirisch entwickelten Beantwortungsschema statt. Das Kompetenzniveau für jede Aufgabe (= Lösungswahrscheinlichkeitsrate) wurde anhand von durchgeführten Pretests vor der Haupterhebungsphase festgelegt. (vgl. Döring et al. 2014: 432ff.).

Die folgende Untersuchung in der hier vorliegenden Arbeit, ab II. Empirischer Teil 1 - Datenerhebung, beschränkt sich auf das „TEMA“-Kompetenzmessinstrument in der pflegerischen Bildung und einen Abriss von ausgewählten Kontextmerkmalen aus „SiKoFak“, die nach Sichtung innerhalb des Fachartikels von Döring et al. 2016 als zentral von der Forscherin eingeschätzt werden und einen Einfluss auf die Gesamtevaluation des anschließenden, weiterführenden Projekts in der Gesundheits- und Krankenpflege erwarten lassen (vgl. Döring et al. 2016a: 253ff.). Als signifikante Ergebnisse zeigen sich innerhalb der Kontextfaktoren, dass die Entscheidung für die Berufswahl zur Ausbildung in der Altenpflege für die Auszubildenden mit dem Hintergrund von „Spaß an der Arbeit“ und „Beruf eine sichere Berufsperspektive“ sich verknüpfen lassen (vgl. Döring et al. 2016a: 255).

Der Focus in der vorliegenden Arbeit liegt aber, bezogen auf die im Kapitel 1.2 benannte zentrale Forschungsfrage, auf dem Messinstrument aus „TEMA“. Weitere Details zur Erhebung der personenbezogenen Daten bzw. Kontextfaktoren im Rahmen des Projektes in der Gesundheits- und Krankenpflege finden sich im Kapitel 10.2.3 mit dem Hinweis zu den einzelnen Auswahlkriterien, wie z.B. Geschlecht, Lebensalter, Familienstand.

---

[39] Systemische und individuelle Kontextfaktoren der beruflichen Ausbildung (angelegt als Querschnittsprojekt u.a. zu „TEMA“)

[40] Messung allgemeiner Kompetenzen - adaptiv (angelegt als weiteres Querschnittsprojekt u.a. zu „TEMA“)

[41] hier werden institutionelle (z.B. Ausbildungsqualität) und individuelle soziale (z.B. Schulabschluss) Kontextfaktoren erfragt und erfasst (vgl. Baethge, Seeber 2016: 16; vgl. Baethge-Kinsky, Baethge, Lischewski 2016: 267).

[42] einstufig, zweistufig, dreistufig und vierstufig mit Angabe von gängigen Quellennachweisen als Begründung der korrekten Antwortoption bzw. Antwortoptionen (z.B. Item DR21A34/diagnostisch-reflexive Kompetenz mit dreistufigem Scoring: 0,1,2 - wenn korrekt -> A+B+D = 1 Punkt, wenn korrekt -> A+B+C+D = 2 Punkte)

In dem groß angelegten, statistischen Auswertungsportfolie des „TEMA“-Kompetenztests zeigen sich multidimensionale Ergebnisse. Zu berücksichtigen ist, dass bei der doch relativ geringen TN-Zahl von 402 TN und keiner breit angelegten Zufallsstichprobe (Einschränkung auf nur zwei Bundesländer) die Ergebnisse im „TEMA“-Kompetenztest bei der Datenanalyse, Datenbewertung einschließlich Dateninterpretation einzubeziehen sind (vgl. Wittmann, Kaspar, Döring 2017: 200).

Es stellten sich aus den 77 entwickelten Testitems 62 Items als statistisch valide und signifikant geeignet für die

> „… Messung der unmittelbar bewohner-/klienten-orientierten pflegerischen Kompetenz zum Ende des dritten Ausbildungsjahres in der Altenpflege …“ (Döring et al. 2016a: 255)

heraus, gleichzeitig wurden emotionale Merkmale mitberücksichtigt (Döring et al. 2016a: 256).

Im Rahmen der Item-Response-Theorie konnte eine Raschvalidität[43] bezüglich der entwickelten Testitems festgestellt werden. Das Auffinden einer lokalen stochastischen Unabhängigkeit einschließlich seiner itemcharakteristischen Funktionen (= ICC) wurde bestätigt (vgl. Hahnel 2018: 16). Eine weitere Vertiefung der Befunde findet sich in Kapitel 15 und wird an dieser Stelle nicht weiter analysiert werden.

Die ursprünglichen und weiteren Befunde im Rahmen der Konfirmatorischen Faktorenanalyse, der Dimensionalität der Leistungsdaten und der Auffindung eines psychometrisch passenden Pflegekompetenzmodells werden in Kapitel 16 erläutert und weiter betrachtet.

In der Beantwortung der Items im computergestützten, technologieorientierten Gesamttest „Pflegekompetenz“ erreichten die Auszubildenden bzw. TN in der Altenpflege nach Normierung[44] im Mittelwert 50 Punkte mit einer vorliegenden Standardabweichung von 10 Punkten, d.h. leicht unter der Hälfte der zu erreichenden normierten Gesamtpunktzahl. Insgesamt haben 294 TN von 402 TN mindestens 36 Punkte bzw. nach Normierung mindestens 50.42 Punkte und mehr erreicht. Wiederum 108 TN lagen unter den zuvor genannten Punktwerten (vgl. Kaspar et al. 2016: 19f.). Die Testitems sind aufgrund der vorliegenden Ergebnisse teilweise als zu schwer einzustufen, trotz der im Vorfeld stattgefundenen Beratung und Analyse durch Fachleute aus dem gesamten Pflegekontext (vgl. Wittmann, Kaspar, Döring 2017: 199).

---

[43] Es liegt eine moderate Modelpassung, reale Aufgabenschwierigkeit, usw. bezüglich der Abbildung des bewohner-/klientenorientierten Pflegekompetenzbereiches vor (vgl. Hahnel 2018: 34).

[44] normed person ability; ohne Normierung wurden zwischen 10 und 59 Punkten erreicht; bei den beobachteten Häufigkeiten (= observed frequency in TN-zahl und Prozent) waren in einem Ranking von oben nach unten auf den ersten drei Plätzen: 104 TN (25,9%) bei 50 Punkten bzw. 65.72 Punkten nach Normierung, 103 TN (25,6%) bei 36 Punkten bzw. 50.42 Punkten nach Normierung, 89TN (22,1%) bei 30 Punkten bzw. 43.50 Punkten nach Normierung. Im Gesamtergebnis des Kompetenztests gilt ein Wert zwischen 36 und 79 Punkten bzw. 50.42 Punkten und mehr nach Normierung als bestanden (vgl. Kaspar et al. 2016: 19f.).

Die anonymisierten einzelnen Schulen zeigen im Vergleich einen normierten WLE[45]-Wert zwischen 45,4 und 56,0 (> 1 Standardabweichung) an, d.h. es sind signifikante Unterschiede in den Messergebnissen der einzelnen Schuleinrichtungen im unmittelbar bewohner-/klientenbezogenen, pflegerischen Kompetenzniveau zu erkennen. Die Gesamtreliabilität[46] im Rahmen des Rasch-Modells zeigt sich nach Ausschluss von 14 Items mit 0.76 als solide. Vor dem Ausschluss lag der Wert bei 0.73. In der Einzelbetrachtung wiederum waren die Werte im diagnostisch-reflexiven Teilkompetenzbereich (0.546), praktisch-technischen Teilkompetenzbereich (0.437) und interaktiv-kommunikativen Teilkompetenzbereich (0.496) zu gering in ihrer ermittelten Reliabilität (vgl. Wittmann, Kaspar, Döring 2017: 198; vgl. Döring et al. 2016b: 126). Daraus folgt, dass der statistische Focus nur in der Gesamtbetrachtung des Kompetenztests und der ermittelten Reliabilität aussagekräftig ist.

Summa summarum ist der technologiebasierte, computergestützte Kompetenztest machbar, aber mit durchschnittlich 45% der möglichen Gesamtpunktzahl der TN zum Ende der dreijährigen Ausbildung in der Altenpflege als „anspruchsvoll“ oder negativ formuliert als „nicht gerade einfach“ einzustufen, jedoch eine gute Handlungsgrundlage für die Vorbereitung auf die praktische Prüfung in der Pflege sowie den bevorstehenden Arbeitsalltag und darüber hinaus (vgl. Wittmann, Kaspar, Döring 2017: 199f.).

Nach der Vorstellung des Teilprojektes „TEMA“ ist zu berücksichtigen, dass nicht die kompletten Befunde der breiten statistischen Datenaufbereitung und Datenanalyse dargestellt wurden. Der Focus lag auf den notwendigen Determinanten zur Koordinierung und Durchführung des Folgeprojektes in der Gesundheits- und Krankenpflege und folglich zur Beantwortung der gestellten Forschungsfragen aus Kapitel 1.2. Die zentralen Ergebnisse der Übertragung und Verknüpfung der beiden Projekte finden sich im Teil III - Empirischer Teil 2.

Wie bereits zum Ende von Kapitel 4.1 erwähnt, wurde darüber hinaus eine selektive Ergänzung an Testitems aus dem Teilprojekt „CoSMed“ im verwendeten Testinstrument vorgenommen.

4.3 Teilprojekt: „CoSMed“

Das technologiebasierte, computergestützte „TEMA“-Kompetenzmessinstrument wurde im Anschluss an die empirisch entwickelten Testitems aus der Altenpflege mit 12 videounabhängigen Items aus dem Teilprojekt „CoSMed“ für Medizinische Fachangestellte erweitert. Die gewählten Items[47] zeigen eine berufliche Schnittstelle zwischen den beiden Berufsbildern (vgl. Seeber et al. 2016: 206). Es waren keine bildgebenden und einführenden Simulationen mitintegriert, sondern es wurde nur das klassische „Frage-Antwort-Prinzip“ gewählt, so dass eine autonome Betrachtungs-

[45] Normierte Weighted Likelihood Estimates -> „gewichtete Wahrscheinlichkeitsschätzung“
[46] Messgenauigkeit eines Tests, Wert soll gegen 1 gehen -> dann besonders hohe Genauigkeit!
[47] Diese wurden in der nummerischen Reihenfolge im Kompetenztest mit den Kürzeln A83 bis A94 festgelegt. Eine konkrete Widergabe der einzelnen Iteminhalte darf aufgrund des Einsatzes für weitere Forschungsvorhaben nicht erfolgen. Es besteht sonst die Gefahr eines Wiedererkennungseffektes für die TN und somit wäre keine objektive Einschätzung mehr möglich.

weise vom TN und später bei der Auswertung auch von der Forscherin ermöglicht wurde. Das Konstrukt der entwickelten Items bezog sich fast (11 von 12 Items) ausschließlich auf den „Frage-Antwort-Typus“ Multiple-Choice mit vier Antwortmöglichkeiten und der Wahl einer Option. Die Abfrage nach semantischem[48] und episodischem[49] Wissen beim TN stand bei der Itemauswahl im Vordergrund. Eine curriculare Begutachtung auf Relevanz und Eignung für den Pflegeberuf fand zuvor statt (vgl. Döring et al. 2014: 238ff.; vgl. „TEMA“-Projektgruppe 2014: 33ff.).

Beispielsweise wird in Item A 93 in einem zuvor vorgestellten, fiktiven „Patientenfall“ nach der korrekten Anzahl der Pulsschläge pro Minute gefragt (vgl. „TEMA“-Projektgruppe 2014: 33ff.).

Ein weiteres Item (A87) verlangte dagegen, von vier Antwortmöglichkeiten zwei richtige Antworten zu finden. Dabei geht es inhaltlich um die Überwachung einer Bewohnerin mit vorliegender Kreislaufinstabilität und daraufhin folgender Verabreichung von Kochsalzlösung per Dauertropfinfusion (vgl. „TEMA“-Projektgruppe 2014: 34).

Die im Anschluss erfolgte Datenaufbereitung und Datenanalyse der 12 „CoSMed“-Testitems wurde separat von den eigentlichen „TEMA“- Testitems im Ursprungsprojekt durchgeführt. Der Akzent lag hier auf der Häufigkeitsverteilung der richtigen Antwortoptionen der TN, um die Eignung der Items für die Berufsgruppe Pflege einschätzen zu können. Eine detaillierte statistische Analyse und Weiterführung im Bereich der Item-Response-Theorie oder darüber hinaus wurde im „TEMA“-Projekt nicht vorgenommen (vgl. Döring et al. 2014: 244ff.).

Zur Übertragung in die Gesundheits- und Krankenpflege wurde der gleiche statistische Weg gewählt, um die Eignung der Testitems zu prüfen und Vergleiche mit der Altenpflege zu erzielen. Wie die Umsetzung im Folgeprojekt erfolgte, wird im Teil II - Datenerhebung und Teil III - Statistische Aufbereitung und Datenanalyse dargestellt.

[48] abstraktes Wissen
[49] situationsgebundenes Wissen

5 Erkenntnisgewinn aus „TEMA“ mit Einschluss von „CoSMed“

Nach den in Kapitel 4.2 und 4.3 vorgestellten und erläuterten Inhalten des computergestützten, technologiebasierten Kompetenztests für Auszubildende in der Altenpflege ist festzustellen, dass in den vorliegenden Ergebnissen weitere multiple Anschlussstudien folgen müssen. Der Mehrwert für eine „kompetente“ und praktische Tätigkeit in der Pflege muss überprüft werden. Das beruflich orientierte Projekt gilt in der bisher entwickelten Form als „Premiere“ mit hoher Bedeutung für das Feld der technologiebasierten, computergestützten Kompetenzmessungen im Bereich der Pflege in Deutschland. Als zentrales Ergebnis zeigt sich, dass eine Übertragung in weitere Berufsgruppen des Pflegeberufs nach empirischer Datenanalyse und daraus folgendem Ergebnis nach Meinung der Projektgruppe in „TEMA“[50] möglich ist, da die Items so gewählt wurden, dass sie auch in der Gesundheits- und Krankenpflege, der Gesundheits- und Kinderkrankenpflege sowie im generalistischen Studiengang Pflege angewendet werden können[51]. Dies gilt es aber erst noch durch weitere Projektinitiativen fundamental empirisch zu bestätigen bzw. zu diskutieren. Darüber hinaus sollte der Focus auf den zeitlichen Aufwand der Erstellung und den Einsatz von computergestütztem Kompetenztest versus Paper-Pencil-Methode-Testung gelegt werden. Hier gilt es, die Vor- und Nachteile der beiden unterschiedlichen Testinstrumente in der Analyse für weitere Projektvorhaben abzuwägen (vgl. Döring et al. 2016a: 261f.; Döring et al. 2016b: 127ff.; Kaspar et al. 2016: 25).

Gleichzeitig ist die zu gestaltende curriculare, gesetzliche und inhaltliche Relevanz der Kompetenzmessung, bezogen auf den Fachbereich Pflege, für die Zukunft weiterzuverfolgen. Es gilt, Möglichkeiten einer vorbereitenden, innovativen Lernüberprüfung sowie Kompetenzeinschätzung zu finden, die Lehrende wie auch Lernende auf anstehende praktische Abschlussprüfungen und das zukünftige Berufsfeld Pflege einstimmen und vorbereiten. Eine statistische Überprüfung auf Itemvalidität, Personenfähigkeit[52] und Itemschwierigkeit[53], Gesamtreliabilität, Modellgeltung usw. sind für ein empirisch und quantitativ überprüftes Messinstrument in der Pflegewissenschaft einschließlich Pflegepädagogik unerlässlich. Dies ist im aktuellen Projektverlauf zu bedenken und findet im Methodenteil seine Umsetzung.

[50] nach zusätzlichem Einbezug diverser Curricula, Lehrplanvorgaben, Modulhandbücher und von Experten aus dem Pflegekontext der Bereiche der Pädagogik, Management und Wissenschaft des Pflegekontextes möglich (vgl. Döring et al. 2016b: 129).

[51] = bestätigt in punktueller Erhebung und Pretest in der Gesundheits- und Krankenpflege bei „CoSMed“ mit einem N von 52 TN und richtig gewählten Lösungen zwischen 28,85% und 82,69% im Bereich der Häufigkeitsstatistik (vgl. Döring et al. 2014: 244).

[52] person ability

[53] item difficulty

6 Abschließender Kommentar zum „Theoretischen Teil“

In den Kapiteln 1 bis 5 wurden Hinführung und Begründungsrahmen einschließlich der Entwicklung des empirisch entwickelten „Heuristischen Modells“ als Grundlage für das zu prüfende Messinstrument zum anstehenden Projekt der zuvor vorgeschlagenen wissenschaftlichen Weiterführung und Übertragung des Kompetenzmessinstrumentes in die Gesundheits- und Krankenpflege, Gesundheits- und Kinderkrankenpflege einschließlich generalistischen Studiengangs im Bereich Pflege herausgearbeitet. In den literarischen Analysen zum „TEMA“ Projekt ist festzustellen, dass sich aufgrund der Komplexität des Forschungsauftrages mehrere Projektgruppen mit unterschiedlichen Arbeitsaufträgen und Schwerpunkten gebildet haben. So sind die mehrfach veröffentlichten, literarischen, wissenschaftlichen Buch- und Zeitschriftenbeiträge einschließlich vorhandener Abstracts der vier Projektpartner überwiegend aus der jeweiligen perspektivischen Herangehensweise bzw. Aufgabenbetrachtung des jeweiligen Projektpartners zu betrachten. Im Abschlussbericht wird deshalb auch von einem „Verbundprojekt“ gesprochen (vgl. Döring et al. 2014: 1).

Dazu gehören das Forschungsinstitut Betriebliche Bildung[54], die Otto-Friedrich-Universität Bamberg[55], das Deutsche Institut für Internationale Pädagogische Forschung[56] und die Fachhochschule Bielefeld[57] („ASCOT“-Forschergruppe 2014: 7).

Im nun folgenden empirischen Projekt werden eine Zusammenführung der einzelnen Projektpartner und deren Aufgabenschwerpunkte aufgezeigt, um eine Facettenbetrachtung zu vermeiden und einen Überblick zur Gesamtkomplexität des Forschungsauftrags herzustellen.

Im Ergebnis erhofft sich die Forscherin ein pädagogisches, standardisiertes „Assessmentinstrument“ als Möglichkeit der praktischen Prüfungsvorbereitung zur Fähigkeitserfassung für pflegepädagogische Lehrkräfte an Bildungseinrichtungen bzw. Hochschulen und deren Auszubildenden mit der Spezialisierung auf Pflege. Dies soll als Anreiz und Hilfestellung für weitere zukünftige Forschungsvorhaben im Setting Pflege dienen. Alternativ bietet es eine weitere Diskussionsgrundlage über den Einsatz von quantitativen Messinstrumenten zur Einschätzung von praktischer Pflegekompetenz.

[54] „Verbundkoordination, Technikentwicklung, Feldtests und Ergebnistransfer“ unter Leitung von: Herrn Dr. Ottmar Döring

[55] „Kompetenz- und Aufgabenmodellierung“ unter Leitung von: Frau Prof. Dr. Eveline Wittmann

[56] „Psychometrische Testverfahren“ unter Leitung von: Herrn Prof. Dr. Johannes Hartig

[57] „Curriculare und ökologische Validierung“ unter Leitung von: Frau Prof. Dr. Ulrike Weyland und Frau Prof. Dr. Annette Nauerth

II Empirischer Teil 1 - Datenerhebung

Im Teil II der Arbeit werden die Vorbereitungen für den Feldzugang ethisch zu berücksichtigende Aspekte, vorliegende Rahmenbedingungen, die Population innerhalb des Feldzugangs einschließlich der Durchführung der Datenerhebung in den einzelnen Einrichtungen aufgezeigt. Den Abschluss bildet ein kurzes Fazit zum gesamten abgeschlossenen Feldzugang innerhalb des Kompetenzprojektes im Setting der Gesundheits- und Krankenpflege.

7 Methodische Vorgehensweise

In der quantitativen Forschung wird mit einem „theoretical construct“ innerhalb der deduktiven[58] Methoden gearbeitet. Das bedeutet für das anstehende Projekt und die methodische Aufbereitung, dass die Entwicklung des zu prüfenden Testinstruments bereits im Ursprungsprojekt stattgefunden hat. Die Operationalisierung beschränkt sich auf die Wahl des „TEMA“-Testinstruments einschließlich der angeschlossenen 12 Testitems aus „CoSMed“. Eine Veränderung oder Ergänzung innerhalb der bestehenden Testitems wird nicht vorgenommen, sondern 1:1 auf den Kontext der Gesundheits- und Krankenpflege übertragen, damit Vergleiche zum Ursprungsprojekt und Beantwortung der zentralen Forschungsfrage erfolgen. Alle weiteren methodischen Aspekte der Datenerhebung werden in den folgenden Unterkapiteln angesprochen.

7.1 Vorbereitungen für den Feldzugang

Vor dem eigentlichen Einsatz des computergestützten Kompetenzmessinstruments waren zunächst verschiedene Forschungsfacetten zu berücksichtigen, zu bewerten bzw. einzuschätzen und abzuklären, bevor eine erste „Feldbegehung“ möglich wurde.

Als vorbereitende Maßnahmen wurden ein standardisiertes Informationsschreiben zur Gewinnung von Auszubildenden am Kompetenztest, eine Einverständniserklärung mit Datenschutzinformation und ein personenbezogener Fragebogen[59] entwickelt (Anhang 1 bis 3). Ein Forschungstagebuch wurde angelegt (Anhang 4). Weitere Details finden sich in Kapitel 4.2 und 9.

Als Kontakt bei Unklarheiten diente während des gesamten Projekts der direkte schriftliche, telefonische oder persönliche Austausch mit Personen aus den Forscherteams der „ASCOT“-Initiative und vor allem schwerpunktmäßig aus „TEMA“ sowie der Verleih des computergestützten Messinstruments (siehe ergänzend Kapitel 8.2 u. 8.3). Ohne diese Möglichkeiten wäre eine Annäherung und Verwirklichung eines weiteren anschlussfähigen Kompetenzmessprojektes gescheitert, da die untersuchten diversen, perspektivischen literarischen Veröffentlichungen für eine vollständige empirische Untersuchung nicht ausreichen.

[58] es wird vom „Allgemeinen“ auf das „Besondere“ geschlossen

[59] Das kontextuelle Querschnittprojekt „SiKoFak“ (systemische und individuelle Kontextfaktoren) aus der „ASCOT“-Initiative wurde bei der Entwicklung des Fragebogens mitberücksichtigt (vgl. Döring et al. 2016a: 253ff.; vgl. Baethge-Kinsky et al. 2016: 265ff.).

7.1.1 Ethische Reflexion

Bei der Durchführung eines wissenschaftlichen Forschungsprojekts, wie hier vorliegend, darf ein präventiver Blick auf die Einhaltung und Beachtung des Probandenschutzes bzw. auf die mögliche Vulnerabilität der TN nicht außer Acht gelassen werden. Es bedarf höchster Sensibilität mit „ethischer Strenge“[60].

Als unterstützende Assessments zur Prüfung der forschungsethischen Kriterien wurden u.a. der Ethikkodex der DGP[61] und ein Begutachtungsformular der Ethikkommission der DGP nach Schiffer[62] einschließlich der forschungsethischen Prinzipien nach Schnell mit einbezogen (vgl. DGP 2016; vgl. Schiffer 2018; vgl. Schnell 2007: 165ff.; [63]). Hier werden nachfolgende ethische Blickwinkel betrachtet.

Zunächst wurden zur Gewinnung von Bildungseinrichtungen ausführliche Informationen über den geplanten Einsatz des Messinstrumentes per Mail verschickt oder per Telefon ausführlich erläutert, um die Notwendigkeit für eine Teilnahme an der Studie anzuregen. Die Schulleitungen bzw. deren Lehrkräfte fungierten als Multiplikatoren bei den Auszubildenden ihrer Einrichtungen. Dabei standen sie im kontinuierlichen Austausch bezüglich Rückfragen, Vorbereitungen zur Testung usw. mit der Forscherin. Es wurde hier auch auf die Freiwilligkeit, die Anonymisierung der Daten, den Schutz der Einrichtung und der TN hingewiesen. Ein Muster der Einverständnis- und Datenschutzerklärung wurde versandt.

Die Freiwilligkeit und informierte Einwilligung wird umgesetzt, indem die Auszubildenden im Vorfeld durch einen schriftlichen Informationsbogen über die zwanglose Möglichkeit zur Teilnahme am Forschungsprojekt informiert werden. Die Regeln des „Informed Consent“ werden eingehalten, d.h. die gewonnen TN werden vor der eigentlichen Kompetenzmessung über das gesamte Projekt und den Ablauf der Testung mündlich informiert. Im Anschluss unterschreiben die TN eine Einverständnis- und Datenschutzerklärung. Ein Zurücktreten von der Studie ist jederzeit bei vollständiger Löschung der erhobenen TN-Daten möglich. Bei Unklarheiten sind Fragen an das Forscherteam jederzeit zu stellen bzw. erwünscht.

Der Punkt der Anonymisierung und Vertraulichkeit des Datenmaterials wird vollständig eingehalten. In der ersten Phase der Erhebung „personenbezogene Daten“ (Anhang 3) per Paper-Pencil-Methode werden keine Angaben von Name, Adresse, Telefonnummer, E-Mailadresse usw. erfragt, die Rückschlüsse auf die einzelnen TN zulassen könnten. In der zweiten Phase der Datenerhebung zur Verwendung des technologiebasierten, computergestützten Messinstruments konnte nur mit einem speziellen Einwahlkode, der per Zufallsprinzip auf den einzelnen Laptops nach dem Aufbau des gesamten Equipments verteilt wurde, an der praxisorientierten Kompetenzmessung teilgenommen werden. So befand sich in einem Umschlag jeweils ein Kärtchen das mit einem anonymisierten Benutzernamen (z.B. km_0007) und einem

[60] Qualitätskriterium im Bereich der Datenerhebung nach Döring, Bortz 2016a: 94ff.
[61] Deutsche Gesellschaft für Pflegewissenschaft - Ethikkommission
[62] u.a. Mitglied der Ethikkommission der DGP
[63] darüber hinaus wurde Döring, Bortz 2016c: 122ff.; Monteverde, Kesselring 2012: 142ff.; Monteverde 2012 bei der ethischen Betrachtung mitberücksichtigt.

zusätzlichen individuellen, spezifischen Einwahlkode versehen war. Es sind folglich weder Rückschlüsse auf die jeweilige Schuleinrichtung noch deren Auszubildende in der Gesundheits- und Krankenpflege bzw. TN möglich. Die erhobenen Rohdaten werden unter Verschluss verwahrt und stehen in vollständiger Form nur der Forscherin und ihrem Projektbetreuer zur Verfügung.

Eine Beeinträchtigung oder Schädigung durch den technologiebasierten, computergestützten Kompetenztest wird ausgeschlossen, da die vorhandenen Videosequenzen, Image-Maps usw. den Alltag eines Auszubildenden im dritten Lehrjahr abbilden sowie den Alltag einer zukünftigen examinierten Pflegekraft real darstellen. Eine zusätzliche Sicherheit bildet für den TN, dass die Datenerhebung in den jeweiligen gewohnten Räumlichkeiten der Bildungseinrichtungen stattfindet. Als präventive Maßnahme wurde zur endgültigen Verhinderung von persönlichen Vulnerabilitäten jeder TN am Ende der Kompetenzmessung von der Forscherin nach persönlichem Befinden und gewonnenen Erfahrungen mit dem „TEMA“-Testinstrument befragt. Sollte es folglich zu persönlich negativen Emotionen oder sonstigen Auswirkungen kommen, besteht zunächst die Möglichkeit, ein ausführliches Feedbackgespräch mit der Forscherin zu führen. Falls weitere Maßnahmen notwendig wären, wird dem TN die Möglichkeit eröffnet, sich an den Vertrauenslehrer der Schuleinrichtung oder auf neutraler Ebene an einen Psychologen zu wenden. Darüber hinaus kann der TN auch nach Abschluss der Kompetenzmessung jederzeit den Kontakt zur Forscherin suchen (Anhang 2).

Als Fazit zeigt sich, dass die Durchführung eines ethischen Clearings für das Projekt nicht notwendig ist, da keine persönlichen Gefährdungskriterien für die TN gegeben sind. Somit entfällt ein Antrag zur Begutachtung durch die Ethikkommission, da kein Mehrwert für den weiteren Projektverlauf besteht.

7.1.2 Gütekriterien der quantitativen Forschung

Die empirischen Bewertungskriterien der Objektivität, Reliabilität, Validität werden im Kompetenzprojekt überprüft und im gesamten Forschungsprozess, beispielsweise bei der Datenanalyse des Messinstrumentes, in den jeweiligen Kapiteln beantwortet. Zum konkretisierten Verständnis erfolgt ein kurzer Abriss über die essentiellen genannten Gütekriterien.

Unter „Objektivität“ oder auch „intersubjektiver Vergleichbarkeit“ ist zu verstehen, dass, egal welcher Forscher bzw. Forschergruppe, unter Annahme der identischen Bedingungen innerhalb eines Tests immer wieder dieselben Ergebnisse erzielt werden. Eine Standardisierung, die sich nach Moosbrugger & Kelava in Durchführungs-, Auswertungs- und Interpretationsobjektivität unterteilt, ist deshalb erforderlich. Ohne Objektivität kann keine Reliabilität und Validität erzielt werden (vgl. Moosbrugger & Kelava 2012: 8ff.; vgl. Hussy 2013: 23f.).

Hohe „Reliabilität“ zeichnet sich durch zuverlässige und beständige Messergebnisse aus. Es liegen nur geringe bis keine Messfehler vor. Die Angabe erfolgt im sogenannten „Reliabilitätskoeffizienten“ mit einem Wert zwischen 0 und 1. Ab Werten von 0.7

ist der Test als positiv zu bewerten. Umso mehr sich der Wert der Zahl 1 nähert, je weniger Messfehler sind vorhanden[64] (vgl. Moosbrugger & Kelava 2012: 11ff.).

Bezogen auf die Durchführung eines Kompetenztests heißt das, dass bei vorhandener Objektivität und hoher Reliabilität überhaupt erst eine signifikante Validität ermittelbar ist (vgl. Döring, Bortz 2016a: 95).

Vorliegende „Validität“ beschreibt die „Gültigkeit“ eines Tests. Es wird dabei untersucht, ob ein statistisches Messinstrument wirklich das gemessen hat, was, bezogen auf den Kompetenztest, die „praktische Pflegekompetenz“, ermittelt werden soll.

Die Facetten der Validität auf Grundlage der Untersuchungen von Campbell[65] werden überprüft durch die Konstruktvalidität[66], interne Validität[67], externe Validität[68] und statistische Validität[69] (vgl. Döring, Bortz 2016a: 97ff.). Beispielsweise kann durch einen zu niedrigen Stichprobenumfang, wie heterogene TN bei der Kompetenzerfassung oder unterschiedliche Rahmenbedingungen der TN im Testfeld, eine ausreichende statistische Validität verhindert werden.

Die Überprüfung der Gütekriterien im Ursprungsprojekt „TEMA“ wurden bereits im Kapitel 4.2 erläutert und zeigen in der Gesamtbetrachtung der Datenauswertung ein anspruchsvolles, aber valides pflegerisches Kompetenzmessinstrument für die Berufsgruppe der Altenpflege.

Im nachstehenden Forschungsprojekt wurden die genannten Gütekriterien einbezogen und berücksichtigt. Die Bewertungen und Spezifikationen finden sich u.a. in Kapitel 14.3 und Kapitel 15.4.

---

64 Es existiert eine hohe Messgenauigkeit! Der Test misst die latente Variable, wie im Projekt „TEMA“, die der praktischen Pflegekompetenz!

65 bereits verstorbener Psychologe und methodisch versierter Wissenschaftler Donald T. Campball (vgl. Campbell 1957: 297ff., vgl. Döring, Bortz 2016a: 93).

66 „construct validity“ -> auf Grundlage von „Konzeptspezifikation“ und „Operationalisierung“ (Döring, Bortz 2016: 97). Im Projekt „TEMA“ wurde u.a. ein „heuristisches Modell“ entwickelt und durch Einsatz eines faktorenanalytischen Verfahrens eine Modellierung von Testpaketen vorgenommen.

67 „internal validity“ -> auf Grundlage von „Untersuchungsdesign“ und seiner „Anwendung“ im Untersuchungsfeld (Döring, Bortz 2016: 97).

68 „external validity“ -> auf Grundlage von „Untersuchungsdesign“ und gewählter „Stichprobenziehung“ mit Übertragung auf beispielsweise andere Berufsgruppen (Döring, Bortz 2016: 97).

69 „statistical inference validity“ -> auf Grundlage von „statistischer Datenanalyse“ und „forschungsbezogenen Aspekten“ (u.a. gewählte Stichprobengröße) (Döring, Bortz 2016: 97).

7.1.3 Bestimmung der Stichprobengröße

Auf Grundlage des Ursprungsprojekts „TEMA“ mit einer Stichprobengröße von 402 TN[70] im Rahmen der Haupterhebungsphase und der Möglichkeit des validen Vergleichs von ermittelten Parametern wurde die Stichprobengröße von 350 bis 400 Teilnehmer festgelegt (vgl. Kaspar et al. 2016: 1; vgl. Döring et al. 2016: 252).

Unter Berücksichtigung der späteren methodischen Vorgehensweise und Datenanalyse ist an dieser Stelle auf die „Simple Size“- Bestimmung von Linacre hinzuweisen. Beim Schätzen der Fallzahl von Rasch-Modellen steht nicht die TN-Zahl, sondern die Stabilität des Modells und seiner Passung im Mittelpunkt. Er weist in diesem Zusammenhang auf die „Blindheit“ von Person und Item hin.

> „The requirements are symmetric for the Rasch model so you need as many items for a stable person measure as you need persons for a stable item measure. Consequently, 30 items administered to 30 persons (with reasonable targeting and fit) should produce statistically stable measures.“ (Linacre 1994: 328).

Für die durchzuführende Studie bedeutet das, dass bei 76 verwendeten Items mindestens 76+ TN notwendig sind.

Zusätzlich wurde zur Überprüfung der bestimmten Stichprobengröße eine „Sample Size“-Berechnung durchgeführt. Hierzu wurde die folgende Formel verwendet (Gl.1):

$$\textit{sample size} = \frac{[z^2 * p\,(1-p)]}{e^2}$$

z = Z-Wert oder Konfidenzniveau
p = p-Wert oder Standardabweichung
e = Fehlermarge oder Konfidenzintervall

Bei Annahme von verwendeten statistischen Durchschnittswerten mit einem Konfidenzniveau von 95% und einem daraus resultierenden Z-Wert von 1,96, einem p-Wert von 0,5 (entspricht 50%) und einem Konfidenzintervall[71] von 0,05 ergibt sich eine ermittelte Stichprobengröße von 384 TN. Daraus folgt, dass eine mögliche Gewinnung von statistisch signifikanten Effekten mit einer Stichprobengröße zwischen 350 und 400 TN erreicht wird und demzufolge als realistisch bzw. ausreichend gilt (vgl. Bortz, Lienert 2003: 48ff.; vgl. Bortz, Schuster 2010: 97ff.; vgl. Koller 2014: 23ff.). Die gewonnene TN-Zahl aus dem Ursprungsprojekt „TEMA“ wird bestätigt!

[70] im dritten Jahr der Altenpflegeausbildung
[71] Weitere statistische Analysen zum Konfidenzintervall finden sich in Kapitel 15.1 und Kapitel 16.

8 Rahmenbedingungen für den Feldzugang

Im Vorfeld der Datenerhebung waren noch verschiedenste Aspekte für die Feldstudie[72] vorzubereiten und abzuklären, wie u.a. der Gewinnung von Bildungseinrichtungen, Beschaffung des Forschungsequipments oder monetärer Aufwand. In den folgenden Unterkapiteln wird über die wichtigsten Rahmenbedingungen berichtet.

8.1 Kontextuelle Rahmenbedingungen

An erster Stelle stand die Akquise der Schuleinrichtungen mit deren Auszubildenden bzw. Studierenden für die Durchführung einer technologiebasierten Kompetenzmessung. Im Kapitel 9.1 wird darüber ausführlich berichtet.

8.2 Technische Rahmenbedingungen

Für die Durchführung der Kompetenztestungen wird der speziell konfigurierte Server mit seiner vorhandenen Software aus „TEMA“ benötigt. Das Testinstrument wurde zur Ausleihe am Forschungsinstitut für Betriebliche Bildung in Nürnberg schriftlich und telefonisch angefragt und die Zustimmung zur Verwendung im Feld der Gesundheits- und Krankenpflege sowie der erforderlichen Datenerhebung und Datenauswertung eingeholt. Eine technische, computergestützte Einweisung einschließlich des Auf- und Abbaus in den Server fand vor Ort mit einem dortigen IT-Fachmann statt (Anhang 5 u. 6).

Darüber hinaus wurde weiteres Forschungsequipment zur allgemeinen und technischen Datengewinnung benötigt. Im Mittelpunkt stand dabei die Ausleihe von 20 Notebooks für die TN[73]. Weiterer Bedarf an Hilfsmitteln bestand in zusätzlichen Adaptersteckern, Verlängerungskabeln, Mehrfachsteckdosen, Kabeltrommeln, optischen Kabelmäusen, Mousepads, Kopfhörer bzw. Earphones, Schubwagen, Kugelschreiber etc. (Anhang 7). Die weiteren Hinweise und Herausforderungen zur Technik sind in Kapitel 10 u. 10.1 dargestellt.

8.3 Finanzielle Rahmenbedingungen

In der ersten Welle der Datenerhebung lagen die finanziellen Kosten hauptsächlich in den Reisekosten zu den einzelnen Schuleinrichtungen und der wiederholten Abholung, Rücktransporten und Rückgabe der geliehenen Notebooks.

In der zweiten Welle war kein Verleih mehr möglich, deshalb entstanden neue, zusätzliche Kosten von mehreren Tausend Euros durch die Anschaffung des in Kapitel 8.2 erwähnten Equipments (Anhang 7). Die Versuche, einen finanziellen Sponsor für das Projekt zu gewinnen, waren nicht erfolgreich. Zusätzlich fielen zu den bereits angesprochenen Reisekosten noch weitere Kosten für Übernachtungen an. Summa summarum heißt das, dass die gesamten Kosten von der Promovendin zu tragen waren.

---

[72] auch genannt „field study“

[73] In der ersten Welle der Datensammlung wurden über zweidrittel der Notebooks kostenlos von der Evangelischen Hochschule Ludwigsburg für die jeweiligen Testtage zur Verfügung gestellt (siehe ergänzend dazu noch Kapitel 9.4).

8.4 Förderungen und Hemmnisse im Rahmen der Projektarbeit im Feld

Als positive, konstruktive Unterstützer zeigten sich zu fast jeder Zeit Personen aus den jeweiligen „TEMA“ - Projektteams. Eine Annäherung und Übertragung in das Berufsfeld der Gesundheits- und Krankenpflege wäre, wie bereits erwähnt, alleine durch die literarischen Erarbeitungen sonst enorm erschwert bis nahezu unmöglich und somit nicht umsetzbar gewesen. Die Möglichkeit für die Durchführung eines weiteren empirischen, berufsorientierten Forschungsvorhabens wurde von den Akteuren aus „TEMA“ und der Evangelischen Hochschule Ludwigsburg unterstützt.

Als belastende Faktoren innerhalb der Feldphase zeigten sich die bereits erwähnten hohen Kosten im Feld (siehe Kapitel 8.3) und der enorme zeitliche Aufwand zur Gewinnung von Bildungseinrichtungen einschließlich Vorbereitung, Durchführung und Nachbereitung der einzelnen Kompetenzmessungen.

9 Untersuchungsdesign

In der vorliegenden Arbeit wird eine Primärdatenerhebung im Feld mit anschließender deskriptiver und explanativer, statistischer Datenauswertung durchgeführt. Die systematische Replikationsstudie wird unter Diversifikation auf die Gesundheits- und Krankenpflege, ohne Messwiederholungen, in den jeweiligen TN-Kohorten realisiert. Als Hilfsmittel werden diverse Fachartikel und der Austausch mit dem „TEMA“ - Projektteam mit einbezogen (vgl. z.B. BMBF 2012, 2014 u. 2015; vgl. Dietzen et al. 2016; vgl. Beck, Landenberger, Oser 2016).

Diese geplante weitere, empirische, berufsorientierte Feldstudie zeigt sich deshalb als geeignet, weil durch das reale und gewohnte Umfeld der TN lediglich verminderte Störvariablen für die Interpretierbarkeit der zu erhebenden Forschungsdaten vorhanden sind.

Die Forschungsfrage verlangt, den Schwerpunkt der Arbeit auf den pflegerischen Kontext des technologiebasierten, computergestützten Kompetenztests aus „TEMA“, ergänzt mit den eingesetzten Items aus „CoSMed“, zu legen (siehe Kapitel 4.2 und 4.3). Vorab wurde zum Einstieg eine freiwillige, verkürzte „Paper-Pencil“-Erfassung der personenbezogenen Daten (siehe Anhang 3) auf Grundlage von „SiKoFak“ eingesetzt (vgl. Baethge-Kinsky, Baethge, Lischewski 2016: 265ff.).

Zur Erprobung und breiten Verwendung des Kompetenzmessinstrumentes wurde im nationalen Kontext nach Einsatzorten gesucht. Die Datenerhebung wurde nicht wie in „TEMA“ auf zwei Bundesländer beschränkt, damit möglichst viele geeignete Bildungseinrichtungen gefunden wurden, aber immer unter Berücksichtigung, der verschiedenen schulischen Facetten, wie z.B. schuleigenen Curricula.

9.1 Einzugsgebiet

Die schriftlichen[74] und telefonischen Anfragen an einzelne Hochschulen und Bildungseinrichtungen fanden im süddeutschen Raum statt, d.h. in den Bundesländern Bayern, Baden-Württemberg, Hessen, Rheinland-Pfalz, Sachsen, Saarland, Thüringen. Die Auswahl der Einrichtungen erfolgte über Internet, persönliche Kontakte und Empfehlungen von Lehrkräften, bei denen bereits eine Kompetenzmessung durchgeführt worden war. Insgesamt wurden über 60 Bildungseinrichtungen zur Durchführung einer Kompetenzmessung angefragt. Die Zusage zur Testung in den einzelnen Erhebungswellen gaben insgesamt 5 +1[75] Hochschulen und 16 + 1 Bildungseinrichtungen in der „Gesundheits- und Krankenpflege“.

Im Vorfeld erhielten die Interessenten über ihre jeweilige Bildungseinrichtung ein Informationsschreiben mit der Bitte und Anregung zur Teilnahme an der Kompetenztestung in den Räumlichkeiten ihrer eigenen Schuleinrichtung (Anhang 1 Informationsschreiben an die Schuleinrichtungen).

[74] per Mail
[75] +1 bedeutet, dass an einer Schule und Hochschule 2 Messungen bei 2 Gruppen stattgefunden haben.

9.2 Teilnehmende Hochschulen und Bildungseinrichtungen

Die Anzahl der teilgenommenen Bildungseinrichtungen je Bundesland zeigt u.a. die Kartenübersicht (Abb. 5).

Abb. 5: Anzahl der teilgenommenen Bildungseinrichtungen (Deutschlandkarte, entnommen aus Internet: http://www.weltkarte.com/europa/landkarten-und-stadtplaene-von-deutschland/landkarte-deutschland-politische-kartebundes-laen-der.htm [Stand: 03.05.2019] und leicht modifiziert von Ries 2019).

Von den 21 verbindlich zugesagten Bildungseinrichtungen hat eine nach Anreise mangels TN kurzfristig abgesagt. Im Gesamtergebnis haben somit 20 Bildungseinrichtungen[76] teilgenommen, es wurden 22 Messungen durchgeführt.

Ein separates Anschreiben wurde im Jahr 2018 per Mail an die einzelnen Schulen zur Anfrage auf Zustimmung zur Veröffentlichung der Namensnennung der Bildungseinrichtung verschickt. Dies hatte den Grund, dass der umfangreiche Erhebungsradius, verbunden mit dem Faktor Zeit und darüber hinaus die sehr kooperative Unterstützung und Zusammenarbeit, herausgestellt werden sollte. Am computergestützten Kompetenztest in der Gesundheits- und Krankenpflege haben u.a. Bildungseinrichtungen teilgenommen:

[76] in zwei Bildungseinrichtungen zweimal Anfahrt zur Kompetenzmessung bei unterschiedlichen TN

- Berufsfachschule für Gesundheits- und Krankenpflege Donauwörth
- Evangelisches Bildungszentrum Stuttgart
- Schule für Pflegeberufe am Klinikum Crailsheim
- Diakonie-Klinikum SHA gGmbH - Schule für Pflegeberufe Schwäbisch Hall
- Gesundheitsakademie Bodensee-Oberschwaben GmbH - Schulen für Pflegeberufe Weingarten
- Krankenpflegeschule der Rhein-Mosel-Fachklinik Andernach
- Gesundheits- und Krankenpflegeschule Künzelsau
- Bildungszentrum für Gesundheits- und Krankenpflege der DRK Krankenhaus GmbH RLP Hachenburg
- SLK-Klinikum - Gesundheitsakademie Heilbronn
- Klinikum Esslingen - Schule für Pflegeberufe
- Berufsfachschule für Krankenpflege am Krankenhaus Martha-Maria Nürnberg
- Katholische Hochschule Freiburg
- Evangelische Hochschule Ludwigsburg
- Technische Hochschule Deggendorf

Weitere Informationen zu den einzelnen Bildungseinrichtungen können im Anhang 8 u. 9 nachgelesen werden.

Der Datenschutz der einzelnen TN einschließlich Details der einzelnen Kompetenzmessungen des Forscherteams bleibt weiterhin berücksichtigt und beruht nicht auf einer möglichen Veröffentlichung.

9.3 Ein- und Ausschlusskriterien der Teilnehmer

Genauso wie die Gewinnung der Schuleinrichtungen war die Auswahl geeigneter TN für die Kompetenztestung zu berücksichtigen. Die Kontaktaufnahme und Anfrage wurde mit Unterstützung der Bildungseinrichtungen ermöglicht.

Tab. 2: Teilnehmerauswahl

Einschlusskriterien:	Ausschlusskriterien:
Hauptfokus: Auszubildende und „Dual“ Pflege Studierende im dritten Ausbildungsjahr in der Gesundheits- und Krankenpflege (im dualen Studium auch Auszubildende in der Altenpflege)	Auszubildende in der Altenpflegehilfe oder Gesundheits- und Krankenpflegehilfe oder auch andere Heilberufe (z.B. Rettungsassistent)
Auszubildende und „Dual“ Studierende im dritten Ausbildungsjahr bzw. im jeweiligen Semester der Abschlussprüfung in der beruflichen Praxis	Auszubildende und „Dual“ Studierende in den ersten zwei Ausbildungsjahren bzw. 3. und 4. Semester an der Hochschule
zusätzlich: Gesundheits- und Kinderkrankenpflege, da teilweise gemeinsame Ausbildung mit der Gesundheits- und Krankenpflege in den Bildungseinrichtungen	Kohorten in der Gesundheits- und Kinderkrankenpflege
Auseinandersetzung in Theorie und Praxis in verschiedenen Pflegesettings im Rahmen der dreijährigen Gesundheits- und Krankenpflegeausbildung	Theoretischer Unterricht und Praxiserfahrungen sind noch nicht in unterschiedlichen Pflegesettings erfolgt, wie z.B. im Bereich der stationären und ambulanten Pflege
Verstehen der deutschen Sprache in Wort und Schrift[77]	Sprachkenntnisse in der deutschen Sprache sind nicht vorhanden
Computergrundkenntnisse sind vorhanden	Computerkenntnisse fehlen vollständig

Insgesamt haben sich 408 TN am Kompetenztest beteiligt. Zwei Testungen haben nicht den vollständigen Einschlusskriterien entsprochen (nicht im dritten Ausbildungsjahr, zeitliche Begrenzung des „Powertests“ fehlt). Der organisatorische Rahmen wurde vom Forscherteam vor der Haupterhebungsphase der Kompetenztests neu justiert (wie z.B. Bildschirme der einzelnen Notebooks stehen gleich auf der Startseite zur Einwahl in den eigentlichen Kompetenztest). Nach Abzug der 19 TN aus Pretestungen ergeben sich 389 TN in der Haupterhebungsphase.

[77] Die Deutschkenntnisse wurden im Bewerbungsverfahren auf einen Ausbildungsplatz bzw. Studienplatz durch Einreichung eines Sprachnachweises auf B2 bzw. C1 Niveau von den Schuleinrichtungen abgefragt.

9.4 Zeitlicher Rahmen für die Datengewinnung

In der Ursprungsstudie wurde die Haupttestung von März bis Juni 2014 durchgeführt (vgl. Döring et al. 2016a: 252). Die Datenerhebung für diese Arbeit fand in zwei Wellen statt, die erste von Anfang Juni bis Ende Juli 2017, die zweite von Mitte Oktober bis Mitte Dezember 2017. Dies war notwendig, weil die Gewinnung der einzelnen Schulen bzw. Hochschulen einen teils hohen zeitlichen Aufwand und Vorlauf[78] erforderte und gleichzeitig noch die Voraussetzung, dass die TN sich auch immer im dritten Ausbildungsjahr bzw. im dualen Studium im 5. oder 6. Semester[79] befinden sollten. Die vorgeschlagenen TN zum Kompetenztest in den einzelnen Bildungseinrichtungen sollten im Rahmen ihrer Gesundheits- und Krankenpflegeausbildung mit den Settings „Wohngruppe“, „Ambulante Pflege“ und „Stationäre Pflege“ in einer theoretischen oder praktischen Form in Kontakt gekommen sein (vgl. Döring et al. 2016b: 124).

[78] TN befinden sich in ihrer Ausbildungs- oder Hochschuleinrichtung im Blockunterricht oder haben Schul- bzw. Studientag.

[79] Hier wird vorausgesetzt, dass der Ausbildungsbeginn der Studierenden im dualen Studiengang Pflege mit gleichzeitigem Beginn von Lehre und Studium, das 5. oder 6. Semester erreicht hat. Der Wissensstand der Semester im dritten Ausbildungsjahr kann je nach Hochschule variieren.

10 Datengewinnung

Im folgenden Kapitel werden die einzelnen Schritte der gesamten Datengewinnung einschließlich der Aufgaben der TN und des Forschungsteams erläutert.

Bevor der erste Pretest stattfinden konnte, wurde das komplette Equipment in einem Lehrsaal zur Probe für die kommenden Kompetenztestungen aufgebaut und eine zeitliche und technische Einschätzung für die zukünftigen Messungen vorgenommen. Die zeitliche Einschätzung beläuft sich für eine komplette Erhebung im Durchschnitt mit Auf- und Abbau auf ca. 5 bis 6 Stunden. Davon sind 2 Stunden für die eigentliche Kompetenzmessung eingeplant.

10.1 Pretests

Vor Beginn der Haupttestungsphase fand an einer Hochschule und an einer Bildungseinrichtung für Gesundheits- und Krankenpflege je eine Probemessung statt, welche nicht den vollständigen Einschlusskriterien entsprochen haben (siehe auch Kapitel 9.3). Es zeigten sich außerdem technische Probleme mit den Notebooks bei Einwahl in den externen „TEMA“-Server bei mehr als 10 TN. Erst durch den Einsatz von Zusatzadaptern und Einwahl in ein zweites Netzwerk[80] konnte dies reguliert werden (Anhang 10).

Die Sicherstellung der Stromversorgung durch die erforderliche Verkabelung der Notebooks mit Netzladegeräten und Mehrfachsteckdosen hin zu den jeweils vorhandenen stromführenden Wandsteckdosen erwies sich als sehr umfangreich und aufwendig[81]. Deshalb wurden noch vier zusätzliche Kabeltrommeln und vier Sechsfachsteckdosen angeschafft.

10.2 Haupterhebungsphase und Ablauf der Kompetenzmessungen

Nach Anpassung aller Auffälligkeiten aus den Pretests begann die eigentliche Erhebungsphase im Untersuchungsfeld. Das durchgängige Forscherteam für die Datenerhebung bestand aus einem IT-Fachmann und der Promovendin. Die erste und zweite Welle der Haupterhebung verlief dabei immer nach dem gleichen Muster, das jetzt in seiner jeweiligen Abfolge beschrieben wird.

[80] Netzwerke - „na“ und „ng“; bei veralteten Notebooks nur „ng“ vorhanden; die Laptops mussten einwahlfähig in die Netzwerke Juniper „na“ und „ng“ sein, damit eine Verbindung in das jeweilige Testheft möglich wurde und eine Stabilität im Testheft beim Abspielen der einzelnen pflegerischen Filmsequenzen vorhanden war. Bei Laptops, die älter als 5 Jahre waren, musste zusätzlich ein spezieller steckbarer W-LAN-Stick angebracht werden, damit beide Netzwerke (nicht nur Juniper „ng“!) sichtbar wurden.

[81] Bei einer Bildungseinrichtung in einem renovierten Altbau musste das Stromnetz zusätzlich abgesichert werden.

10.2.1 Einführungsvortrag in die Kompetenzmessung

Zu Beginn jedes „Speedtestings“[82] fand ein kurzer Einführungsvortrag über das Ursprungsprojekt, die Weiterführung in der Gesundheits- und Krankenpflege sowie den Ablauf und Hinweise zur Kompetenzerfassung statt. Es konnten im Vorfeld Fragen von den TN gestellt werden, was jedoch nur vereinzelt vorkam. Es handelte sich dabei um organisatorische Fragen wie die Dauer der Messung, Möglichkeit des Toilettengangs während der Messung usw. Zu Inhalten innerhalb des Kompetenztests wurden keine Hinweise oder Erläuterungen verbalisiert.

An jedem Notebookplatz wurden bereits nach dem Aufbau des IT-Equipments zwei Teilnehmerbögen und ein Kuvert mit einem Einwahlcode in den Kompetenztest ausgelegt (Anhang 11). Die Anonymisierung der erhobenen Daten war somit gewährleistet.

10.2.2 Einverständniserklärung und Datenschutzerklärung

Zuerst wurden die TN gebeten, die Erklärung zu lesen und anschließend mit Ort und Datum zu unterschreiben. Dies geschah freiwillig und mit dem Hinweis auf den jederzeit möglichen Abbruch des Kompetenztests. Alle TN haben zugestimmt und dies mit ihrer Unterschrift bestätigt. Keine Person hat wegen der Kompetenzerfassung den Test abgebrochen.

10.2.3 Erhebung der personenbezogenen Daten

Die TN wurden gebeten, den zweiten einseitigen Bogen durchzulesen, nach Möglichkeit eines der aufgeführten, betreffenden Merkmale zu kennzeichnen und bei Bedarf Ergänzungen vorzunehmen oder auch zu streichen (Anhang 3). Drei TN haben den Bogen nicht abgegeben. Dies geschah innerhalb des Pretests ohne Angabe von Gründen.

10.2.4 Durchführung der „Technologiebasierten“ Kompetenzmessung

Der Beginn des Kompetenztests erfolgte durch die persönliche Code-Einwahl des Teilnehmers in die Lernplattform „ILIAS“[83] in der Version 4.2.5. Der Code konnte nur einmal vergeben werden, oder es musste eine Löschung der bereits erhobenen Daten direkt am externen zentralen „TEMA“-Server vorgenommen werden.

Es standen zwei Testhefte aus dem „TEMA“-Ursprungstest[84] zur Verfügung.

Der eigentliche computergestützte Kompetenztest dauerte pro Messung bis zu zwei Stunden (= Power Testung). Die TN, die länger für die Beantwortung des jeweiligen

---

[82] zeitliche Begrenzung der Bearbeitung des jeweiligen Testheftes vorgegeben

[83] Integriertes Lern-, Informations- und Arbeitskooperations-System; geeignete freie Software zum Einpflegen und Einsetzen von E-Learning-Einheiten beispielsweise an Hochschulen in Deutschland

[84] Testheft 2 und 3; durch die bis zu zweistündige Dauer des Tests kann es zu Konzentrationseinbußen und leichter Müdigkeit kommen. Deshalb bearbeitete die eine Hälfte der Teilnehmer den Test von vorne nach hinten, die andere von hinten nach vorne. Eine Stabilität im Test wurde hierdurch angestrebt.

Testheftes benötigten, wurden aufgefordert, diesen zu beenden, und die nicht beantworteten Items wurden als nicht gelöst (= 0) später codiert.

In den gesamten 389 durchgeführten Kompetenzmessungen gab es drei Zeitüberschreitungen. Der Gesamttestzeitraum lag zwischen 01:03 Stunden und 02:17 Stunden. Die einzelnen Items wurden in einer Zeitspanne von 0:21 Minute bis 1:08 Minuten beantwortet.

Zur Beantwortung der einzelnen Items durften keine Hilfsmittel verwendet werden. Auch wurden keine inhaltlichen Fragen vom Forscherteam zum Kompetenztest beantwortet (vgl. Ursprungstest „TEMA“ Kapitel 4.2; vgl. Anhang 11).

Der eigentliche Kompetenztest gliederte sich in die bereits erwähnten drei Settings: „Wohngruppe“, „Ambulante Pflege“ und „Stationäre Pflege“ (vgl. Döring et al. 2016b: 124; vgl. Döring, Möllers, Schöpf 2014: 41), ergänzt mit 12 Items aus „CoSMed“. Alle weiteren Aspekte zum Inhalt des Kompetenztests wurden bereits in den Kapiteln 4.2 und 4.3 angesprochen.

10.2.5 Beendigung der Kompetenzmessung vom TN und Forschungsteam

Sobald der TN am Ende des Testheftes innerhalb des „TEMA“-Kompetenztests angekommen war, wurde von der Forscherin die Sichtung und Auslösung des Abschlussbuttons auf dem jeweiligen Laptop überprüft. Dies war deshalb so wichtig, weil dadurch die automatische Datensicherung der gewonnenen Ergebnisse auf dem externen „TEMA“-Server erfolgreich ausgelöst wurde. Eine Einsicht in die erhobenen Daten hatte ausschließlich die Promovendin, wenn die Aktivierung des Servers veranlasst wurde.

Mit jedem einzelnen TN fand nach jeweiliger Beendigung des Kompetenztests ein kurzes Feedbackgespräch über das persönliche Befinden und den Eindruck über die Testung statt. Die Aussagen waren überwiegend sehr positiv. Wesentliche Zitate sind: „Ist endlich mal was Neues!“; „Hat mir was gebracht!“; „Sollte man öfters machen!“; bis hin zu: „Jetzt brauche ich erst mal eine Pause!“; „Test ist schon sehr lang!“; „Würde ich gerne wieder machen!“; „Schade, dass wir nicht gleich ein Ergebnis zum Test bekommen!“; usw.

Die abschließende Aufgabe des Forscherteams bestand nach jeder durchgeführten Kompetenzmessung in einem durchschnittlich einstündigen Abbau des gesamten Forschungsequipments.

Nach Abschluss aller Kompetenzmessungen fand die Endabspeicherung der gesamten gewonnenen „Teilnehmer-Kompetenzdaten“ von der Lernplattform ILIAS auf zwei externe USB[85]-Sticks und zusätzlich auf einen separaten Laptop der Forscherin statt. Nachdem alle erhobenen Daten vom geliehenen externen, zentralen „TEMA“-Server gelöscht waren, wurde er dem f-bb in Nürnberg wieder zurückgegeben.

[85] Universal Serial Bus

11 Abschließender Kommentar zum „Empirischen Teil 1“ - Datenerhebung

Der „TEMA“-Kompetenztest zeigt sich im organisatorischen Rahmen des Aufbaus, der Durchführung und des Abbaus als zeitlich und technisch sehr aufwendig. Inhaltlich ist er für die einzelnen TN eine Bereicherung für die bevorstehende praktische Prüfung in der Gesundheits- und Krankenpflege und ergänzend sogar für die Gesundheits- und Kinderkrankenpflege, trotz der Fokussierung auf „alte Menschen“. Die gewonnenen 408 bzw. 389 TN-Datensätze haben einen vergleichbaren Umfang wie in „TEMA“ erreicht und entsprechen der Schätzung einer ermittelten „Sample Size“ von 384 TN. Eine valide, statistische Evaluation kann daraus gefolgert werden.

III Empirischer Teil 2 - Statistische Aufbereitung und Datenanalyse

Im Teil III wird die Vorgehensweise und Darstellung der Datenaufbereitung für die statistischen Auswertungen, verwendete Software- und Statistikprogramme, Inhalte und Ergebnisse der personenbezogenen und kompetenzbezogenen TN-Daten aus den Kompetenzmessungen, Erläuterungen zu verwendeten quantitativen Theorien und Modellen einschließlich der statistischen Resultate aus der Item-Response-Theorie und der Konfirmatorischen Faktorenanalyse aufgezeigt.

12 Statistische Aufbereitung der gewonnenen Daten

Zunächst war eine spezielle Aufbereitung des gewonnenen Rohdatenmaterials der Gesamt-TN notwendig, damit eine anschließende Datenauswertung bzw. Datenanalyse überhaupt möglich wurde. Diese zeigte sich ebenfalls als sehr aufwendig und komplex. Der kalkulierte Aufwand einschließlich der Neucodierung von ursprünglich geplanten ca. 4 Wochen, dehnte sich letztendlich auf mehr als 14 Wochen aus. Grund dafür war weitere multiple und intensive Literaturrecherchearbeit über das Ursprungsprojekt, um die notwendigen Skalenniveaus[86] der einzelnen Items aus dem vorhandenen Textmaterial aus „TEMA“ herauszufinden (Anhang 14). Gleichzeitig wurden noch mehr Inhalte des Ursprungsprojektes, wie beispielsweise aus „CoSMed“, inhaltlich aufbereitet und mit den erhobenen Rohdaten in das Dissertationsprojekt überführt (vgl. Kaspar et al. 2016; vgl. Döring et al. 2014; vgl. Wittmann, Kaspar, Döring 2017: 185ff.).

12.1 Software: „Open Office Calc“ und „Microsoft EXCEL“

Die erhobenen Rohdaten standen nach den durchgeführten Kompetenzmessungen auf dem „TEMA“-Server in einer kostenfreien Version von „Open Office Calc“[87] zur Verfügung. Die Abspeicherung dieser Tabellen-Rohdaten[88] erfolgte im „.ods-Format“[89], damit das Einlesen in unterschiedliche Software- und Statistikprogramme, wie beispielsweise „Microsoft EXCEL“ und folgend in SPSS, erleichtert bzw. möglich wurde (vgl. interne Unterlagen auf „TEMA“ Server; vgl. Döring et al. 2014).

Die anschließende vollständige Anonymisierung des Datensatzes erfolgte nach Konvertierung in „Microsoft EXCEL“ durch das Streichen der Usernames (z.B. km_001) im Rahmen der Einwahlcodes der einzelnen Teilnehmer. Die Itembezeichnungen wurden im Verlauf der weiteren Aufbereitungen spezifiziert aufgeführt, um eine verbesserte Übersichtlichkeit für die später folgende Datenauswertung bzw. Dateninterpretation zu erhalten (z.B. aus Item „A34“ wird Item „DR21A34_polytom“[90]).

---

86 unterschiedlich abgestufte Bewertungen der einzelnen Itemantworten - „Scoring“ in Anlehnung an „TEMA“ vorgenommen

87 vergleichbar mit der kostenpflichtigen Software: „Microsoft_Exel“ für Tabellenbearbeitungen, statistischen Auswertungen usw.; „Calc“ steht für Tabellenkalkulation

88 bestehend aus zwei „Open Office Calc“-Dateien bzw. Testheft 2 und 3

89 „Open Document Standard Format“; Gruppe der kostenfreien „Open Document Formate“

90 d.h. Teilkompetenz: Diagnostisch-Reflexiv21_Item A34 mit polytomem Beantwortungsmuster

Ein individuelles Scoring der einzelnen Itemantworten wurde in Anlehnung an „TEMA“ angelegt (Anhang 12 u. 14) und im vorhandenen „EXCEL“-Rohdatensatz korrekt angepasst bzw. aktuell codiert, d.h., die Skalierungen mussten für jedes einzelne Item festgelegt werden[91]. Es fanden nach Abschluss der Recodierung nochmals zwei Überprüfungen auf deren Richtigkeit statt.

12.2 SPSS und „Missing Values“

Bevor die ersten Datenergebnisse ermittelt werden konnten, wurde der bearbeitete EXCEL-Datensatz in das Statistikprogramm „SPSS“[92] übertragen. Die fehlenden Werte[93] wurden mit 0 im Datensatz ergänzt, da alle Items einer vollständigen Datenauswertung unterliegen sollen (Anhang 13). Dies ist deshalb so festgelegt, weil es wie in einer händisch durchgeführten, schriftlichen Klausur als „nicht beantwortet“ bzw. „falsch“ einzustufen ist. Gleiches gilt für nicht beantwortete Items bei Zeitüberschreitung des TN im Gesamtkompetenztest. Es werden keine Minuspunkte oder ein Zahlenwert außerhalb der Skalierung bei falscher Beantwortung in der Auswertung vergeben, wie teilweise von Autoren empfohlen (vgl. Kunina-Habenicht 2018a u. 2018b).

[91] 0, 1 oder 0, 1, 2 oder 0, 1, 2, 3 bei nicht Beantwortung des Items wurde ein N/A = not available = nicht verfügbar, gesetzt!

[92] Superior Performing Software System oder auch Statistical Package fort the Social Sciences

[93] Missing Values (in vorliegendem Projekt als systemdefinierte Missing Values, da die Angabe N/A aus „EXCEL“ nicht von SPSS erkannt und nur als ein Punkt in der jeweiligen Zeile dargestellt wird)

13 Verwendete Statistikprogramme für die Datenanalyse

Um eine umfassende Untersuchung des erhaltenen Datensatzes durchführen zu können, wurden verschiedene Statistikprogramme verwendet. Diese sollen kurz mit dem Bezug zum Einsatz im Forschungsprojekt erläutert werden, um weitere kompetenzorientierte Forschungsvorhaben für Dritte im Umgang mit genannter statistischer Software zu erleichtern, und mit dem Hinweis eines zusätzlich notwendigen Forschungsteams (Universität Göttingen) in „TEMA“. Die Herausforderung bestand hier in der Aneignung und Handhabung der einzelnen relevanten Programme für die notwendige Datenauswertung.

13.1 Auswertungen in SPSS

Wie bereits in Kapitel 12.1 und 12.2 erwähnt, wurde zur statistischen Analyse das kostenpflichtige Programm SPSS Statistics in der Version 25[94] verwendet. Es bietet multiple Möglichkeiten der Datenanalyse, wie beispielsweise differenzierte statistische Analysen in Form von Kennzahlen- und Häufigkeitsermittlungen und der graphischen Darstellung von statistischen Ergebnissen. SPSS zeigt nach Öffnen für den Einstieg in die Analysen in seiner Ansicht die Wahl zur Datenmatrix oder Variablenansicht und dient als Grundlage für die Berechnungen. Bei der Anwendung ist SPSS benutzerfreundlich und nicht nur für Profis geeignet. Weitere Hinweise zur Handhabung von SPSS werden in dieser Arbeit nicht gegeben, da dies über die Fragestellung hinausgeht, sind jedoch in einschlägiger Literatur nachzulesen (vgl. Koller 2014: 47ff.; vgl. Döring, Bortz 2016b: 618ff.).

In der vorliegenden Arbeit wurde SPSS verwendet für die Ermittlung von Häufigkeiten der gelösten und nicht gelösten Items aller TN, die erreichte Gesamtpunktzahl der TN in „TEMA“ und „CoSMed“ und passend zu diesen Ergebnissen die graphischen Darstellungen in Form von Balkendiagrammen. Weitere Vertiefungen erfolgen in Kapitel 14[95].

13.2 Auswertungen in R bzw. RStudio

Über das Statistikprogramm SPSS hinaus, wurde zur vertieften statistischen Datenauswertung und Datenvisualisierung das kostenfrei im Internet zur Verfügung stehende Programm R[96] für Windows in der Version 3.5.1[97] und seine entsprechende Entwicklungsumgebung RStudio[98] mit dem Vorteil der verbesserten Handhabung und übersichtlicheren, kompakteren Darstellung der Rechenoptionen heruntergela-

---

[94] Hinweis: Im weiteren Verlauf der Arbeit wird nur noch von SPSS gesprochen, es ist hier aber immer die benannte Version zu verstehen.

[95] Anmerkung: Bei den Auswertungen wurde zu Beginn das Ergänzungsprogramm SPSS AMOS Version 25 für die graphische Darstellung und Entwicklung von Strukturgleichungsmodellen versucht. Im Verlauf der Datenanalyse und den Datenauswertungen stellte sich das Programm R als die umfassendere Software dar mit übersichtlicheren graphischen Darstellungen der einzelnen Strukturmodelle.

[96] Das Programm R hat seinen Namen durch die beiden Vornamen der Entwickler Ross Ihaka und Robert Gentelman bekommen.

[97] unter http://www.r-project.org/

[98] unter https://www.rstudio.com/products/rstudio/download/

den und genutzt. Die Herausforderung bei diesen Programmen liegt im Erlernen von Befehlen und der Entwicklung der notwendigen Syntax für die gewünschten Rechenoptionen. Dafür bieten sie einen hohen Objektivitätsgrad durch die Festschreibung von R Skripten. Die Folge ist, dass ein reger, verlässlicher Austausch in der statistischen R Community entsteht. Gleichzeitig bieten sie im Rahmen von Projektaufträgen Nachvollziehbarkeit und Transparenz für das gesamte Projektteam und deren Auftraggeber (vgl. Wollschläger 2017: 5ff.; vgl. Koller, Alexandrowicz, Hatzinger 2012: 218ff.).

Zur Einarbeitung in die Software R und RStudio wurde diverse statistische Fachliteratur herangezogen und bearbeitet (vgl. Alexandrowicz 2013; vgl. Hatzinger et al. 2014; vgl. Wollschläger 2017; vgl. Koller, Alexandrowicz, Hatzinger 2012: 217ff.). Gleichzeitig wurden zum Ausbau der eigenen Fähigkeiten mehrere Schulungen in R besucht. Ein Eindringen in die Welt der statistischen R Community und deren Biometrikern wurde möglich. Das zeigte sich im Projekt der Gesundheits- und Krankenpflege als sehr hilfreich und positiv, gerade wenn Fragen der Forscherin zu R und deren Packages nicht durch autonome Evaluation und Interpretation zu klären waren. Genauso wurde mit erstellten R Skripten verfahren.

R kam im Projekt u.a. zum Einsatz bei der Modellgeltungstestung (siehe Kapitel 16.1 und 16.2) und der Beurteilung der Kompetenztestitems auf deren Schwierigkeit und der Fähigkeit der TN, die jeweiligen Items zu beantworten (siehe Kapitel 15.4). Ergänzend wurden zur verbesserten Darstellung mehrere graphische Übersichten erzeugt. Dafür stehen in R diverse Packages[99] zur Verfügung, die kurz Erwähnung finden sollen.

13.3 Verwendete Packages in R

Die notwendigen Packages für die durchzuführende, quantitative Datenermittlung in R bzw. RStudio sind in der vorliegenden Arbeit im Schwerpunkt:

- “ltm“ (Latent Trait Models under IRT Version 1.1-1; Rizopoulos 2018)
- “TAM” (Test Analysis Models Version 3.1-45; Robitzsch, Kiefer, Wu 2019)
- “Wright Map” (IRT Item-Person Map with ‘ConQuest’ Integration Version 1.2.1; Irribarra & Freund 2016)
- “foreign” (Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase', ... Version 0.8-71; R Core Team et al. 2018)
- “irtoys” (A Collection of Functions Related to Item Response Theory (IRT) Version 0.2.1; Partchev, Maris, Hattori 2017)
- “haven” (Import and Export 'SPSS', 'Stata' and 'SAS' Files Version 2.1.0; Wickham, Miller, RStudio 2019)
- “readr” (Read Rectangular Text Data Version 1.3.1; Wickham et al. 2018)

[99] Erweiterungen der Funktionen in Standard R, hierfür wird eine spezielle Packages-Bibliothek in R angelegt (vgl. Wollschläger 2017:19ff.; vgl. Hahnel 2018: 47)

- “lavaan” (Latent Variable Analysis Version 0.6-3; Rosseel et al. 2018)
- “semPlot” (Path Diagrams and Visual Analysis of Various SEM Packages Output Version 1.1.1; Epskamp et al. 2019)
- “eRm” (Extended Rasch Modeling Version 1.0-1; Mair et al. 2020)

Zum Package “lavaan” ist anzumerken, dass dies eine BETA-Version der kostenpflichtigen[100] Software „Mplus“[101] ist (vgl. Muthén & Muthén 1998-2012). Diese ist sehr teuer und wird zur Beurteilung von Strukturgleichungsmodellen und Modellvergleichen im Rahmen der Konfirmatorischen Faktorenanalyse eingesetzt (siehe Kapitel 16). In „TEMA“ wurde die Originalsoftware „Mplus“ Version 7 verwendet. Aus Kostengründen und letztendlich aber auch gleichen Berechnungsmöglichkeiten wird in der vorliegenden Arbeit die Software R unter Einbezug von „lavaan“ für die Berechnungen benutzt.

[100] wenn keine Einschränkung der Variablenanzahl vorhanden!

[101] Hinweis: Alternativ könnten zur Berechnung von Strukturgleichungsmodellen auch die Statistiksoftware LISREL, EQS, PLS-Graph oder Smart PLS angewendet werden (vgl. Döring, Bortz 2016b: 619).

14 Deskriptive statistische Datenauswertung

Im folgenden Kapitel werden die Datenanalysen, bezogen auf die TN, und die erhobenen Kompetenztestergebnisse vorgestellt. Zu berücksichtigen ist, dass keine Einzelauswertungen von TN bzw. einer Bildungseinrichtung aufgeführt werden. Es handelt sich deshalb generell um Gesamtergebnisse aller TN. Die Anonymität der einzelnen TN einschließlich der Bildungseinrichtungen bezüglich der Kompetenztestung wird somit geschützt. Der Einstieg erfolgt in Kapitel 14.1 mit einer Übersicht zu den individuellen TN-Daten.

14.1 Personenbezogene Datenanalyse

Insgesamt haben 408 TN zu Beginn der Kompetenztestung einen Bogen „Personenbezogene Daten zur Kompetenzmessung in der Gesundheits- und Krankenpflege“ erhalten (Anlage 3). Davon waren 106 TN aus Hochschulen und 302 TN aus beruflichen Bildungseinrichtungen. Aus der Gesundheits- und Kinderkrankenpflege kamen 17 TN. Der Rücklauf lag nach Ausfüllen des Bogens bei 405 TN-Bögen, dies entspricht einer Teilnahme von fast 100% (= 99,27%).

Erfragt wurden Informationen zu „Soziodemographischen Merkmalen“, „Ausbildungsbezogene Hintergrundmerkmale“ und ein „Besonderes Qualitätsmerkmal der beruflichen Gesundheits- und Krankenpflegeausbildung bzw. des Studiums in der Pflege“. Jeder TN wurde zunächst in einer „EXEL“-Tabelle erfasst, anschließend die Gesamthäufigkeiten ermittelt. Die folgende Tabelle (Tab. 3) stellt die Gesamtauswertung der TN dar.

Tab. 3: Personenbezogene Daten der TN

Merkmal		N	Gesamte TN
Soziodemographische Merkmale			
Geschlecht	männlich	76	402
	weiblich	326	
Geburtsland	Deutschland	353	398
	Europa	19	
	außerhalb der EU	26	
Lebensalter	unter 20	43	404
	20-30	325	
	30-40	27	
	über 40	9	
Höchster Schulabschluss	Hauptschule mit Zusatzabschluss	32	402
	Realschule	125	
	Gymnasium[102]	245	

[102] beinhaltet Fachhochschulreife, allgemeine Hochschulreife, Fachoberschule, Berufskolleg

Familienstand	ledig[103] verheiratet verwitwet alleinerziehend keine Angabe	366 29 0 3 4	402
Familiäre Pflege (Versorgung eines Angehörigen zu Hause)	findet statt fand statt trifft nicht zu	53 37 312	402
Lesen von Büchern und Fachzeitschriften	nie einmal pro Jahr einmal im Monat einmal in der Woche häufiger als einmal pro Woche	30 71 182 66 56	405
Computer- oder Internetnutzung	nie weniger als zwei Stunden pro Woche mehr als zwei Stunden pro Woche	1 30 370	401
Ausbildungsbezogene Hintergrundmerkmale			
Übergang von Schule zu Ausbildung bzw. Studium der Pflege	direkt erst später	177 217	394
Ausbildungszeit bzw. Studienzeit	verkürzt drei Jahre 8 Semester 9 Semester andere Studien- bzw. Ausbildungszeit[104]	4 295 6 92 37	397[105]
Primärer Grund für die Berufswahl	* weil der Beruf Spaß macht * weil der Beruf eine sichere Berufsperspektive versprach * weil ich mich für den Beruf für geeignet gehalten habe * weil mir meine Eltern, Verwandte, Freunde, PartnerIn dazu geraten haben bzw. hat * weil ich für diesen Beruf später respektiert werde * weil andere Menschen zu begleiten, einfach eine große Erfüllung bringt * anderer Grund[106]:	162 125 137 22 6 161 44	402[107]

[103] Ein TN hat bei „ledig“ den Zusatz „geschieden“ aufgeführt.
[104] Ein TN hat vermerkt: „keine Angabe“.
[105] Teilweise doppelte Benennungen, deshalb andere Gesamt-TN-Zahl in Summe bei Beantwortung.
[106] Am häufigsten wurde genannt: warten auf ein Medizinstudium; darüber hinaus wurde u.a. genannt: möchte Menschen helfen, spontan nach Pflege eines Angehörigen, weil ich Herausforderungen liebe, Interesse, Erfahrungen sammeln in der Pflege, Einsetzen für ein verbessertes Gesundheitswesen, Fortschritt in der Pflege erzielen, Interesse am Thema, keine andere Idee gehabt, Bedarf an Pflegefachkräften, großes Arbeitsfeld, Nähe zum Ausbildungsort, gute Karrieremöglichkeiten, finanzielle Absicherung, bereits die Eltern sind in der Pflege tätig, usw.

Besonderes Qualitätsmerkmal der beruflichen Gesundheits- und Krankenpflegeausbildung bzw. des Studiums in der Pflege			
1	Betriebliches Interaktionsklima	69	403[108]
2	Unterstützung durch Ausbilder	148	
3	Unterstützung durch Kollegen	148	
4	Betriebliche Gestaltungsqualität	27	
5	Betriebliche Instruktionsqualität	9	
6	Betriebliche Lernprozesskontrolle	63	
7	Strukturqualität betrieblicher Arbeitsaufgaben	28	
8	Aufgabenvielfalt, Fähigkeitsentsprechung	263	
9	Betriebliche Bedeutsamkeit	41	
10	Inhaltliche Autonomie der betrieblichen Aufgabenbearbeitung	51	
11	Zeitliche Dispositionschancen der betrieblichen Aufgabenbearbeitung	12	
12	Förderung des Lernklimas im familiären Umfeld	30	
13	Anderes Qualitätsmerkmal[109]	9	

Abschließende Bemerkung zu den Tabellenergebnissen: Die gesamten TN haben zu über 98% die einzelnen Fragestellungen beantwortet. Bereits im Vorfeld fiel die positive Einstellung, Offenheit und Bereitschaft zu einer Teilnahme am Kompetenztest auf. Wissenschaftlich interessant wäre die Einschätzung von höherer Schulbildung auf die Lösungswahrscheinlichkeit der Items im Kompetenztest noch genauer zu untersuchen (62,9% der TN haben eine gymnasiale Schuleinrichtung besucht, in „TEMA“ sind es nur 26,4% der TN!). Genauso bemerkenswert ist die Weiterverfol-

[107] Mehrfachbenennungen der einzelnen TN waren möglich, deshalb ergeben sich in N höhere Zahlenwerte! Die Gesamtanzahl der TN konnte trotzdem ermittelt werden!

[108] Mehrfachbenennungen der einzelnen TN waren möglich, deshalb ergeben sich in N höhere Zahlenwerte! Die Gesamtanzahl der TN konnte trotzdem ermittelt werden!

[109] Antworten: Selbständiges Lernen, familiär, große Auswahl an externen Einsätzen, Anerkennung durch Kollegen, Ausbildung mit Anspruch und Lernerfolgen, vergleichsweise gute Ausbildungsvergütung, Sozialkompetenz der Ausbilder, oder auch keine vertiefte Angabe in der Kategorie.

gung der beruflichen Lesekompetenz, wie sie sich aus den Ergebnissen aus Lesen (Häufigkeit bei 1x im Monat am stärksten ausgeprägt) und der Nutzung von Computer bzw. Internet (erhöhte Häufigkeit bei mehr als 2 Stunden am Tag) ergibt. Die Voraussetzungen für die Erfassung eines vorliegenden Textes sowie der grundlegende Umgang mit einem Notebook können angenommen werden. Auffällig zeigen sich außerdem als primäre Punkte der Berufswahl die Faktoren Spaß (162 Nennungen) und Erfüllung (161 Nennungen), wiederum als besonderes Qualitätsmerkmal der Punkt der „Aufgabenvielfalt, Fähigkeitsentwicklung“ (263 Nennungen). Dies weiterzuverfolgen würde aber den Rahmen des Forschungsauftrages bzw. der Forschungsfrage übersteigen, regt aber zu weiteren pflegewissenschaftlichen Untersuchungen an.

14.2 Kompetenzbezogene Datenanalyse aus „TEMA“ und „CoSMed“

Zunächst wurden die Häufigkeiten der 77 bzw. 76[110] „TEMA“-Items bezüglich Beantwortung/Nichtbeantwortung einschließlich der Scoringzugehörigkeit bestimmt, danach die von den „CoSMed“-Items. Bei den 389 TN-Datensätzen haben 194 TN Testheft 2 und 195 TN Testheft 3 bearbeitet. In den folgenden Darstellungen und Tabellen sind beispielhaft das dichotome Item A31 mit leichtem, einfachem Beantwortungsmuster (66,3%) und das polytome Item A75 mit schwerem Beantwortungsmuster (8,7% haben es vollständig gelöst, 86,1% haben es nicht gelöst) aus dem ausgewerteten Datensatz aufgeführt.

[110] Das Item A48 wurde für die weiteren Auswertungen entfernt, da von keinem TN beantwortet. Es handelte sich um eine „Image-Map“-Fragestellung mit einer Fallbeschreibung zur Bestimmung der korrekten Injektionsorte (s.c. = subcutan) von Insulin. Hier bestätigt sich das gleiche Item wie bei „TEMA“ als „Ausreißer“.

Tab. 4: Zwei Itemauswertungsbeispiele zur Häufigkeitsverteilung in SPSS

A31

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	124	31,9	32,5	32,5
	1	258	66,3	67,5	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A75_polytom

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	335	86,1	87,7	87,7
	1	7	1,8	1,8	89,5
	2	6	1,5	1,6	91,1
	3	34	8,7	8,9	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

Alle weiteren Häufigkeitsverteilungen zu den nicht aufgeführten Items befinden sich im Anhang 15 und können dort eingesehen werden.

Innerhalb der weiteren polytomen und bei einer Vielzahl von dichotomen Items zeigen sich eine hohe Anzahl an TN bei den nicht korrekt beantworteten Aufgabenstellungen (= 0), wie z.B. A42, A73, A74 usw. Dies lässt auf eine hohe Schwierigkeit der jeweiligen Aufgaben schließen. Der ermittelte Schwierigkeitsgrad aus „TEMA“ bestätigt sich somit in den Häufigkeitsergebnissen aus der Gesundheits- und Krankenpflege (vgl. Döring et al. 2014: 456ff.).

Anschließend wurden zur besseren Veranschaulichung mehrere Gesamtübersichten über die Beantwortung der Items, der Berücksichtigung des Scorings in den Ergebnissen einschließlich der Häufigkeitsverteilungen bezogen auf die TN und den korrekt beantworteten Items erzeugt. Die folgenden zwei Balkendiagramme zeigen die Ergebnisse auf (Abb. 6 u. 7).

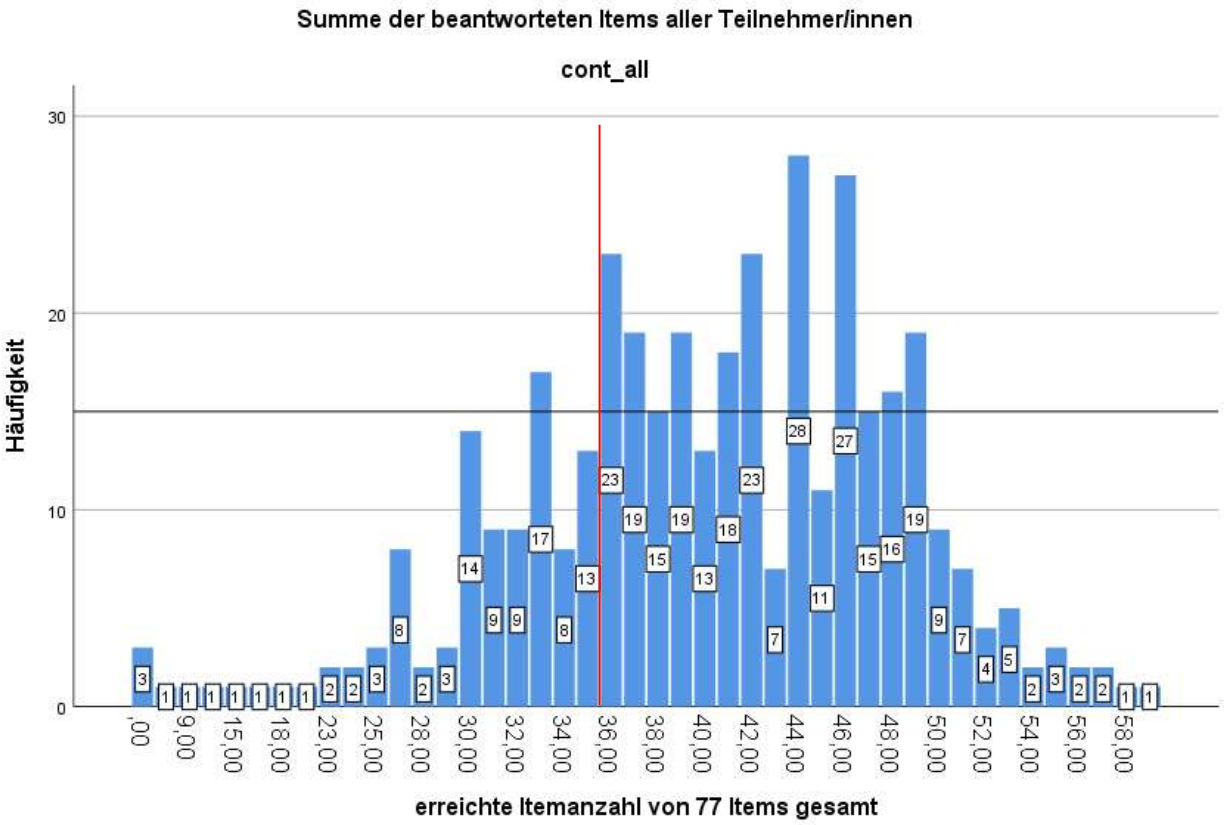


Abb. 6: Übersicht zur Verteilung der beantworteten Items aller TN aus SPSS[111]

Von den 77 Items mussten mindestens 36 Items (Abb. 6 roter Strich/Cut-off-Wert[112]) beantwortet werden, damit der Kompetenztest als bestanden gilt. Unter der Bedingung cont_all haben 289 TN bestanden und 100 TN nicht bestanden. Die Häufungen lagen bei 44 Items = 28 TN, 46 Items = 27 TN, gefolgt von 42 Items = 23 TN und 36 Items = 23 TN ohne eine vorgenommene Normierung. Die Differenz liegt zwischen 0 und 59 Items. In „TEMA“ lagen die Ergebnisse zwischen 10 und 59 Items. Es zeigen sich fast vergleichbare Werte, außer den Ausreißern von unter 10 Items. Die Gründe dafür könnten in mangelnder Motivation, reduziertem Textverständnis und damit eingeschränkter pflegerischer Fallerfassung sowie Schwierigkeiten mit der Konzentration sein. Warum nicht mehr als 59 Items korrekt beantwortet wurden, lässt auch die Frage zu, ob die aufwendig gescorten Items zur Ergebniserzielung überdacht werden sollten. Dies lässt sich abschließend aber nicht vollständig klären. Weshalb drei TN, ohne Angabe von Gründen, kein Item richtig beantwortet haben, trotz korrektem Abschluss des Kompetenztests, lässt sich nicht ermitteln. Die nächste Tabelle zeigt eine Gesamtverteilungsübersicht unter Einbezug des vorgenommenen Scorings.

[111] Hinweis: In der aufgeführten Abbildung sind auf dem Zahlenstrahl der x-Achse „erreichte Itemanzahl von 77 Items gesamt“ zwischen 28 und 59 immer zweierschritte angegeben. Unterhalb dieser Angabe sind nur noch die erreichten Punktwerte aufgeführt, d.h. 0, 1, 9, 13, 15, 17, 18, 19, 23, 24, 25, 27.

[112] Mindestanforderung des TN ist erfüllt, d.h. der TN hat bestanden!

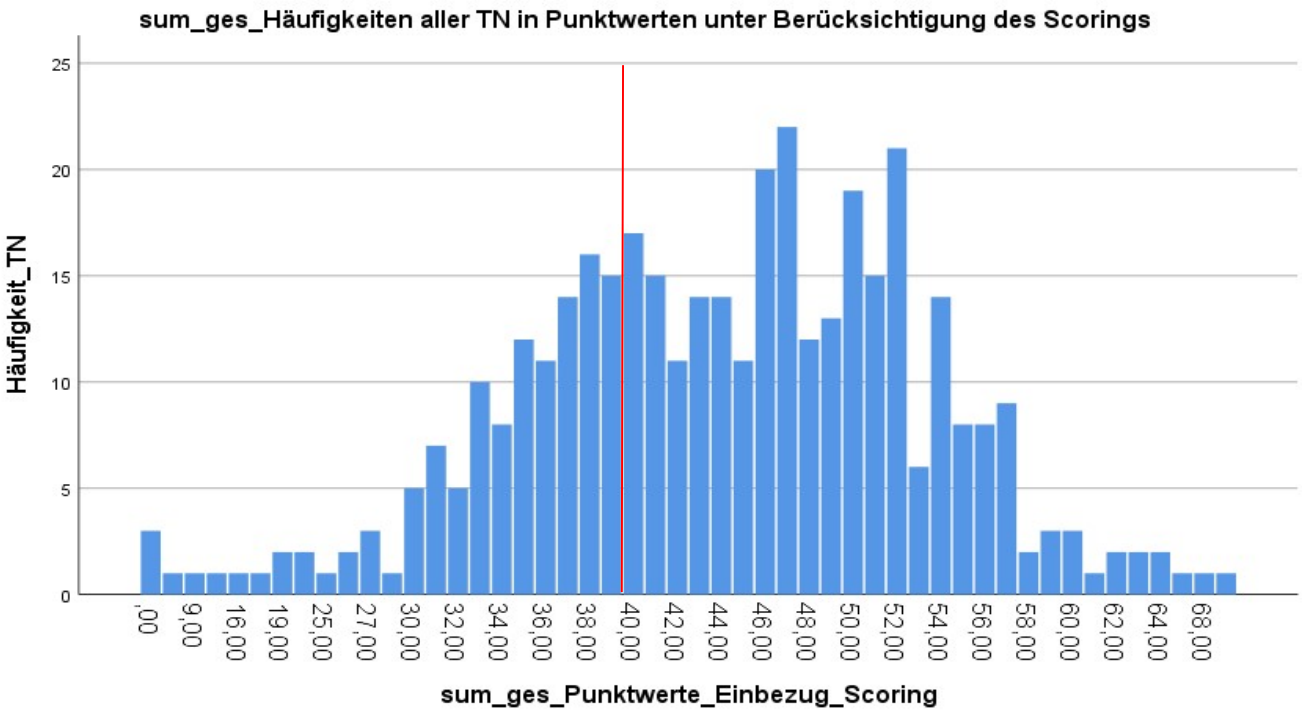


Abb. 7: Gesamtpunktzahlen der TN unter Scoringbedingungen in SPSS[113]

Bei einer Annahme von sum_ges von 95 gescorten Punkten und einem Bestehen ab 40 gescorten Punkten (Abb. 7 roter Strich/Cut-off-Wert) hätten 267 TN den Test bestanden und 122 TN nicht bestanden[114]. Die Häufigkeiten der Gesamtpunktwerte der TN unter Berücksichtigung des Scorings jedes Items sind in der Spanne von 0 bis 69 Punkten zu finden, bei einem MW[115] von 43,75 Punkten und einem Median von 45 Punkten. Punkthäufungen bezüglich der TN zeigen sich bei 47 Punkten mit 22 TN, 52 Punkten mit 21 TN, 46 Punkten mit 20 TN und 50 Punkten mit 19 TN[116].

Bezogen auf die Forschungsfrage zeigt sich in der Gesamtbetrachtung der „TEMA"-Kompetenztest auch in der Gesundheits- und Krankenpflege als erfolgreich einsetzbar, aber durchaus herausfordernd in der Beantwortung der einzelnen Items für die Auszubildenden bzw. Studierenden. Dies stützt die vorhandenen Ergebnisse aus der Altenpflege (vgl. Kaspar et al. 2016: 1 u. 24ff.).

Weiterführend wurden die Auswertungen zum erhobenen „CoSMed"-Datensatz vorgenommen. Diese waren, wie in Kapitel 4.3 erwähnt, mit in den technologiebasierten, computergestützten Kompetenztest integriert. Hier zeigten sich gute Gesamtergebnisse der TN, d.h., ca. 2/3 (= 239 TN) haben die einzelnen Aufgabenstellungen

[113] Hinweis: In der Abbildung sind im Zahlenstrahl auf der x-Achse „sum_ges_Punktwerte_Einbezug_Scoring" zwischen 30 und 64 zweierschritte angegeben, jedoch unterhalb und überhalb dieser Werte werden nur noch die erreichten Gesamtpunktzahlen aufgeführt, d.h. 0, 2, 9, 14, 16, 18, 19, 24, 25, 26, 27, 28 und 64, 66, 68, 69.
[114] ohne den „CoSMed" Ergebnissen
[115] Mittelwert
[116] in absteigender Reihenfolge aufgeführte TN-Zahl!

innerhalb der Testitems mit 9 Punkten und mehr beantworten können. Es gab keine fehlenden Werte von den gesamten 389 TN.

Die folgende Tabelle (Tab. 5) zeigt die Häufigkeiten der TN im Beantwortungsmuster der einzelnen Items, darauf folgt ein Balkendiagramm (Abb. 8) mit der Verteilung der korrekt gelösten Gesamtitemanzahl der TN im Überblick.

Tab. 5: Beantwortungsmuster „CoSMed“ in SPSS

		Anzahl
A83neu Items CosMed	0	65
	1	324
A84 Items CosMed	0	171
	1	218
A85 Items CosMed	0	92
	1	297
A86 Items CosMed	0	214
	1	175
A87_polytom Items CosMed	0	17
	1	149
	2	223
A88 Items CosMed	0	170
	1	219
A89 Items CosMed	0	165
	1	224
A90 Items CosMed	0	138
	1	251
A91 Items CosMed	0	78
	1	311
A92 Items CosMed	0	55
	1	334
A93 Items CosMed	0	73
	1	316
A94 Items CosMed	0	218
	1	171

Die Items A86 (= 44,99% gelöst) und A94 (= 43,96% gelöst) stellen sich als besonders schwierig zu lösende Items für die TN heraus. Sie wurden von jeweils über 200 TN nicht korrekt beantwortet. Im Vergleich zeigen sich aber im Beantwortungsmuster der Ergebnisse der Altenpflege noch niedrigere Werte von Item A86 (= 24,42% gelöst) und Item A94 (= 40,21% gelöst).

Inhaltlich geht es um die Themen der urologischen Krankheitsbildeinschätzung anhand einer fiktiven Falldarstellung und medizinische Mathematik im Bereich Tropfenverabreichung. Dieses Spektrum wird in der theoretischen sowie praktischen Gesundheits- und Krankenpflegeausbildung umfassend vermittelt. Die Grundlagen der Mathematik werden aber in der Regel als gegeben vorausgesetzt und nicht nochmals mit den Auszubildenden vertieft oder wiederholt. Die Ermittlung der Tropfgeschwindigkeit findet wiederum regulär im ersten Ausbildungsjahr statt. Die urologischen Krankheitsbilder werden je nach Lehrplan der Bildungseinrichtung im zweiten und/oder dritten Ausbildungsjahr besprochen. Darüber hinaus findet durch praktische Schülereinsätze auf vermehrt interdisziplinären Fachabteilungen und Multimorbidität der einzelnen Patienten von Beginn der Ausbildung bzw. des Studiums die Auseinandersetzung mit urologischen Krankheitsbildern statt.

Als extrem niedrig im Beantwortungsmuster zeigte sich darüber hinaus im Ursprungsprojekt Altenpflege das Item A88 (= 4,70%) mit einem jetzigen Lösungswert in der Gesundheits- und Krankenpflege von 56,3%. Dies beruht sicher auf der Tatsache, dass der Schwerpunkt Krankheitsbilder (Diagnostik und Therapie) einschließlich Ermittlung von Laborparametern mit in den Lehrplänen und Curricula verankert ist und bei praktischen Begleitungen und Pflegeskills eine zentrale Rolle in der Akutversorgung spielt (vgl. Döring et al. 2014: 244ff.).

Gleichzeitig ist darauf hinzuweisen, dass die guten Ergebnisse auf die Komplexität der Fragestellungen zu den möglich Antwortalternativen, überwiegend in kurzen Sätzen oder nur einzelnen Begrifflichkeiten, zurückgeführt werden können. Dieser Frage-Antwort-Stil erleichtert dem TN die kognitive Erfassung der abzufragenden einzelnen Items und sein darauf reaktives Verhalten im Kompetenztest (vgl. Porst 2014: 17).

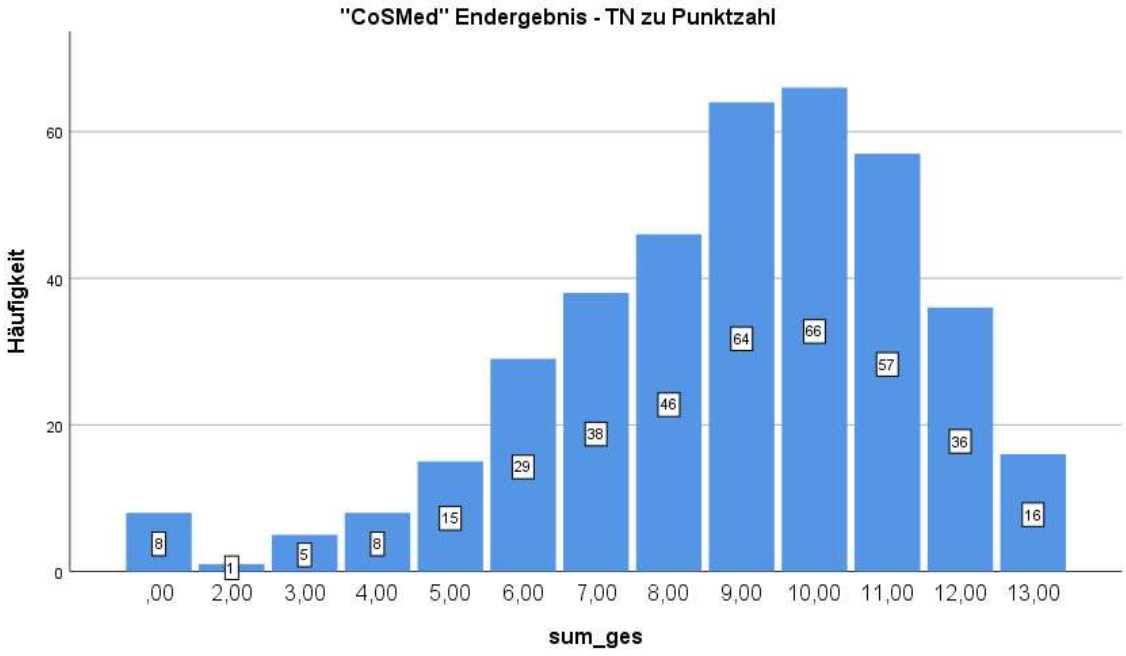


Abb. 8: Verteilungsdiagramm der korrekten Antworten aus „CoSMed“

Wie bereits erwähnt, zeigt das oben aufgeführte Diagramm die TN-Verteilung mit jeweils erreichter Gesamtpunktzahl in „CoSMed“. 8 TN haben kein Item korrekt beantwortet. Kein TN hat nur eine richtige Antwort erreicht, deshalb ist der Wert hier nicht aufgeführt. Im Endergebnis beginnt die positive Lösungsspanne bei 2 korrekten Antworten und endet bei der maximal zu erreichenden Punktzahl von 13 möglichen Antwortpunkten.

Daraus folgt, dass die aktuellen „CoSMed“- Endergebnisse dazu aufrufen, dass bei einer weiteren Überarbeitung des Messinstrumentes aus „TEMA“ eine durchaus essenzielle Ergänzung im akuten stationären klinischen Bereich für die zukünftige Ausrichtung in der generalistischen Ausbildung ab 2020 zu sehen ist.

14.3 Weitere statistische Teilergebnisse in SPSS

Als erste Ergebnisse bezüglich der Messgenauigkeit des Kompetenztests wird im Bereich der Reliabilitätsstatistik[117] geschätzt, wie hoch die „interne Konsistenz“ ist, d.h., wie stark stehen die einzelnen Items miteinander in Verbindung. Dies wird unter der genutzten Anwendung auch bei dichotomen Items mit der gängigen Methode der Bestimmung des „Cronbach Alpha“-Koeffizienten angegeben. Der Wert ist abhängig von den mittleren Werten der Inter-Item-Korrelationen und der vorhandenen Gesamtanzahl an Testitems (vgl. Bühner 2011: 166ff.; vgl. Kunina-Habenicht 2018: 44ff.).

Die dafür einzusetzende Formel lautet (vgl. Döring, Bortz 2016: 468) (Gl.2):

$$\text{Rel}_{\text{interne Konsistenz}} = \alpha = \frac{p}{p-1} \cdot \left(1 - \frac{\sum_{i=1}^{p} Var_{xi}}{Var_x}\right)$$

p = Anzahl der Items
xi = Itemwerte von 1 bis p
x = Testwert als Summe aller Itemwerte

Der ermittelte Cronbach Alpha liegt bei einer berücksichtigten Anzahl von 76 Items bei einem Wert von 0,786[118], d.h., der Test zeigt eine tendenziell gute Reliabilität bzw. interne Konsistenz an. Es ist aber kein valider Beweis für eine „Eindimensionalität“ im vorliegenden Testmodell (vgl. Schermelleh-Engel & Werner 2012: 133f.). Eine Verbesserung des Cronbach Alpha wird ausschließlich signifikant bei Weglassen des Items IK33A63 mit einem Folgewert von 0,800 (siehe Anhang 16). Zu berücksichtigen ist, dass der ermittelte Wert des Cronbach Alpha als nicht bindend gesehen werden muss, da die vorhandenen Daten nicht stabil sind. Somit gilt

[117] „… der Anteil der Varianz der wahren Werte an der Varianz der beobachteten Werte …“ (Bühner 2011: 144). Die Festlegung dieses Differenzwertes erfolgt über eine Korrelation. Weitere Hintergründe zu den Grundlagen zur Reliabilität wurden bereits in Kapitel 7.1.2 „Gütekriterien der quantitativen Forschung“ erläutert.

[118] zufriedenstellende Reliabilitätswerte liegen bei einem Mindestwert von 0,7, bei Werten von 0,8 gilt der Test als gut.

der ermittelte Wert als Ersteinschätzung für die Messgenauigkeit des Kompetenztests, hat aber nicht genügend Aussagekraft (vgl. Döring, Bortz 2016: 468f.; vgl. Schermelleh-Engel & Werner 2012: 130ff.). Weitere Berechnungen und Einschätzungen zum Strukturmodell werden deshalb diesbezüglich noch in Kapitel 16 weiter vertieft werden.

Darüber hinaus wurden zur Itemanalyse zwei Berechnungen vorgenommen. Zunächst wurden die berechneten Item-Mittelwerte[119] betrachtet, um die Itemschwierigkeiten der einzelnen Items einschätzen zu können. Je niedriger der MW eines Items, umso schwieriger ist es. Die Angabe erfolgt in % der gelösten Items[120] der TN in SPSS. Die MW lagen bei hohem Schwierigkeitsgrad u.a. bei Item DR23A42 bei 0,06 und bei Item IK14A20 bei 0,10. Bei niedrigem Schwierigkeitsgrad lagen die MW u.a. bei Item IK22A27 bei 0,96 und bei Item IK23A39 bei 0,90. Im mittleren Schwierigkeitsbereich finden sich u.a. das Item IK21A32 bei 0,54 und das Item IK22A30 bei 0,48. Im Gesamtergebnis zeigen sich die Schwierigkeiten der einzelnen Items als durchaus unterschiedlich, d.h., das komplette Bewertungsspektrum ist vertreten (siehe Anhang 17).

Abschließend wurden die Item-Trennschärfen berechnet und bewertet. Diese sind in SPSS als korrigierte Item-Skala-Korrelation angegeben (siehe Anhang 16) und werden durch die Itemschwierigkeiten mit beeinflusst. Die Lösung eines Items steht in Verbindung mit der Zuverlässigkeit der Leistung im Gesamttest. Es wird auch von „Item-Test-Korrelation“ gesprochen (vgl. Pospeschill 2010: 80; vgl. Döring, Bortz 2016b: 477f.). In den ermittelten Diskriminationen[121] in der Gesundheits- und Krankenpflege stellt sich im Range-Index das Spektrum von 0,09 bis an die 0,383 dar. Folglich sollten schlechte Diskriminationen (unter 0,19) ein Eliminieren von niedrig bewerteten einzelnen Items oder eine nochmalige Überarbeitung bzw. Anpassung (ausreichende Diskrimination zwischen 0,20 und 0,29) der Items zur Folge haben, um die Diskriminationswerte zu verbessern und bei weiteren Projekten erneut zu überprüfen (vgl. Pospeschill 2010: 82f.; vgl. Kunina-Habenicht 2018b: 48).

Hohe Diskriminationen der entwickelten Items sind aus diesem Grund erwünscht. Oder nach Eid u. Schmidt:

> „Der Trennschärfenkoeffizient eines Items entspricht der Korrelation des Items mit der Testwertvariablen. Je größer die Korrelation ist, umso stärker ist der Zusammenhang eines Items mit dem Gesamttestwert und umso stärker wirken sich auch Unterschiede in den Testwerten auf Unterschiede in den Itemantworten aus. Anzustreben sind daher Items mit hohen Trennschärfekoeffizienten.“ (Eid u. Schmidt 2014: 216).

[119] abgekürzt „Item-MW“

[120] zwischen 0 und 0,20 wird als sehr schweres Item, zwischen 0,40 und 0,60 als mittelschweres Item und zwischen 0,81 und 1 wird als sehr leichtes Item bezeichnet (vgl. Kunina-Habenicht 2018b: 33).

[121] „Unterscheidung“ oder „Differenzierung“ eines gewählten einzelnen Items zum Gesamttest; eine sehr gute Diskrimination liegt bei > 0,40 (vgl. Kunina-Habenicht 2018b: 48; vgl. Pospeschill 2010: 80ff.).

Zur weiteren Datenanalyse mit Beurteilung des Messinstrumentes werden die statistischen Verfahren aus „TEMA“ aufgegriffen und damit eine vollständige Übertragung in die Gesundheits- und Krankenpflege realisiert. Im folgenden Kapitel 15 wird zunächst die Itemkonstruktion innerhalb der entworfenen Testmodelle untersucht. Es wird dafür eine Fortentwicklung der Klassischen Testtheorie[122], die sogenannte Item-Response-Theorie, zur Beurteilung der Messung eines latenten Konstrukts, eingesetzt. Im Anschluss wird in Kapitel 16 mit einer Dimensionalitätsprüfung das passende Messmodell (ein- oder mehrdimensional?) ermittelt.

[122] Abkürzung: KTT oder auch genannt „true score theory“; Hier wird die „wahre“ Ausprägung des Merkmals unter Einbezug von Messfehlern bestimmt. Der Wert des Testfehlers muss aber vorliegen. Im Mittelpunkt steht die Reliabilitätsbestimmung. (vgl. Döring, Bortz 2016: 461).

15 Die Item-Response-Theorie

Die Item-Response-Theorie[123] ist eine mögliche Testtheorie, mit der von manifesten Variablen[124] auf eine latente Variable[125] oder ein Konstrukt geschlossen werden kann. Dabei wird versucht, Dinge zu messen, die man normalerweise nicht messen kann, wie z.B. in der hier vorliegenden Arbeit praktische Pflegekompetenz[126]. Das pflegerische Verhalten soll unter Zuhilfenahme dieser Theorie transparent gemacht werden. Exaktes Messen ist aber nicht möglich. Es wird von Wahrscheinlichkeiten ausgegangen. Die IRT ist der Überbegriff für eine ganze Reihe von verschiedenen entwickelten Messmodellen, je nach Modell unter möglicher Berücksichtigung von u.a. Itemschwierigkeit, Itemtrennschärfe[127] usw. Es kommt auf die durchzuführende Testung an, welches Messmodell seine Anwendung findet. Das Item und das darauffolgende Antwortverhalten bilden ein sich ergebendes Konstrukt. Dieses kann sehr unterschiedlich gestaltet sein (vgl. Pospeschill 2010: 114; vgl. Moosbrugger, Kelava 2012: 228f.; vgl. Döring, Bortz 2016b: 482ff.; vgl. Brühl 2016: 3ff.).

Die folgende Abbildung (Abb.9) soll diesen Zusammenhang nochmals darstellen.

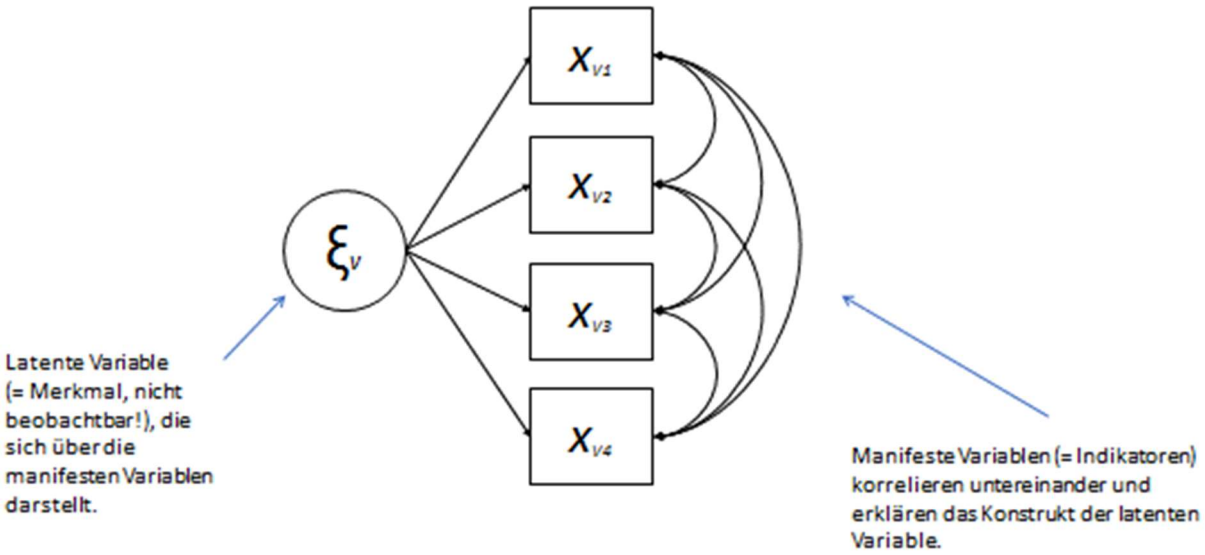


Abb. 9: Korrelationen zwischen latenter Variable und ihren manifesten Variablen (entnommen aus: Pospeschill 2010: 115 leicht modifiziert von Ries 2019)

[123] abgekürzt „IRT“ oder auch probabilistische Testtheorie (= PTT); Komplettierung zur Klassischen Testtheorie (= KTT)

[124] beobachtetes Antwortverhalten auf unterschiedliche Items in der jeweiligen Testung; Cave: dichotomes oder polytomes Testitem!

[125] nicht exakt beobachtbar

[126] diese wird dann als „kardinalskaliertes latentes Merkmal“ und in Form eines vorliegenden Konstrukts „Latent-Trait-Modell“ genannt, da es die zu erkennende praktische Pflegekompetenz messen soll (vgl. Döring, Bortz 2016: 483). Die Betrachtung bestimmter Merkmale gibt dabei Hinweise, obwohl zur Analyse ausschließlich beantwortete Items verwendet werden.

[127] „Item Discriminability“; Lösung eines bestimmten Items zeigt eine mögliche Vorhersage zum gesamten Testergebnis an. Die Angabe erfolgt im sogenannten Trennschärfe-Index (vgl. Döring, Bortz 2016: 478).

Die IRT kann für die Entwicklung und Auswertung psychometrischer Tests in Form eines „Large-Scale-Assessment“[128] u.a. als computergestützter Test oder per „Paper & Pencil“-Fragebogen eingesetzt werden (vgl. Hahnel 2018: 35). Dies ist der Grund, weshalb die Methode in vorliegendem Projekt zum Einsatz kommt.

Mit der IRT können Gütekriterien wissenschaftlich aufbereitet werden, wie z.B. Skalierungen, Bestimmung von Itemparametern ohne die Berücksichtigung der Stichprobe. Der Entwurf eines IRT-Testmodells verlangt nach einem Itemschwierigkeitsparameter[129], Personenfähigkeitsparameter[130], einer „Joint Scale“[131], spezifischer Objektivität[132] und einer statistischen Testbarkeit, bezogen auf die erhobene Stichprobe in Relation zur Grundgesamtheit. Darüber hinaus sind die Annahmen der Itemhomogenität[133], die lokale stochastische Unabhängigkeit[134] und die itemcharakteristische Funktion[135] bzw. die monotone itemcharakteristische Kurve[136] (Abb. 10) zu berücksichtigen und anzugeben (siehe Kapitel 4.2).

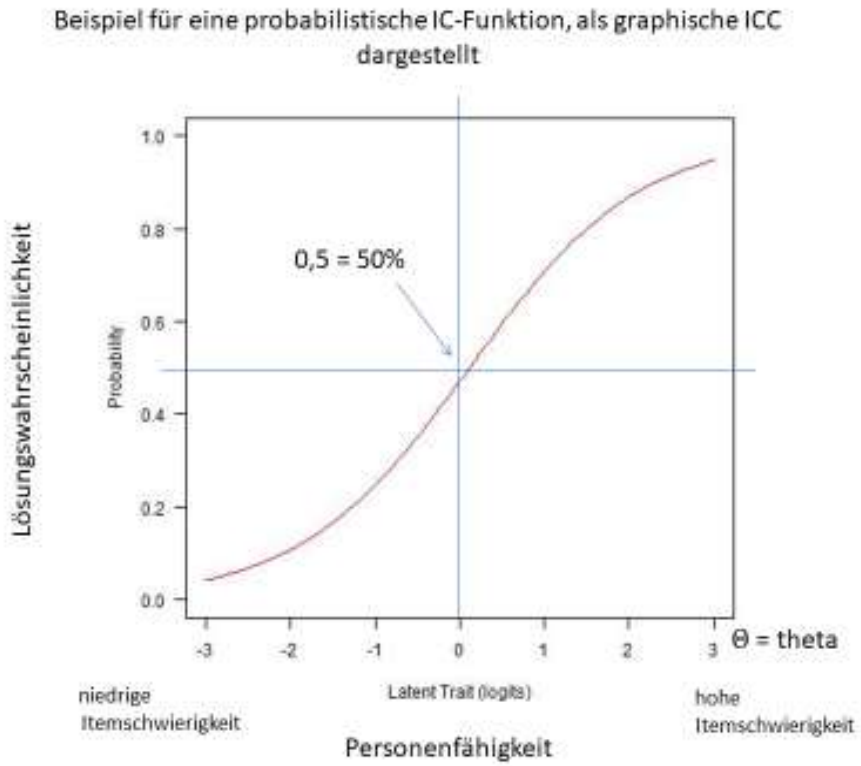


Abb. 10: Erklärungen zum Aufbau einer ICC in der PTT aus den verwendeten Testitems in der Gesundheits- und Krankenpflege (vereinfachte Darstellung in Anlehnung an Mai 2010: 37, erstellt u. modifiziert von Ries 2019)

[128] wie beispielsweise in der „PISA“-Studie oder in der „TEMA“-Studie
[129] Angabe als sigma bzw. σ
[130] Angabe als theta bzw. θ oder xi bzw. ξ
[131] eindimensionale Skala für Itemschwierigkeit und Personenfähigkeit
[132] Itemschwierigkeit und Personenfähigkeit werden getrennt voneinander bestimmt bzw. „unabhängig“ ermittelt (vgl. Hahnel 2018: 15; vgl. Bensch 2013: 103).
[133] Nur eine einzige latente Variable wird gemessen (z.B. nur praktische Pflegekompetenz).
[134] Es ist kein Zusammenhang der manifesten Items untereinander zu erkennen.
[135] oder „IC”-Funktion bei dichotomen Items genannt
[136] „Item Characteristic Curve” oder Kürzel „ICC“

Die monotone ICC wird in Anlehnung an eine logistische Funktion[137] angegeben. Auf der Abszisse werden in der oben angeführten Abbildung die Personenfähigkeit und die Itemschwierigkeit abgetragen, während auf der Ordinate die Lösungswahrscheinlichkeit[138] des vorliegenden Items zu erkennen ist. Beispielsweise lässt es die Schlussfolgerung zu, dass ein TN mit sehr hohem Fähigkeitsmuster auch eine sehr hohe Lösungswahrscheinlichkeit hat. Bei einer vorliegenden Personenfähigkeit von 0[139] liegt eine Lösungswahrscheinlichkeit von fast 50% vor. Es ist im angeführten Beispiel eine minimal erhöhte Itemschwierigkeit zu erkennen. Je höher die Personenfähigkeit θ, desto höher die Ausprägung des latenten Merkmals, d.h. bei Bestätigung eines eindimensionalen Strukturmodells die praktische Pflegekompetenz. Darüber hinaus zeigt das vorliegende Itembeispiel eine geringere Trennschärfe, da die Steigung im Nullpunkt flacher verläuft zur Ordinate. Ein Vergleich verschiedener ICC kann durch vorliegendes Beispiel aber noch nicht erfolgen, da in der Abbildung nur eine Kurve angegeben wurde und diese zum Verständnis des Aufbaus einer ICC dient. Zur endgültigen Einschätzung müsste das Itemergebnis mit dem gesamten Ergebnis aller verwendeten Testitems verglichen werden. Eine Konkretisierung der einzelnen Itemergebnisse wird dadurch möglich. Zudem besteht die Möglichkeit einer adaptiven Anpassung[140] an die einzelnen TN, d.h., diese erhalten die Items mit geeignetem Schwierigkeitsparameter (vgl. Döring, Bortz 2016b: 485; vgl. Kaspar et al. 2016: 11).

Es gibt diverse probabilistische Testmodelle, die unterschiedliche Einflussgrößen[141] mit einbeziehen können, je nach vorhandenem Itemformat[142] und unter Berücksichtigung des notwendigen zu messenden latenten Merkmals (vgl. Döring, Bortz 2016b: 486f.). In der hier vorliegenden Arbeit liegt der Schwerpunkt beim 1 PL-Modell[143] bzw. in Form des „Partial-Credit-Modells“ und 2 PL-Modell[144] bzw. in einer erweiterten Form als „Generalisiertes Partial-Credit-Modell”. Diese sollen zum Verständnis und Hintergrund der Datenanalyse kurz umrissen werden. Anschließend werden die Ergebnisse der Datenauswertungen in R vorgestellt und für die Überführung zu einem Gesamtergebnis einer vorliegenden Raschhomogenität[145] interpretiert.

[137] S-geformte Verlaufskurve
[138] wird angegeben von 0 bis 1 oder auch 0% bis 100% Lösungswahrscheinlichkeit (vgl. Döring, Bortz 2016: 484).
[139] Merkmal ist im mittleren Bereich der Abszisse der oben angeführten Abbildung zu finden bzw. vorliegend.
[140] auch „adaptives Testen“ genannt
[141] Personenparameter, Itemschwierigkeit, Trennschärfe, usw.
[142] dichotom oder polytom
[143] „Einparameter-Logistisches-Modell“, auch „Rasch-Modell“ genannt
[144] „Zweiparameter-Logistisches-Modell“, auch „Birnbaum-Modell“ genannt
[145] unterschiedliche Schwierigkeit der Items liegt vor (vgl. Pospeschill 2010: 123).

15.1 Rasch-Modell

Als Grundlage für eine ganze Reihe weiterer IRT-Modelle wurde das Rasch-Modell entwickelt, das auch als dichotomes Modell bezeichnet wird. Im Zentrum steht der Itemschwierigkeitsparameter. Das Antwortformat stellt sich in Form von zwei Antwortoptionen dar, beispielsweise „ja/nein“ oder „richtig/falsch“. Damit es angewendet werden kann, ist es notwendig, dass zwei Bedingungen erfüllt sind: Itemhomogenität und stochastische Unabhängigkeit (vgl. Moosbrugger 2012: 229; vgl. Rost 1999: 140ff.).

Alle weiteren Bedingungen entsprechen den in Kapitel 15 erläuterten Grundlagen der IRT.

Die Anzahl der richtigen Antworten zeigt die Fähigkeit einer Person bzw. des jeweiligen TN an[146]. Jedes Item besitzt im Gesamtmodell die identische IC-Funktion. Die anzuwendende Formel für das Rasch-Modell lautet (vgl. Pospeschill 2010: 122; vgl. Hahnel 2018: 26) (Gl.3):

$$P\ (X_{vi} = 1) = \frac{e^{\theta_v - \sigma_i}}{1 + e^{\theta_v - \sigma_i}}$$

X_{vi} = 1 -> Item wird korrekt beantwortet.
e = Eulersche Zahl 2,718 oder auch exp[147]
θ oder auch ξ = Personenfähigkeit
σ oder auch β = Itemschwierigkeit

Ebenso wie die Lösungswahrscheinlichkeit (X_{vi} = 1) ermittelt werden kann, ist auch die Gegenwahrscheinlichkeit (X_{vi} = 0) zu berücksichtigen. Bei Betrachtung einer IC-Funktion sind unterschiedliche Konstellationen im Verlauf möglich. Diese wurden ebenfalls in Kapitel 15 bei den Grundlagen des Aufbaus einer IC-Kurve erläutert.

Weiterhin ist die sogenannte „Rasch-Homogenität“ zu prüfen. Sie ist dann erfüllt, wenn sich die vorhandenen Items ausschließlich in ihrem Schwierigkeitsparameter unterscheiden. Anhand von abgebildeten IC-Funktionen kann abgelesen werden, je weiter nach rechts die IC-Funktionen der einzelnen Items auf einer Skala abgetragen werden, desto anspruchsvoller sind sie (vgl. Pospeschill 2010: 123).

Um ideale Parameterwerte zu gewinnen, werden die nicht bekannten Aufgabenparameter und Personenparameter geschätzt und nach dem Maximum-Likelihood-Prinzip ermittelt. Ein Gesamtbild aller beobachteten Werte wird sichtbar, und Schätzfehler werden folglich minimiert. Die Likelihood-Funktion unter Bezugnahme eines bestimmten Datenmusters X[148] wird mit folgender Formel ermittelt (vgl. Pospeschill 2010: 124; vgl. Hahnel 2018: 60; vgl. Kreienbrock, Pigeot, Ahrens 2012: 313: ff.)(Gl.4):

[146] erschöpfende Statistik (vgl. Döring, Bortz 2016: 489)
[147] Expotentialfunktion
[148] θ und σ

$$L = P\ (X|\ \theta, \sigma) = \prod_{v=1}^{n} \prod_{i=1}^{m} P\ (x_{vi}) = \prod_{v=1}^{n} \prod_{i=1}^{m} \left(\frac{exp\ \left(x_{vi}(\theta_v - \sigma_i)\right)}{1 + exp(\theta_v - \sigma_i)} \right)$$

$\prod$ = Multiplikationssymbol
n = Anzahl der Personen bzw. Stichprobengröße
m = Anzahl der Items

Alle anderen Symbolerläuterungen können aus der zuvor aufgeführten Formel entnommen werden. Für die Stichprobengröße sowie die Items wird die Wahrscheinlichkeit $P\ (x_{vi})$ ermittelt. Im abschließenden Schritt werden die beiden Ergebnisse mit allen weiteren Wahrscheinlichkeiten zu einem Gesamtwert multipliziert. Der ermittelte Wert liegt zwischen 0 und 1. Bei geeigneten Werten, Werten gegen 1, ist die Wahrscheinlichkeit erhöht, bei ungeeigneten Werten, Werten gegen 0, erniedrigt. Das Maximum wird erreicht, wenn der Wert der Probanden und der Wert der Items sich nach den zuvor erläuterten Kriterien, als „ideal“ darstellen, mit der Bezeichnung als „erschöpfende Statistik“ (vgl. Pospeschill 2010: 126).

Je komplexer das Datenmuster X, desto kleiner ist der ermittelte Wert der Likelihoodfunktion. Deshalb kann stattdessen zum Schätzen von noch genaueren und geeigneteren Parameterwerten die „logarithmierte Likelihood“[149] bestimmt werden. Das Maximum der Funktion wird durch die erste Ableitung[150] und darauf folgend die zweite Ableitung[151] erreicht. Die dafür anzuwendende Formel lautet (vgl. Hahnel 2018: 61f.) (Gl.5):

$$log\ (P\ (X|\theta, \sigma)) = \sum_{v=1}^{n} \sum_{i=1}^{m} log\ \left(P_{vi}^{X_{vi}}\ Q_{vi}^{1\ -\ X_{vi}}\right)$$

X = gesamte Datenmatrix bzw. Datenmuster
$\sum$ = Summenzeichen
P_{vi} = Wahrscheinlichkeitswert von TN v, Item i richtig zu beantworten
Q_{vi} = Wahrscheinlichkeitswert von TN v, Item i nicht richtig zu beantworten

Die iterative Suche nach geeigneten Personen- und Itemparametern kann über verschiedene Schätzverfahren erfolgen, bei den Itemparametern beispielsweise über die „Conditional Maximum Likelihood“[152] unter Berücksichtigung der Summenscores oder den Personenparametern wie etwa über die „Weighted Likelihood Estimates“[153], mit der auch Extremwerte geschätzt werden können. Alternativ gilt auch der „Expected A Posteriori Estimates“ als geeignetes Schätzverfahren von Personenparametern unter der Annahme, dass jeder einzelne Personenparameter eine jeweils kleine „A-posteriori“-Verteilung aufweist (vgl. Hahnel 2018: 64ff.).

[149] oder auch „Log Likelihood-Funktion“
[150] im Ergebnis = 0
[151] im Ergebnis = 1
[152] Abkürzung CML; wird bei unterschiedlichsten Rasch-Modellen eingesetzt
[153] Abkürzung WLE; wird auch als „Korrekturterm“ bezeichnet

Zur Überprüfung der Modellkonformität[154] können der Likelihood-Ratio-Test[155] und weitere diverse Informationskriterien, wie die Devianz[156], ermittelt und zur Sichtung herangezogen werden.

> „Ein Likelihoodquotient ist der Quotient zweier Likelihoodwerte derselben Datenmatrix unter zwei unterschiedlichen Modellen …“ (Rost 1996: 330).

Eine Überführung des ermittelten Quotienten in eine χ^2 - Verteilung ist nach zuvor bestimmten Prüfungen, wie u.a. der Gültigkeit des komplexeren Modells und weiteren Kriterien, möglich[157]. Zusätzlich müssen für die Testung die Anzahl der Freiheitsgrade ermittelt werden. Danach wird der kritische Wert aus der χ^2 - Tabelle entnommen[158] (vgl. Rost 1996: 330ff.).

> „Die statistische Hypothese wird *verworfen*, wenn der errechnete Wert der Prüfgröße *größer* ist als der kritische Wert, also unwahrscheinlicher als 5%.“ (Rost 1996: 331).

Am häufigsten wird der Likelihood-Ratio-Test beim Vergleich zweier vorliegender Modelle eingesetzt. Weitere Modelle sind dabei nicht eingeschlossen. Oder nach Rost:

> „Der Vorteil von Modellvergleichen mit dem saturierten Modell[159] liegt darin, daß das saturierte Modell nicht irgendeine beliebige Alternative darstellt, sondern es das höchste Kriterium verkörpert, das es zu erreichen gilt …“ (Rost 1996: 333).

Im Gegensatz dazu ist zu berücksichtigen, dass bei einem ablehnenden Ergebnis (Die beiden Modelle unterscheiden sich! Das saturierte Modell ist weiterhin zu bevorzugen.) keine genaueren Aussagen über die Ursachen gemacht werden können.

Alternativ zum Likelihood-Ratio-Test kann ein globaler Wald-Test zur Überprüfung der Modellgeltung durchgeführt werden. Die gesamten Items werden gemeinsam auf Gruppenabweichungen geprüft (vgl. Strobel 2015: 44ff.).

> „Der globale Wald-Test ist asymptotisch (d.h. für große Stichproben) äquivalent zum Likelihood-Ratio-Test.“ (Strobel 2015: 47).

Aufgrund einer höheren Power wird der Likelihood-Quotienten-Test nach Andersen zur Überprüfung der globalen Modellannahme im vorliegenden Projekt eingesetzt (vgl. Strobel 2015: 42ff.; Strobel 2020). Es werden alle Items simultan geprüft. Der Test überprüft dabei die Nullhypothese. Das Rasch-Modell wird als gültig angenommen, wenn der Likelihood-Quotient gleich 1 ist.

---

154 auf Gültigkeit, Einfachheitskriterium und Geltungsbereich

155 auch Likelihoodquotiententest genannt; Nullhypothese wird überprüft und soll beibehalten werden. Ist dies der Fall, kann von Modellkonformität ausgegangen werden (vgl. Pospeschill 2010: 128). Im vorliegenden Projekt würde das bedeuten, die beiden Modelle (PCM und GPCM) unterscheiden sich nicht.

156 engl. Deviance, weitere Erläuterungen finden sich in Kapitel 15.4

157 = -2 log(Likelihoodquotient) ergibt χ^2 - Wert; unter Berücksichtigung einer ausreichend großen vorliegenden Datenmenge (vgl. Rost 1996: 330)

158 Das Signifikanzniveau wird in der Regel bei 95% festgelegt.

159 Die Erklärung des beobachteten Datenmaterials ist bei Vorliegen eines solchen Modells am besten (vgl. Rost 1996: 333).

Beim Vorliegen von nur kleinen Stichproben wird mit dem sogenannten „Bootstrapping“ gearbeitet. Es werden dafür „künstliche“ Prüfverteilungen aus den gewonnenen Daten erzeugt (vgl. Pospeschill 2010: 142). Diese Möglichkeit kommt in der vorliegenden Arbeit aber nicht zum Einsatz, da die Stichprobengröße für die notwendigen Berechnungen ausreicht.

Als spezielle informationstheoretische Kriterien können u.a. der Akaike`s Information Criterion[160], das Bayesian Information Criterion[161] oder auch der Consistent Akaike`s Information Criterion[162] bestimmt werden. Im Ergebnis wird der jeweilige Informationsindex bzw. letztendlich immer das Modell mit dem niedrigsten Wert priorisiert. Über die angeführten Informationskriterien können aber keine Interpretationen zu den Abweichungen zweier Modelle geschlossen werden (vgl. Pospeschill 2010: 127ff.; vgl. Strobl 2015: 65; vgl. Rost 2004: 157 u. 220; Bühner 2011: 541ff.). Die Entscheidung, welcher Informationsindex herangezogen wird, entscheidet das „Kriterium der Brauchbarkeit“[163] bezogen auf die gewünschte Bestätigung der Raschvalidität im Kompetenztest (vgl. Rost 2004: 342). Das geeignete Maß von AIC und BIC sollte vor allem klein ausfallen. Beim BIC wird eine möglichst niedrige Anzahl an Parametern bevorzugt, somit das möglichst sparsamere Modell mit einer besseren Modellpassung. Eid u. Schmidt weisen darauf hin, dass der BIC unter dem Einbezug von N als dominanteres Kriterium gilt und das „passendere“ Modell herausarbeitet (vgl. Eid u. Schmidt 2014: 191).

Die zu verwendenden Formeln für AIC, BIC und CAIC lauten (vgl. Eid u. Schmidt 2014: 190f.; vgl. Rost 2004: 344) (Gl. 6, 7, 8):

$$AIC = -2 \cdot lnL + 2 \cdot n_{par}$$

$$BIC = -2 \cdot lnL + (\ln n) \cdot n_{par}$$

$$CAIC = -2 \cdot lnL + (\ln n) \cdot n_{par} + n_{par}$$

ln = logarithmiert
L = Likelihood eines Modells
n_{par} = Anzahl der Parameter, die geschätzt werden sollen

[160] „Das *Akaike Information Criterion* (AIC) setzt das Fitkriterium - die logarithmierte Likelihood - ins Verhältnis zur Anzahl der unabhängigen Modellparameter im getesteten Modell.“ (Bensch 2013: 124).

[161] „Für das *Bayes Information Criterion* (BIC) wird der Likelihoodwert eines getesteten Modells, die Anzahl der Modellparameter sowie der Stichprobenumfang verwendet.“ (Bensch 2013: 124).

[162] CAIC; „vereinfachter“ AIC, d.h. noch sparsamer im Modell mit der Folge von nur eingeschränkter Beurteilung von Modellkonstrukten einschließlich seiner Modellinhalte (vgl. Anderson, Burnham, White 1998: 263ff.); Rost und Bensch weisen noch auf die mögliche Konsistenz bei großem N hin und gilt somit als Anpassung des AIC (vgl. Rost 2004: 344; vgl. Bensch 2013: 124). Der CAIC wird hier nur der Vollständigkeit halber erwähnt, aber ist für die weiteren Analysen nicht weiter gewinnbringend.

[163] „Das Besondere an diesem Gütekriterium liegt darin, dass die Einschätzung der Brauchbarkeit einer Theorie nicht wertneutral erfolgen kann. Schon die Festlegung, *wofür* die Theorie brauchbar sein sollte, basiert auf subjektiven Wertsetzungen.“ (Rost 2004: 342).

Ergänzend zu den einzelnen Bestimmungen der spezifischen Informationskriterien kann die „spezifische Objektivität der Vergleiche“[164] ermittelt werden (vgl. Pospeschill 2010: 128). Es erfolgt dabei die Betrachtung der gesamten IC-Kurven mit Blick auf die identische Formung und ihrer parallelen Verschiebungen auf der Abszisse. Die Differenzen in den Itemschwierigkeiten[165] von immer jeweils zwei aufgeführten Items kann abgelesen werden, unabhängig davon die separierte Begutachtung der Personenfähigkeiten (siehe Kapitel 15 und vgl. Pospeschill 2012: 129).

Als weiterer Modellgeltungstest kann der globale „Goodness of Fit Pearson-χ^2-Test“ durchgeführt werden[166]. Hier werden die erwarteten Häufigkeiten im Antwortmuster mit den beobachteten Häufigkeiten innerhalb eines Antwortmusters verglichen und auf deren Abweichungen geprüft (vgl. Hahnel 2018: 86; vgl. Eid u. Schmidt 2014: 192; vgl. Bühner 2011: 534f.). Der Grundsatz lautet: Je kleiner der χ^2-Wert bzw. die Abweichungen, desto besser passt das Rasch-Modell.

Die dafür zu verwendende Formel lautet (Gl. 9):

$$X^2 = \sum_{r=1}^{2^P} \frac{(o_r - e_r)^2}{e_r}$$

P = Anzahl der Items
r = vorliegendes Antwortmuster
o_r = Häufigkeiten des beobachteten Antwortmusters
e_r = Häufigkeiten des erwarteten Antwortmusters (Rasch-Modell Berücksichtigung!)

Ein weiteres Indiz über die jeweiligen TN kann die „Person-Fit-Statistik“ liefern. Sie geht der Frage der Personen-Modell-Beziehung nach. Die Testergebnisse einzelner TN werden auf ihre Plausibilität überprüft. Ein Rückschluss auf die Bearbeitung und Beantwortung der Testung vom TN sowie Defizite innerhalb von Testheften können ermittelt und aufgedeckt werden. Als Konsequenz kann es zum Ausschluss von auffälligen TN im Test kommen, um eine valide Testbewertung beibehalten zu können. Pospeschill und Moosbrugger sprechen von einem „modellinkonformen Verhalten“ als wichtigem Aussagekriterium innerhalb der TN-Beurteilung (vgl. Pospeschill 2010: 129; vgl. Moosbrugger 2012: 246).

Wiederum bezogen auf die einzelnen Testitems kann die Testung einen Hinweis auf das zu messende Merkmal, wie in der vorliegenden Arbeit die praktische Pflegekompetenz, darstellen. Die Höhe des Informationswertes bestimmt die eigentliche Bedeutung des Testitems zur Merkmalsbestimmung unter Einbezug der Personen-

[164] Schwierigkeiten und Fähigkeiten werden „unabhängig voneinander“ ermittelt (vgl. Hahnel 2018: 15).
[165] $\beta_1 - \beta_2 = \delta$
[166] Kurzer statistischer Exkurs: „Ein χ^2-Test (Chi-quadrat) ist ein *Signifikanztest*, d.h. ein Verfahren, mit dem man eine statistische Hypothese prüfen kann. ‚Signifikanz‘ heißt ‚Bedeutsamkeit‘ und der Begriff ‚Signifikanztest‘ bezeichnet die Prüfung (=Test), ob die Abweichung einer anhand der Daten berechneten Größe (=Prüfgröße) von dem Idealwert, den diese Größe bei Geltung der Hypothese annimmt, bedeutsam, also signifikant ist.“ (Rost 2004: 333).

fähigkeit. Die Iteminformationsfunktion[167] hat ihren absoluten Wert bei $\theta_v = \sigma_i$ und kann mit folgender Formel angegeben werden (Gl. 10):

$$I_{iv} = I_i \mid \theta_v = \frac{exp\,(\theta_v - \sigma_i)}{(1 + exp\,(\theta_v - \sigma_i))^{\,2}}$$

I_i = Iteminformation
θ_v = Personenfähigkeit
Auf dem zuvor ermittelten Ergebnis aufbauend, kann die Genauigkeit der Messung des einzelnen TN durch Berechnung des Testinformationswerts[168] erhöht werden (vgl. Pospeschill 2012: 130)(Gl. 11):

$$I_v = \sum_{i=1}^{m} I_{iv}$$

m = Anzahl der Items

Die Präzision der Testung eines TN kann durch Hinzufügen oder Anpassung von Testitems noch weiter angehoben werden (vgl. Moosbrugger 2012: 249).

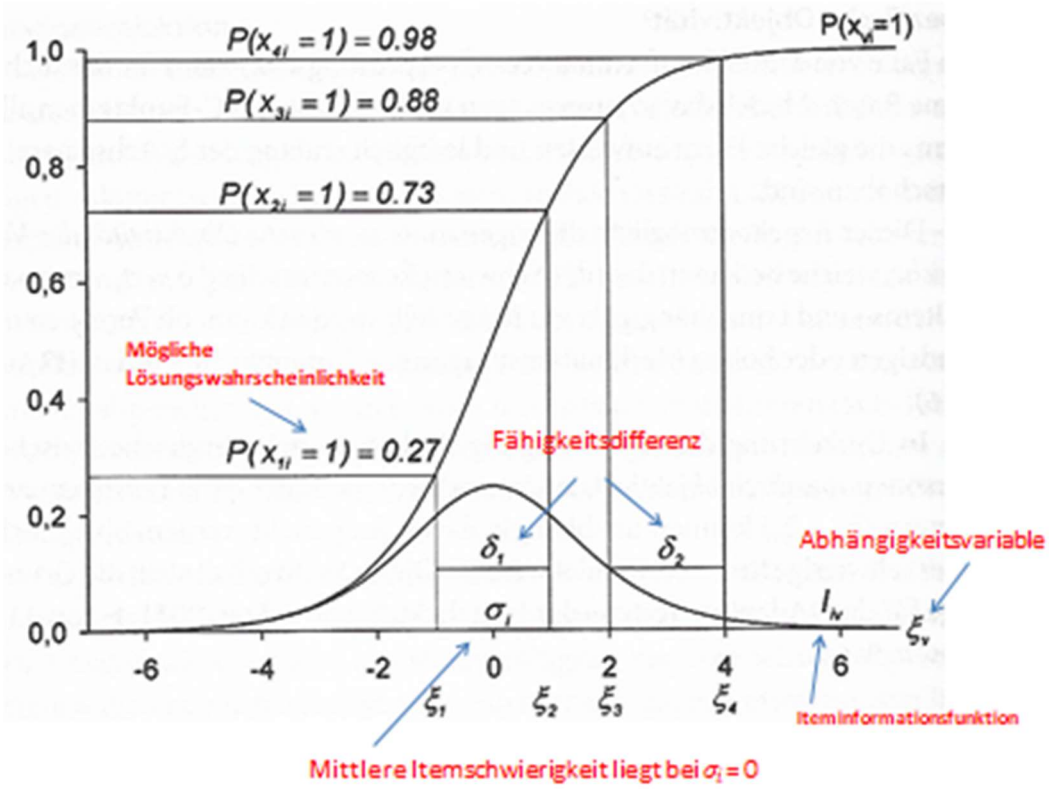


Abb. 11: Spezifizierte ICC zur Iteminformation und Personenfähigkeit (entnommen aus: Moosbrugger 2010: 248 leicht modifiziert von Ries 2019)

[167] abgekürzt mit I_{iv}
[168] abgekürzt mit I_v

Die Ermittlung der zuvor genannten Testparameter soll durch die angefügte, differenzierte Abb. 11 nochmals visuell verdeutlicht werden. Die vorliegende Bedingung ist ein dichotomes Testitem.

Als ergänzendes, statistisches Testinformationskriterium schlägt Pospeschill noch die abschließende Möglichkeit der Bestimmung eines Konfidenzintervalls[169] vor, die die Untersuchung auf ein dichotomes Rasch-Modell abschließt (vgl. Pospeschill 2010: 130).

Sind in den IC-Funktionen verschiedene Steigungen in den dichotomen Testitems aufzufinden, wird ein zusätzlicher Testitemparameter[170] mit einbezogen, der sogenannte „Itemdiskriminationsparameter“[171] aus dem zweiparametrischen-logistischen Modell nach Birnbaum (vgl. Pospeschill 2010: 132). Abzulesen ist er an der steilsten Stelle der IC-Funktion. Der Grundsatz lautet: Je höher die Diskrimination, desto steiler der Anstieg (vgl. Bühner 2011: 504). Zur weiteren Erläuterung und Ergänzung ist die folgende Abb. 12 und Gl. 12 zu berücksichtigen. Das hat zur Folge, dass die spezifische Objektivität aus dem dichotomen Rasch-Modell von TN und vorliegenden Items nicht gegenübergestellt werden kann. Ebenso ist nach Bühner eine Stichprobengröße von über 1000 TN erforderlich, damit eine valide Parameterschätzung stattfinden kann (vgl. Bühner 2011: 514). Eine weitere Vorstellung des 2 PL-Modells übersteigt den Rahmen der vorliegenden Arbeit und geht über die zu Beginn gestellten Fragestellungen hinaus.

[169] „… kennzeichnet denjenigen Bereich von Merkmalsausprägungen, in dem sich 95% (99%) aller möglichen Populationsparameter befinden, die den empirisch ermittelten Stichprobenkennwert erzeugt haben können.“ (Döring, Bortz 2016: 644). Das Konfidenzintervall wird auch als Vertrauensintervall bezeichnet (vgl. Rost 2004: 360f.).

[170] oder auch Trennschärfenparameter genannt (vgl. Bühner 2011: 503)

[171] abgekürzt mit λ_i (= gr. Lambda)

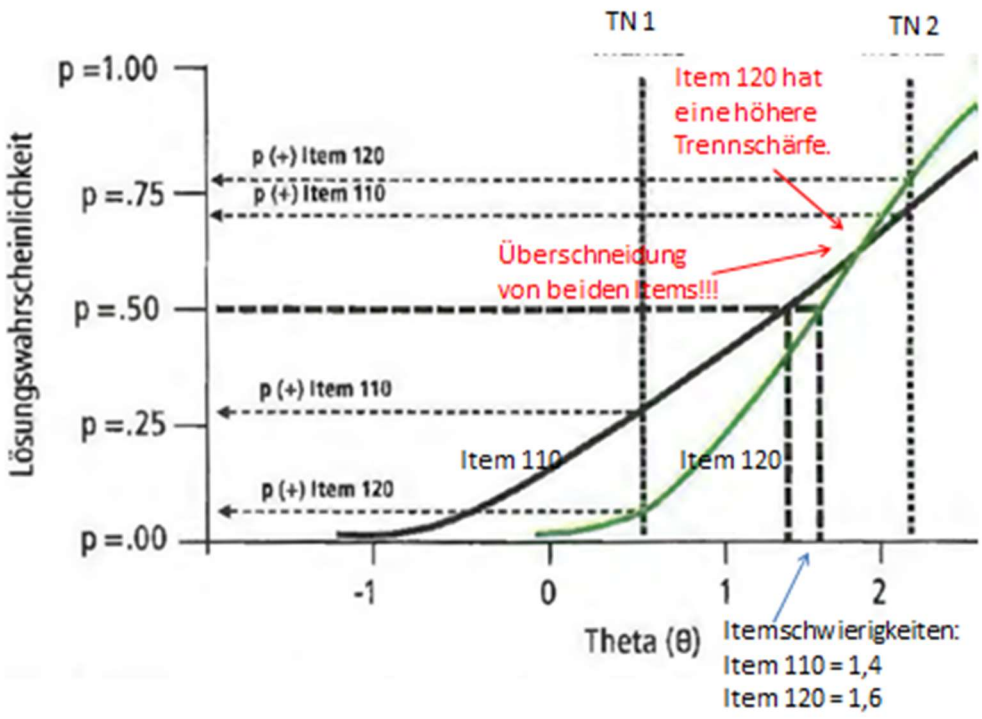


Abb. 12: IC-Funktionen unter Berücksichtigung des 2 PL-Modells (entnommen aus: Bühner 2011: 504 modifiziert von Ries)

Alternativ kann unter Berücksichtigung des Diskriminationsparameters die folgende Formel (entnommen aus: Pospeschill 2010: 132) verwendet werden (Gl. 12):

$$P(x_{vi} = 1) = \frac{exp\ (\lambda_i(\theta_v - \sigma_i))}{1 + exp\ (\lambda_i(\theta_v - \sigma_i))}$$

$P(x_{vi} = 1)$ = Lösungswahrscheinlichkeiten in Verbindung mit θ_v der TN-Fähigkeiten
λ_i = Itemdiskriminationsparameter
In „TEMA“ wurde zur Überprüfung der erhobenen Leistungsdaten aus der IRT der polytomen Modelle das Partial-Credit-Modell und Generalisierte Partial-Credit-Modell eingesetzt. Dies sind spezifizierte Modelle[172], die auf den zuvor erläuterten Grundlagen aufbauen.

15.2 Partial-Credit-Modell

Das Partial-Credit-Modell[173] ist eine Sonderform des Rasch-Modells[174] und wird auch als ordinales Latent-Trait-Modell bezeichnet. Es berücksichtigt nicht nur die dichotomen Items, sondern auch polytome Itemformate. In diesen Fällen werden auch kor-

[172] oder Modellerweiterungen
[173] abgekürzt mit „PCM“; wurde von Masters (1982) entwickelt (vgl. Masters 1982: 149ff.)
[174] als ordinales Rasch-Modell nach Rost 2004 bezeichnet (vgl. Moosbrugger 2012: 266; vgl. Bühner 2011: 515ff.); gehört in die Gruppe der direkten Modelle („divide-by-total-models“), d.h. das Fähigkeitsniveau und die Wahrscheinlichkeit in der jeweiligen Antwortausprägung werden zusammen ermittelt (vgl. Hahnel 2018: 100).

rekte Teilantworten (gestuftes Antwortformat[175]) in die Auswertungen mit einbezogen. Die Darstellung erfolgt auch hier in einer logistischen Funktion, mit dem Unterschied, dass pro einzelner Antwortkategorie bzw. Antwortausprägung eine separate Kurve entwickelt wird. Die einzelnen Kurven werden übereinandergelegt. Die Beurteilung und Einschätzung erfolgt in den jeweiligen Schnittpunkten der „CCC"-Kurven[176] bzw. in der Angabe der Schwellenparameter[177] (vgl. Döring, Bortz 2016b: 487; vgl. Bühner 2011: 516ff.; vgl. Hahnel 2018: 101). Strobl begründet dies damit:

> „… dass die Wahrscheinlichkeit des Übergangs von einer Antwortkategorie c – 1 zur nächsten Antwortkategorie c mithilfe des Rasch-Modells beschrieben wird." (Strobl 2015: 57).[178]

Je höher der ermittelte Wahrscheinlichkeitswert in der jeweiligen Kategorie, desto höher ist die Wahrscheinlichkeit, die nächste Stufe im Kategoriensystem zu beantworten. Dies geschieht in Verbindung mit der Personenfähigkeit auf der Abszisse. Somit wird im Partial-Credit-Modell die Wahrscheinlichkeit ermittelt, in der ein TN sich für eine der vorgegebenen Antwortkategorien[179] entscheidet.

[175] ordinales Antwortmuster, wie z.B. falsch (Kodierung mit 0), könnte zutreffen (Kodierung mit 1), richtig (Kodierung mit 2)
[176] aufgrund des Vorliegens von einzelnen Kategorien bzw. mehrstufigen Antwortmustern, als „Category Characteristic Curves" bezeichnet.
[177] Thresholds; angegeben in τ_i
[178] wird in der Fachliteratur auch mit k bzw. k-1 angegeben (vgl. Bühner 2011: 522)
[179] Kategorienwahrscheinlichkeit nach Bühner (vgl. Bühner 2011: 521)

Um auf die zu verwendende Formel des ordinalen Rasch-Modells zu kommen, wird zunächst ein sogenannter „logarithmierter Quotient“[180] bestimmt (Gl. 13; vgl. Bühner 2011: 521).

$$ln\frac{p_x}{p_0} = \sum_{s=1}^{x}(\theta_v - \tau_{is})$$

ln = natürlicher Logarithmus
p_x = Wert der Wahrscheinlichkeit sich für Antwortkategorie *x* (z.B. Kategorie 1, 2, usw.) zu entscheiden
p_0 = Wert der Wahrscheinlichkeit, sich für Antwortkategorie 0 zu entscheiden
$\sum_{s=1}^{x}$ = Summe der Schwellenparameter bis zur gewählten Antwortkategorie *x*
θ_v = Personenparameter eines TN *v*
τ_{is} = Schwellenparameter (tau) von Item *i* bei Schwelle *s*
Er summiert die Differenzen zwischen θ_v und τ_{is} bis zur ausgewählten Antwortkategorie *x* (vgl. Bühner 2011: 521). Nach Umwandlung der zuvor aufgeführten Gl. 13 ergibt sich die Formel für das Partial-Credit-Modell (Gl. 14):

$$p(X_{vi} = x) = \frac{e^{([x \cdot \theta_v] - \sigma_{ix})}}{\sum_{s=0}^{m} e^{([m \cdot \theta_v] - \sigma_{im})}}, x = 0, 1, 2, \dots, x$$

$p(X_{vi} = x)$ = Wahrscheinlichkeit einer Person *v*, bei Item *i* sich für die Kategorie *x* zu entscheiden
m = Anzahl der Schwellen[181]
s = Schwelle

Im Zähler bedeutet σ_{ix}(= $\sum_{s=0}^{x}\tau_{ix}$) das Vorliegen der Gesamtsumme der Schwellenparameter eines Items bis zur jeweils ausgewählten Antwortkategorie des TN. Eine Aufsummierung wird beim Überspringen von einzelnen Schwellen notwendig. Im Nenner wiederum wird die Summe aller Schwellenparameter σ_{im}(= $\sum_{s=0}^{m}\tau_{im}$) des Items *i* abgebildet (vgl. Bühner 2011: 522). Daran zeigt sich nach Bühner die Konsequenz:

> „…alle Schwellenparameter des Items tragen zur Vorhersage der Wahrscheinlichkeit bei, mit der eine Person eine bestimmte Kategorie wählt.“ (Bühner 2011: 522).

Ein gleichzeitig notwendiges Kriterium für die Definition eines Partial-Credit-Modells ist gegeben, wenn eine Ordnung der einzelnen Schwellenparameter vorliegt (vgl. Bühner 2011: 522; vgl. Rost 2004: 205; vgl. Strobl 2015: 63; vgl. Moosbrugger 2012: 265f.; vgl. Eid u. Schmidt 2014: 231u. 238).

Grundsätzlich sind dieselben Bedingungen wie beim dichotomen Rasch-Modell, z.B. Schätzmethoden und Modellkonformität, zu berücksichtigen (vgl. Bensch 2013: 125ff.).

[180] Schwellenwahrscheinlichkeitsquotient
[181] definiert als (k-1 = Anzahl der vorhandenen Antwortkategorien - 1 Schwellenparameter)

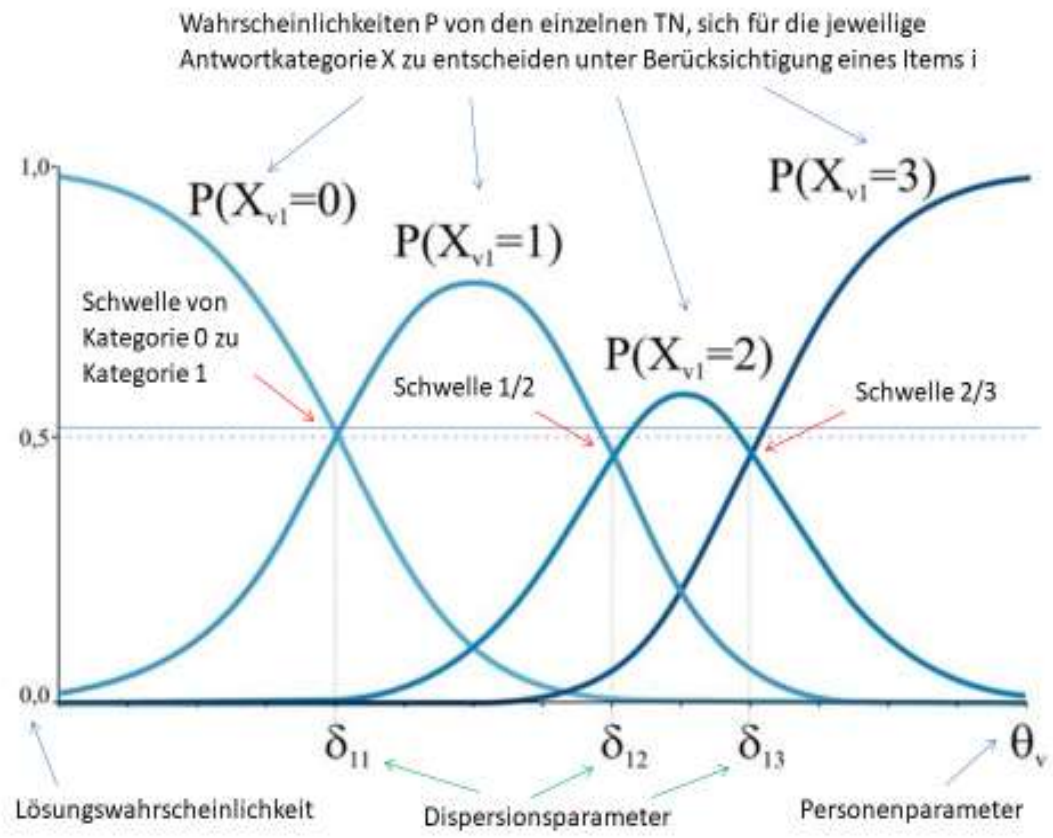


Abb. 13: Beispiel eines Partial-Credit-Modells mit drei Schwellenparametern (entnommen aus: Hahnel 2018: 101 modifiziert von Ries)

Die aufgeführte Abb. 13 stellt exemplarisch ein Testitem mit vier verschiedenen Antwortkategorien (0 bis 3) und drei angezeigten Schwellen dar (4 Antwortkategorien - 1 Schwellenparameter = 3 Schwellen). Die Angabe δ [182] steht für den Schwellenabstand[183] des Items. Je breiter die Schwellendistanz[184], desto häufiger finden sich Antworten in den mittleren Kategorien. Das heißt aber gleichzeitig, dass die Streuung der Antwortvariablen klein ist[185] und daran gekoppelt die Trennschärfe der Items sich ebenfalls als gering zeigt (vgl. Rost 2004: 217). In der statistischen Anwendung sollten auf Empfehlung von Rost Items mit breiten Schwellendistanzen bevorzugt werden, damit eine bessere Trennung von TN in den mittleren Bereichen stattfinden kann (vgl. Rost 2004: 370f.; vgl. Bensch 2013: 131).

[182] Streuungsparameter oder inverser (= „umkehren“) Trennschärfeparameter (vgl. Rost 2004: 217)
[183] Schwellendistanz
[184] Der Abstand kann auf der Abszisse abgelesen werden.
[185] kleine Dispersion über die gesamten TN nach Rost (vgl. Rost 2004: 216)

15.3 Generalized-Partial-Credit-Model

Das ordinale Rasch-Modell bzw. Partial-Credit-Modell kann unter zusätzlicher Ergänzung des Trennschärfenparameters[186] erweitert werden, d.h., es liegt eine spezielle Form eines ordinalen 2 PL-Modells[187] auf Grundlage der Rasch-Modelle vor (vgl. Muraki 1992: 160; vgl. Bühner 2011: 577ff.; vgl. Eid u. Schmidt 2014: 253 u. 255f.).

Das Besondere im Vergleich zum PCM ist, dass pro Schwellenfunktion eine eigene Steigung vorliegt, d.h. jeweils ein separater Diskriminationsparameter mit einer differierenden Steilheit (vgl. Bühner 2011: 579). Die Herausforderungen zeigen sich im GPCM in Anlehnung an das 2 PL-Modell durch u.a. geringere Restriktion[188] (siehe Kapitel 15.1). Die Autoren Eid u. Schmidt ergänzen jedoch, dass durch die Reduktion von „statistischer Strenge“ das GPCM in statistischen Auswertungen deshalb nur selten zum Einsatz kommt (vgl. Eid u. Schmidt 2014: 253).

In „TEMA“ jedoch wurden das PCM und das komplexere GPCM zur Bestimmung der „Model Fit“-Werte für die Testantworten einander gegenübergestellt und verglichen. Die Interpretation auf Raschvalidität zur Messung von praktischer Pflegekompetenz konnte damit nachgewiesen werden.

Die Wahl für die Anwendung der IRT-Methode wurde vom Forscherteam durch die Aussage im Abschlussbericht des „TEMA“-Projektes so begründet:

> „Das Skalierungsmodell repräsentiert den vermuteten Zusammenhang zwischen der abzubildenden latenten Kompetenz und dem Testverhalten, und wird gewöhnlich als logistische Funktion beschrieben. In Abhängigkeit von der Art der entwickelten Testaufgaben finden unterschiedlich komplexe Modelle Anwendung, …“ (Döring et al. 2014: 158).

Aufgrund der durchgeführten Analysen und einer möglichen Parallelität über die Altenpflege hinaus wird die Auswertung für das „Large-Scale-Assessment-Instrument" so weiterverfolgt.

Die zentralen Auswertungen zur IRT im hier vorliegenden Projekt der Gesundheits- und Krankenpflege schließt sich im folgenden Kapitel 15.4 an unter Berücksichtigung der Forschungsfragen aus Kapitel 1.2. Beide Projekte werden miteinander verglichen.

15.4 Ergebnisse aus der IRT-Datenanalyse in R

In Kapitel 13.2 und 13.3 wurde auf die Verwendung des Statistik-Programms von R einschließlich Packages in der vorliegenden Arbeit eingegangen. Zu den notwendigen IRT-Berechnungen ist auf Grundlage der zur Verfügung gestellten R-Skripte aus „TEMA“ ein separates, umfassendes R-Skript entwickelt und programmiert wor-

[186] itemspezifischer Diskriminationsparameter β_i

[187] engl. „Generalized-Partial-Credit-Model” nach Muraki 1992 oder dt. „Generalisiertes Partial-Credit-Modell” mit der Abkürzung “GPCM” (vgl. Muraki 1992: 159ff.)

[188] geringere Strenge von Annahmen und damit Verlust von positiven messtheoretischen Ansätzen

den (Anhang 18). Weitere Programme für die IRT-Analysen[189], wie z.B. ConQuest[190], WINMIRA[191], waren zur Beantwortung der Forschungsfragen nicht notwendig, wie in diverser Literatur vorgeschlagen. Es existieren in der freien Software R bereits geeignete Packages, wie „ltm“, „Wright Map“, „irtoys“ usw., die für die notwendigen Datenanalysen umfänglich sind.

Im statistischen Forschungsüberblick wurden die gesamten Testitems und die TN-Fähigkeit mit selektiver Bewertung der polytomen Items analysiert. Es wurden unterschiedliche Personen- und Itemschätzwerte ermittelt und eine Überprüfung der Modellkonformität von 1 PL-Modell bzw. PCM und 2 PL-Modell bzw. GPCM erreicht, unter Verwendung der möglichen Testungen, wie z.B. Log Likelihood, Deviance, AIC und BIC. Zum Vergleich der beiden Modelle kam ergänzend ein Likelihood-Ratio-Test zum Einsatz. Zur besseren Veranschaulichung wurde für die dichotomen Items das 1 PL-Modell als ICC-Kurven bzw. CCC-Kurven geplottet und bewertet. Eine Gegenüberstellung der Ergebnisse aus der Gesundheits- und Krankenpflege und Altenpflege wurde durchgeführt. Abschließend wurden die Modell- und Itemfit-Werte betrachtet. Aus allen statistisch durchgeführten Analysen werden im Anschluss die jeweiligen zentralen Ergebnisse aufgeführt und erläutert unter Einbezug von vorhandenen Parametern aus dem Ursprungsprojekt „TEMA“.

Tab. 6: Auszug aus den ermittelten Ergebnissen von Itemschwierigkeiten aus TAM

Itembezeichnung	Rasch (PCM)	2PL (GPCM)
PT11A05_Cat1	0.730	0.724
DR11A73_Cat1	0.683	0.729
DR11A74_Cat1	0.135	0.110
IK12A06_Cat1	0.232	0.224
DR12A07_Cat1	-0.447	-0.461
PT12A08_Cat1	-0.658	-0.673
IK12A09_Cat1	-1.044	-1.072
DR12A10_Cat1	-0.719	-0.735
PT12A11_Cat1	-0.610	-0.618
PT12A12_Cat1	0.601	0.587

blau = höhere Itemschwierigkeit

Die Itemschwierigkeiten differieren im PCM und GPCM häufig nur gering. Ein meist minimal höherer Schwierigkeitsgrad findet sich im GPCM. Letztlich zeigt sich überwiegend eine entsprechende Breite zwischen leichten und schweren Items und damit eine passende Komplexität für schwächere und stärkere Auszubildende für den gesamten „Large-Scale-Assessment-Kompetenztest“ (Anhang 19 u. 20).

[189] wurde in z.B. Bühner 2011: 557ff. aufbereitet
[190] wurde vom „Australian Council for Educational Research Ltd“ (ACER) für die PISA-Datenauswertungen entwickelt
[191] Statistiksoftware für Rasch-Modelle, gemischte Rasch-Modelle, usw.

Tab. 7: Auszug aus Diskriminationsermittlungen in TAM

	Cat0	Cat1	Cat2	Cat3
DR21A31	0	0.550495518		
IK21A32	0	0.007571891		
DR21A34_polytom	0	0.705574699	0.7708023	
PT21A35neu_polytom	0	0.436177486	0.7775528	
IK21A72	0	0.615088563		
DR21A75_polytom	0	0.425117942	-0.3795615	0.7468584
PT22A25	0	0.340112443		
PT22A26	0	0.529175315		
IK22A27	0	1.556528129		
IK22A28_polytom	0	-0.068695848	-0.6739307	0.7409547

Die Diskriminationen zeigen in der Gesamtbetrachtung[192] beim GPCM, wie erwartet, unterschiedliche Steigungen an ihren Wendepunkten an. In der Kategorie 1 finden sich Werte von - 0.1817 bis 1.654 (Tab. 7 Auszug von -0.0686 bis 0.6150), Kategorie 2 von - 0.6739 bis 1.4901 (Tab. 7 Auszug von -0.6739 bis 0.7775) und Kategorie 3 von - 0.2066 bis 1.4322 (Tab. 7 Auszug von 0.7409 bis 0.7468). Die Merkmalstrennungen stellen sich als unterschiedlich, aber konform dar, gleichzeitig werden die Merkmalsausprägungen günstig und flexibel erfasst, wie sich in den ermittelten Diskriminationsparametern zeigt (Anhang 21).

Zur Gesamterfassung aller signifikanten Aufgaben- und TN-Ergebnisse wird bei Item-Response-Analysen bevorzugt die folgende Darstellung (Abb.14) erzeugt und interpretiert (vgl. Hahnel 2018: 33).

[192] Die angegebenen Kategorienwerte 1, 2, 3 sind auf alle Itemdiskriminationen bezogen (Anhang 21).

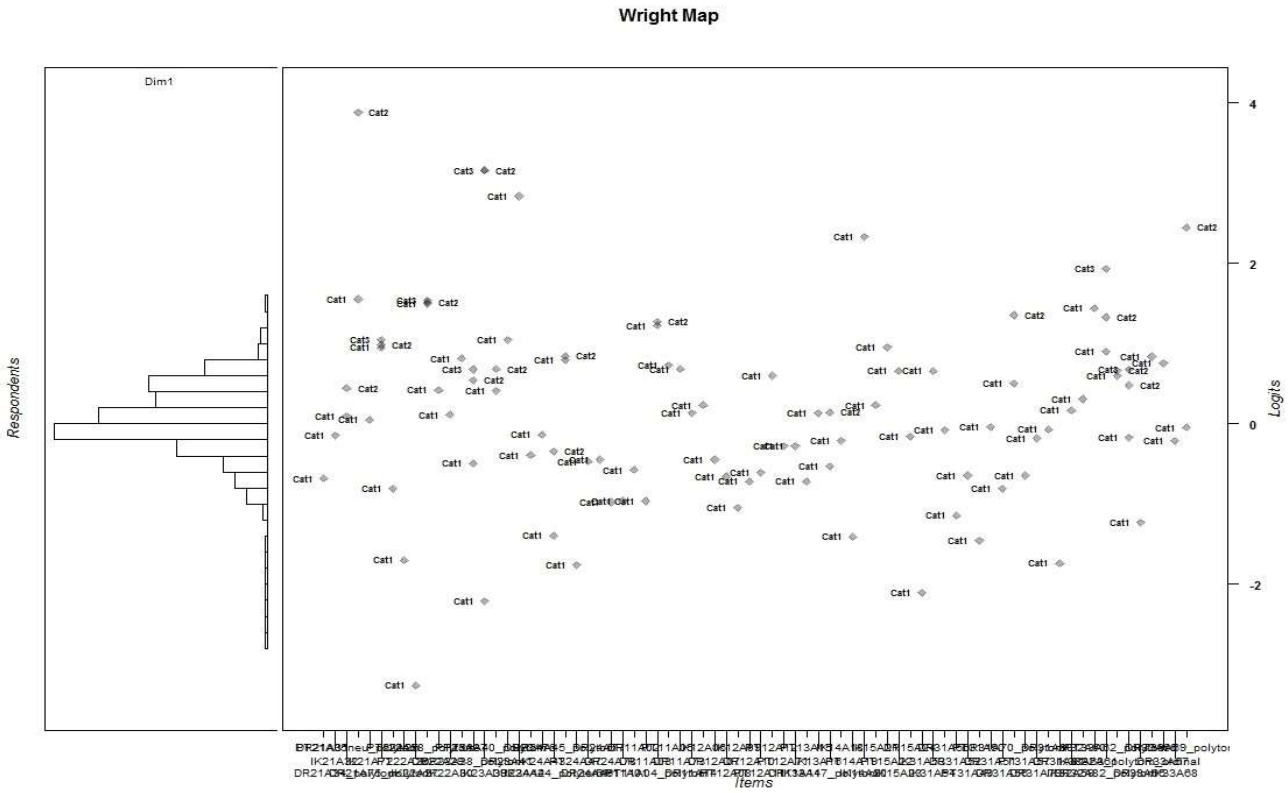


Abb. 14: Übersicht aller dichotomen und ordinalen bzw. polytomen Items im Kompetenztest in der Gesundheits- und Krankenpflege[193]

Im erzeugten Plot (Abb.14), genannt „Wright Map“, werden die Respondents[194] und Logits[195] der Items aufgeführt und gegenübergestellt (vgl. Irribarra & Freund 2016). Links ist ein Histogramm mit den vorliegenden Personenfähigkeiten bzw. den TN aus der Gesundheits- und Krankenpflege zu sehen. Die mittlere Respondents (= 50%) würde bei einem Wert 0 Logits liegen. Die gesamten Items des Kompetenztests werden auf der rechten Seite des Plots aufgeführt. In der Regel sollten die Aufgabenschwierigkeiten bei Logits zwischen -3[196] und +3[197] liegen.

Es zeigt sich in den Ergebnissen die höchste Respondents bei einem Logit von 0 und darunter, d.h. von 50% und leicht weniger. Die Itemschwierigkeit bewegt sich überwiegend im Bereich zwischen -2 und +2. Es stellen sich nur wenige Items über- bzw. unterbewertet dar. Der Anteil an „Ausreißeritems“ im Minusbereich ist für TN mit niedriger Fähigkeit als „Motivator“ und für TN mit höherer Fähigkeit als „Warm up“ und bei Logits über +2 als „Herausforderung“ zu verstehen. Somit zeigt sich anhand der vorliegenden 76 modellierten Itemparameter ein empirisch sehr gut entwickeltes Kompetenzmessinstrument für die Gesundheits- und Krankenpflege zur Messung von praktischer Pflegekompetenz. Bezogen auf die Personenfähigkeit zeigen sich fast identische Ergebnisse wie in „TEMA“ (siehe Kapitel 14.2 u. 14.3).

[193] Die einzelnen Itemparameterwerte finden sich im Anhang 12 und Anhang 13.
[194] Befragte Personen bzw. TN = Personenfähigkeit
[195] Aufgaben- bzw. Itemschwierigkeiten
[196] recht leichte Items bzw. Kategorien
[197] recht schwere Items bzw. Kategorien

Entsprechend heißt das, dass der Anspruch an die TN im dritten Ausbildungsjahr bzw. entsprechenden Studiensemester als überwiegend herausfordernd einzuschätzen ist!

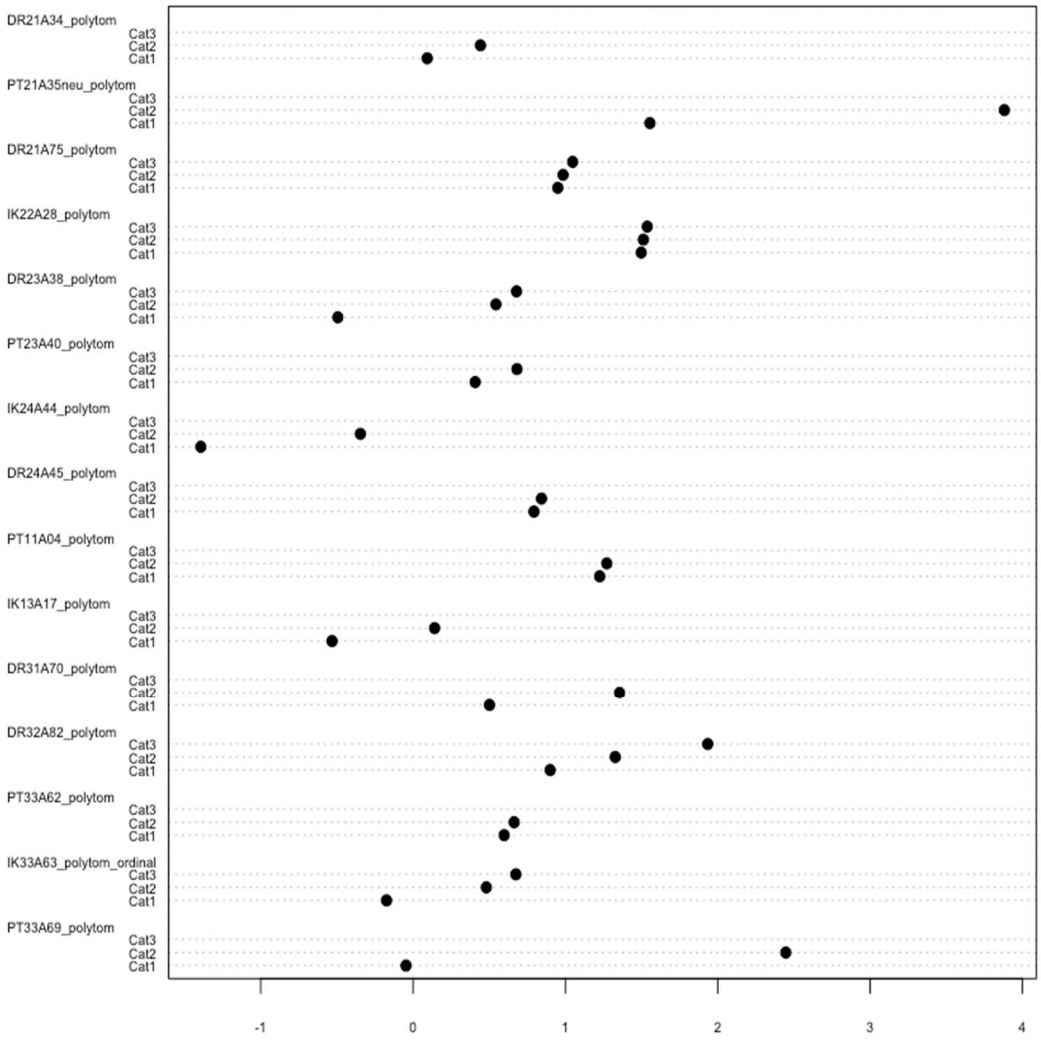


Abb. 15: Übersicht der Itemschwierigkeiten der polytomen Testitems[198]

Der aufgeführte Plot (Abb. 15) stellt nochmals genauer als in der zuvor aufgeführten „Wright Map“ ausschließlich die polytomen Items mit ihren einzelnen Antwortkategorien in Form einzelner Kontinua dar (vgl. Koller, Alexandrowicz, Hatzinger 2012: 180). Die Messung der latenten Variable, der praktischen Pflegekompetenz, soll damit abgebildet werden[199]. Die meisten polytomen Items befinden sich von ihren Schwellenwerten her im Bereich zwischen 0 und 2. Innerhalb der Antwortkategorien 1 und 2 differieren die polytomen Items A35 (gerundet 1,62 / 3,66)[200] und A69 (gerundet 0,02 / 2,22) sehr stark. Dies lässt die Folgerung zu, dass der Anspruch zur vollständigen Lösung der Aufgaben (A35 Verbandwechsel einschließlich Hygie-

[198] Hinweis: Im Plott sind generell drei verschiedene Kategorien (= Cat) aufgeführt, aber die Cat 3 ist bei 10 Items nicht vorhanden, d.h. es liegt hier nur Cat 1 und Cat 2 vor.
[199] unter Verwendung der Werte aus dem GPCM
[200] Hinweis: ermittelte Schwellenwerte immer von Kategorie 1 nach Kategorie 2 angegeben!

neregeln und A69 Versorgung eines Verstorbenen) eine hohe Personenfähigkeit voraussetzt. Dagegen lässt sich das Item A44 (gerundet -1,28 / -0,78) mit einer leichteren Aufgabenschwierigkeit in Antwortkategorie 1 und 2 bewerten (A44 Ernährungsberatung eines Patienten mit Diabetes Mellitus) und zeigt einen niedrigeren Fähigkeitsanspruch. Allgemein sprechen auch hier die polytomen Testitems für eine hochwertige empirische Modellierung.

Im Vergleich zu den „TEMA"-Ergebnissen sind teils ähnliche, teils aber auch deutliche Unterschiede in den Itemschwierigkeiten zu erkennen. Beispielsweise pendeln bei Item 75 (Beurteilung von Wundverhältnissen) die Werte in der Altenpflege leicht unter 0 und fast bis 2, während in der Krankenpflege die Werte fast synchron bei einem Wert um die 1 liegen. Dies lässt den Schluss zu, dass ein Zusammenhang zwischen den Schwerpunkten der praktischen Ausbildungsinhalte und den curricularen Vorgaben in den jeweiligen pflegerischen Berufsfeldern besteht. Das Antwortverhalten ist deshalb von den beiden TN-Gruppen (Altenpflege und Gesundheits- und Krankenpflege) als nicht konform in der Evaluation zu bewerten.

Bei der globalen Beurteilung des PCM und GPCM lassen sich die folgenden zentralen Ergebnisse und Folgerungen in der Tabelle ermitteln und interpretieren.

Tab. 8: Modellvergleich des PCM und GPCM in der Gesundheits- und Krankenpflege für das Antwortverhalten

(mehr Parameter vorhanden)
höhere Komplexität!

	Rasch (PCM)	2PL (GPCM)
Number of estimated parameters	99	196
LogLikelihood (LL)	-18,414.02	-18,169.39
Deviance	36828.04	36338.78
Akaike Information Criterion (AIC)	37,026.04	36,730.78
Bayesian Information Criterion (BIC)	37,418.43	37,507.64
Likelihood-ratio Test	x^2=489.2628, df=97, p=0	

Modell wird bevorzugt!

Vergleich des PCM mit dem GPCM
-> Unterscheidung der beiden Modelle!

Der Vergleich der beiden Modelle unter Berücksichtigung von 76 Itemparametern zur Ermittlung des latenten Kompetenzbereichs Pflege zeigt, dass es zunächst zu einem nicht ganz eindeutigen Ergebnis kommt (Tab. 8). Das PCM und das GPCM zeigen im Likelihood-Ratio-Test eine Unterscheidung auf[201]. Es liegt eine höhere Deviance[202] beim PCM vor, was für das GPCM spricht. Die LogLikelihood[203] und der AIC würden das komplexere GPCM ebenfalls bevorzugen, trotz seiner reduzierten Genauigkeit bzw. höheren Komplexität, während der BIC, als „dominanteres Kriterium“ mit statistischer „Strenge“, sich für das parameterreduzierte PCM entscheiden würde und somit auch für die Entscheidung zum bevorzugt passenden sparsamen PCM-Modell gilt. Ebenso ist bei der Evaluation der ermittelten Daten auf die Berücksichtigung der reduzierten Stichprobengröße (= 389TN) hinzuweisen, bei Betrachtung des komplexeren GPCM von 97 weiteren Modellparametern kommt es zu größeren Ungenauigkeiten in den Berechnungen und als Folge im Gesamtergebnis. Die Befunde decken sich mit den aus „TEMA“ vorliegenden und kongruent der Beibehaltung der Raschvalidität im Modellvergleich.

Somit bestätigt sich bis heute die mittelalterliche Theorie nach v. Ockhams, dass mit der einfachsten Erklärung auch immer die Wahl des mathematisch sparsamsten Modells einhergeht. Es wird als das „Sparsamkeitsprinzip“ bezeichnet und kann auf verschiedenste Gegenstandsbereiche der Wissenschaft übertragen werden (vgl. Beckmann 2015: 42ff.; vgl. Hähnel 2015: 124).

Die EAP-Reliabilität[204] unter der Bedingung des spezifischen 1 PL-Modell und 76 verwendeten Items liegt bei EAP = 0.822 mit einer Varianz von 0.2468[205]. Es zeigt sich somit eine gute Reliabilität der Items. Aus Sicht der TN stellen sich die Items jedoch als anspruchsvoll und schwierig zu beantworten dar. Bei Berücksichtigung der Dichotomie im PCM ist eine hohe Varianz vorhanden, d.h., es liegt ein komplexer Informationsgehalt für die modellierten Items mit einem aussagekräftigen Erklärungscharakter vor. Eine weitere Optimierung wird durch Schätzung der WLE-Reliabilitäten[206] von WLE = 0.868 im gesamten Test erreicht. Es ist darauf hinzuweisen, dass im durchgeführten Kompetenztest in der Gesundheits- und Kran-

[201] Ermittelt über: 1. Schritt -> χ^2-Wert oder auch Prüfgröße von 489.26; 2. Schritt -> Freiheitsgrade von 97 (über Number of estimated parameters GPCM 196 minus PCM 99); 3. Schritt -> Ablesen des kritischen Wertes aus χ^2-Tabelle bei einem Signifikanzniveau von 95% bei fast 100 df liegt bei 124,3 df; 4. Schritt -> Ergebnisinterpretation: Die beiden Modelle unterscheiden sich. Das saturierte Modell (PCM) wird bevorzugt. Es spricht für das spezifizierte Rasch-Modell und seine vorhandene „Reduziertheit“. Weitere Interpretationen können anhand des vorliegenden Ergebnisses nicht vorgenommen werden.

[202] Devianz oder auch Abweichung; wird berechnet über die -2* log Likelihood (Wahrscheinlichkeit), die Differenz der beiden Devianzen ergibt den χ^2-Wert; die Devianz alleine hat aber keine Aussagekraft. Sie wird benötigt für die Bestimmung des Deviance Information Criterion (= DIC) zum Modellvergleich. Das Kriterium ist auf kurzfristige Vorhersagen gerichtet und ist eine Verallgemeinerung des AIC (vgl. Spiegelhalter et al. 2002: 585ff.).

[203] der LL-Wert, der stärker Richtung 0 tendiert, wird bevorzugt (vgl. Hahnel 2018: 60ff.)!

[204] „Expected A Posteriori Estimates“ -> Personenparameterschätzer, der nach der MML-Methode (= „Marginal Maximum Likelihood“ Methode = Beschreibung der Verteilung der Itemparameter) eingesetzt wird zur Itemparameterschätzung unter Verwendung der zuvor ermittelten Verteilungsparameter.

[205] bei einem Maximalwert von 0.25 entspricht dies einer Lösungswahrscheinlichkeit von 50% (0 oder 1) des TN, er gilt als geeignet für die Teilnahme an weiteren Untersuchungen!

[206] „Weighted Likelihood Estimates“ (= gewichtete Wahrscheinlichkeitsschätzung) -> Parameterschätzung/Personen; dieser ist weniger anfällig als der EAP-Schätzer!

kenpflege keine Items weggelassen wurden, aufgrund der vorliegenden positiven Reliabilitätswerte. Im Vergleich dazu zeigte sich in „TEMA“ die maximale WLE-Reliabilität bei WLE = 0.76 nach durchgeführten Ausschlüssen von Testitems.

Bei der Überprüfung der globalen Modellpassung auf Signifikanz mit dem Likelihood-Quotienten-Test nach Anderen zeigt sich eine Ablehnung des Gesamtmodells, d.h. es liegt keine Raschvalidität vor[207]. Eine lokale stochastische Unabhängigkeit ist nicht vorhanden. Als Splitkriterium wurde „der Mittelwert der Rohwerte“ (Strobel 2015: 78) bezogen auf die Personenrandsummen eingesetzt. Der p-Wert liegt bei 0 und somit unter dem erforderlichen p-Wert von 0,05. Zwei emotionale Items wurden bei den Berechnungen in R mit dem Package „eRm“ automatisch eliminiert, da diese (PT11A04_polytom 337 TN und IK22A28_polytom 371 TN) von der überwiegenden Mehrheit der TN nicht beantwortet wurden (Anhang 22).

```
Warning in LRtest.Rm(out2):

The following items were excluded due to inappropriate response pat-
terns within subgroups:

PT11A04_polytom IK22A28_polytom

Full and subgroup models are estimated without these items!

Andersen LR-test:
LR-value: 274.31
Chi-square df: 90
p-value: 0
```

Im durchgeführten „Goodness of Fit Pearson χ^2-Test“ als weiterer globaler Modellgeltungstest bestätigt sich ebenfalls die Ablehnung der Modellgüte nach Rasch (Tab. 9 u. Anhang 18).

Tab. 9: Ergebnisse aus der globalen Modelltteststatistik in „TAM“

Test of Global Model Fit (Maximum Chi Square)

maxX2	Npairs	p.holm
32.39739	2850	0.01123

Fit Statistics

100*MADCOV	SRMR	SRMSR	MADaQ3	pmaxX2
1.418	0.051	0.065	0.045	0.011

Der maximale χ^2-Test (maxX2) im ersten Teil der Ausgabe in R gibt den höchsten Wert der Berechnungen der bedingten Unabhängigkeit für jedes Itempaar von Antworten (Npairs 2850) an. Darüber hinaus ist der korrigierte Wahrscheinlichkeitswert (p.holm) sehr klein. Er berücksichtigt die Gesamtzahl der 2850 durchgeführten Tests. Der maxX2 soll möglichst klein sein und deutet hier auf einige „Hotspots“ hin. Diese Analyse spricht für die Interpretation einer unerwünschten Passform und somit einer fehlenden lokalen stochastischen Unabhängigkeit. So ist es laut Lampria-

[207] siehe folgender Ergebnisausschnitt, das dafür notwendige Skript für R findet sich im Anhang!

nou häufig der Fall, dass die Nullhypothese der perfekten globalen Passform abgelehnt wird. Die Betrachtung der einzelnen Items im Verlauf der weiteren Analysen ist deshalb ein entscheidendes Kriterium zum Auffinden und Entfernen von störenden Mustern (vgl. Lamprianou 2020).

Im zweiten Teil der Ausgabe „Fit Statistics“ sind sehr gute Ergebnisse zu erkennen. Als zentrales Ergebnis zeigt sich der MADaQ3 als mittlerer absoluter Wert der durchgeführten 2850 Korrelationen als geeignet, da dieser gegen 0 geht. Ebenso geht der SRMSR[208] und SRMR[209] gegen 0, was für das Gesamtmodell spricht. Die MADCOV Statistik beruht auf dem Vergleich von Restkovarianzen von Item-Paaren, d.h. die mittlere absolute Abweichung der Restkovarianz wird in der Ausgabe angegeben. Sie wird in „TAM“, aufgrund der Darstellung, mit 100 in der Ausgabe multipliziert und gibt den Wert von 1.418 bzw. 0.01418 an. Sie geht somit, wie gewünscht, ebenfalls auch gegen 0 (vgl. Döring, Bortz 2016: 967; Lamprianou 2020: 153).

Als Anmerkung zu den durchgeführten Untersuchungen zur globalen Modellgüte ist zu sagen, dass im „TEMA“-Projekt in den zur Verfügung stehenden Veröffentlichungen und internen Unterlagen keine globalen Modellgeltungstests zu recherchieren waren, obwohl zur Überprüfung einer möglichen Modellpassung und vorhandenen Raschvalidität notwendig. Der Schwerpunkt war in der Bewertung der einzelnen Items zu finden (vgl. z.B. Kaspar et al. 2016: 13).

Unter Verwendung der Item-Fit-Statistik[210] wurden die einzelnen Kompetenzitems mit dem Focus der TN-Fähigkeiten und ihrem Antwortverhalten über die gesamten Antwortkategorien betrachtet und u.a. auf Overfit[211]-Werte und Underfit[212]-Werte beurteilt.

Für die Überprüfungen wurden die erzeugten „Expected Response Curves“[213] mit den Packages „irtoys“ und „TAM“ verwendet. Es werden die beobachteten mit den erwarteten Lösungshäufigkeiten angezeigt, von der Forscherin verglichen und auf Abweichungen geprüft. Stellvertretend sollen hier zwei Item-Fit Beispiele[214] aus dem ordinalen Rasch-Modell kurz vorgestellt werden. Weitere Itembeispiele finden sich in Anhang 22.

---

208 Standardized Root Mean Square Root of Squared Residual

209 Standardized root mean square residual

210 „Itemfit-Maße zeigen an, wie gut ein Item zu dem Testmodell passt, das als Grundlage der Testauswertung dient.“ (Rost 2004: 371).

211 „Die Kurve der **beobachteten** Werte ist **steiler** als die der **erwarteten** Werte.“ (Hahnel 2018: 90).

212 „Die Kurve der **beobachteten** Werte ist **flacher** als die der **erwarteten** Werte.“ (Hahnel 2018: 90).

213 als „Erwartete Antwortkurven“ in „TEMA“ benannt (vgl. Döring et al. 2014); wird auch als „Expected Scores Curves“ bezeichnet (vgl. Döring et al. 2014: 996ff.)

214 Itembeispiele mit dem fachlichen Inhalt und Hintergrund: A41 Bobath-Konzept und A82 Dekubitusrisiko (vgl. „TEMA“-Projektgruppe 2014: 18 u. 29).

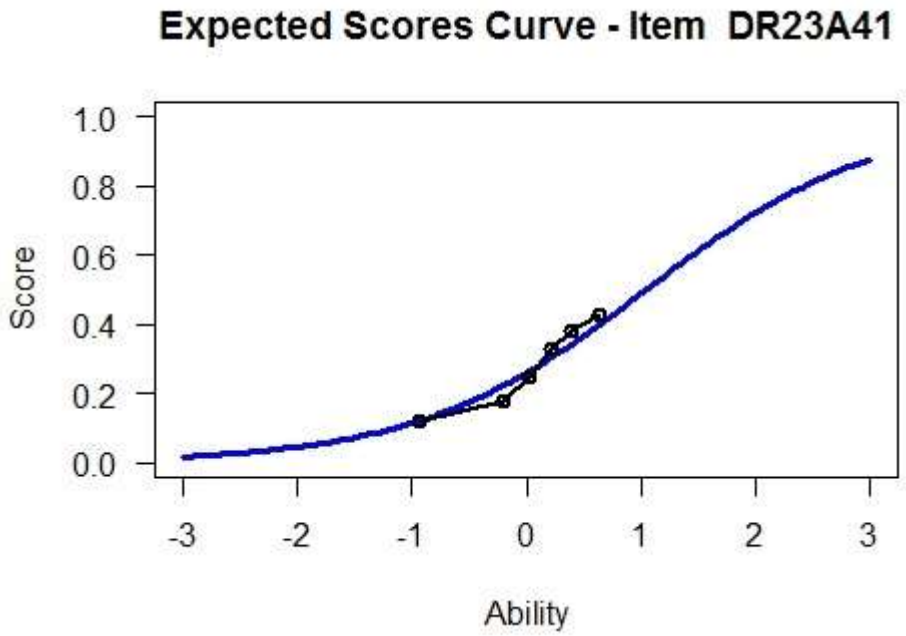


Abb. 16: Dichotomes Itembeispiel A41 aus dem Bereich DR

Die blaue Kurve stellt die erwarteten Lösungshäufigkeiten in einer ICC dar. Die beobachteten Lösungshäufigkeiten werden in einer schwarzen Kurve mit Punkten dargestellt (Abb. 16). Die latente Fähigkeit liegt zwischen -1 und 0.8, es ist eine schmale Spanne von Werten für die Einschätzung des ansteigenden latenten Merkmals „praktische Pflegekompetenz“ vorhanden. Im Score[215] sind es unter 0.2 bis an die 0.5. Die bereits zuvor ermittelte erhöhte Itemschwierigkeit von 1.0451 passt zu dem ermittelten Ergebnis. Die beobachteten Werte sind etwas höher bzw. die Kurve zeigt sich geringfügig steiler als die der erwarteten Werte im Bereich von über 0 und unter 1 in der „Ability“. Ein minimaler Overfit liegt dort vor. Als Gesamtbild zeigen sich beobachtete und erwartete Lösungshäufigkeiten als fast konform.

[215] Ergebnis, Resultat; hier: Score für die jeweiligen Antwortkategorien (z.B. 0-1-2-3)

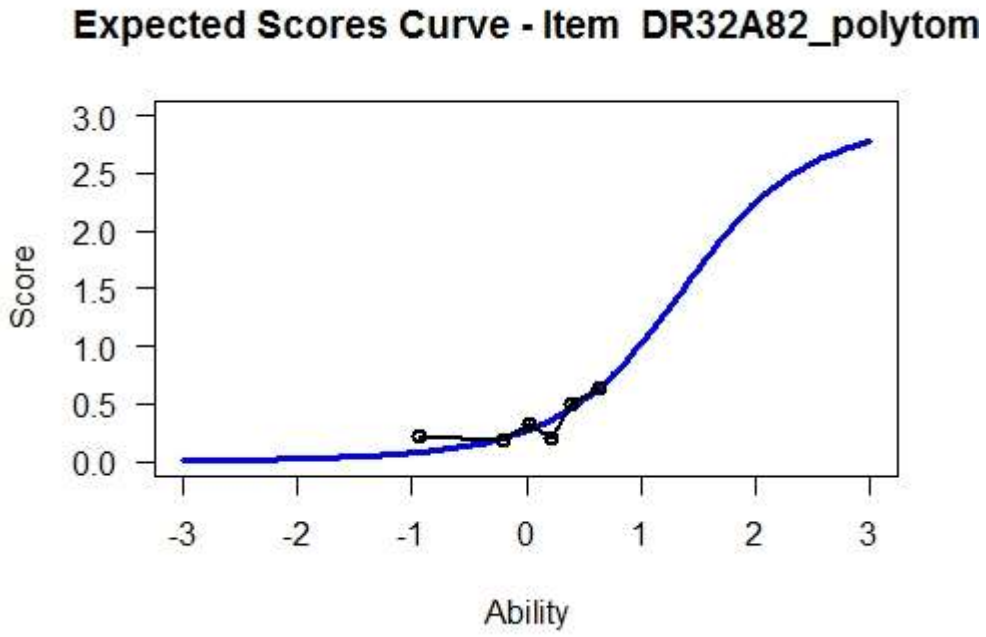


Abb. 17: Polytomes Itembeispiel A82 aus dem Bereich DR

Im zweiten Itembeispiel zeigt sich ebenfalls eine schmale Spanne bei einem etwas breiteren Score von 0.2 bis um 0.7 (Abb. 17). Die Itemschwierigkeiten in den einzelnen Kategorien liegen in Cat1[216] bei 1.7676, Cat2 bei 1.0668 und Cat3 bei 1.3603. Die beobachteten und erwarteten Lösungshäufigkeiten stimmen vom Kurvenverlauf her fast miteinander überein und zeigen nur eine geringe Abweichung, wie bereits im ersten Item-Fit-Beispiel erläutert. Ein bedeutender Over- oder Underfit ist nicht festzustellen. Eine valide Testinterpretation kann, identisch mit den Ergebnissen aus „TEMA“[217], stattfinden (vgl. Pospeschill 2010: 129f.; vgl. Döring et al. 2014: 160 u. 996ff.).

Die ermittelten Ergebnisse im Gesamtitempool zeigen überwiegend nur geringe Abweichungen auf.

Als abschließende Untersuchung wurden die residuenbasierten[218] Infit- und Outfitwerte, beruhend auf Chi-Quadrat-Verteilung, beurteilt. Die idealen Item-Fit-Werte sollten bei 1 liegen. Der Infit zeigt bestimmte Antwortmuster in den Kompetenztestungen an und liefert dadurch entsprechende statistische Informationen, die gewichtet werden. Er ist sensibel gerade bei untypischen Lösungsmustern, z.B. bei willkürlichen Antworten des TN. Beim Outfit werden die sogenannten „Ausreißerantworten“ mit erfasst. Bezogen auf die TN gelten die zu beantwortenden Items als besonders

[216] bedeutet engl. category 1, usw.

[217] „…, dass zur Berechnung der Fit-Statistiken jeweils nur vergleichsweise wenige empirische Werte auf einem engen Ausschnitt der latenten Fähigkeitsdimensionen herangezogen werden können.“ (Döring et al. 2014: 160).

[218] „Als Residuum bezeichnet man die Abweichungen von beobachteten und theoretisch erwarteten Werten.“ (Rost 2004: 371).

leicht oder schwer für einen bestimmten Personenkreis und zeigen sich bei Modellkonformität auch in den Testergebnissen (vgl. Hahnel 2018: 89; vgl. Eid u. Schmidt 2014: 199f.; vgl. Rost 2004: 371; vgl. Bensch 213: 121f.).

Die erzeugten Scatterplotts mit dem Package „TAM“ (Anhang 18) geben zunächst einen Gesamtüberblick aller ermittelten Item-Fit-Werte (Abb. 18)[219]. Anschließend werden in einer Ausschnittsvergrößerung die als passend geltenden zentralen Item-Fit-Werte im Gesamtergebnis[220] dargestellt (Abb. 19).

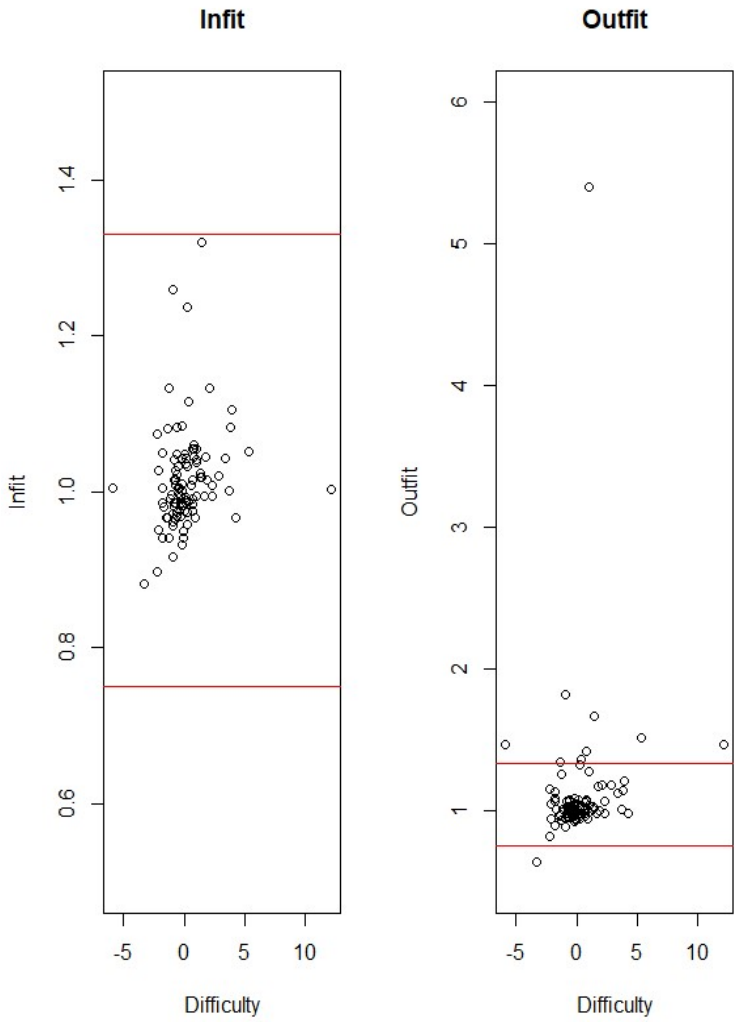


Abb. 18: Gesamtübersicht aller Item-Fit-Werte in der Gesundheits- und Krankenpflege

Die horizontalen roten Linien stellen die Eingrenzung der Normbereiche in den beiden Fitmaßen dar (Abb. 18). Im Abschlussbericht „TEMA“ wird auf einen Toleranzbereich im Infit wie im Outfit von 0.75 bis 1.33 hingewiesen (vgl. Döring et al. 2014: 160). Dieser wird auch in der vorliegenden Arbeit als vergleichbares Bewertungs-

[219] Die einzelnen numerischen Infit- und Outfitergebnisse mit p-Werten finden sich in Anhang 24.
[220] Die überwiegenden Werte häufen sich bei und um die 1, was für eine konkrete und valide Fähigkeitsmessung der latenten Variablen „praktische Pflegekompetenz“ gilt.

maß herangezogen. Als Hinweis sei an dieser Stelle noch erwähnt, dass von einigen Autoren die Ausweitung der Spanne auf Werte von 0.5 bis 1.5 teils diskutiert und als akzeptabel gilt (vgl. Bensch 2013: 122f.).

Der Range der Infitwerte bewegt sich zwischen 0.882 und 1.320 und der Outfitwerte zwischen 0.633 und 5.398.

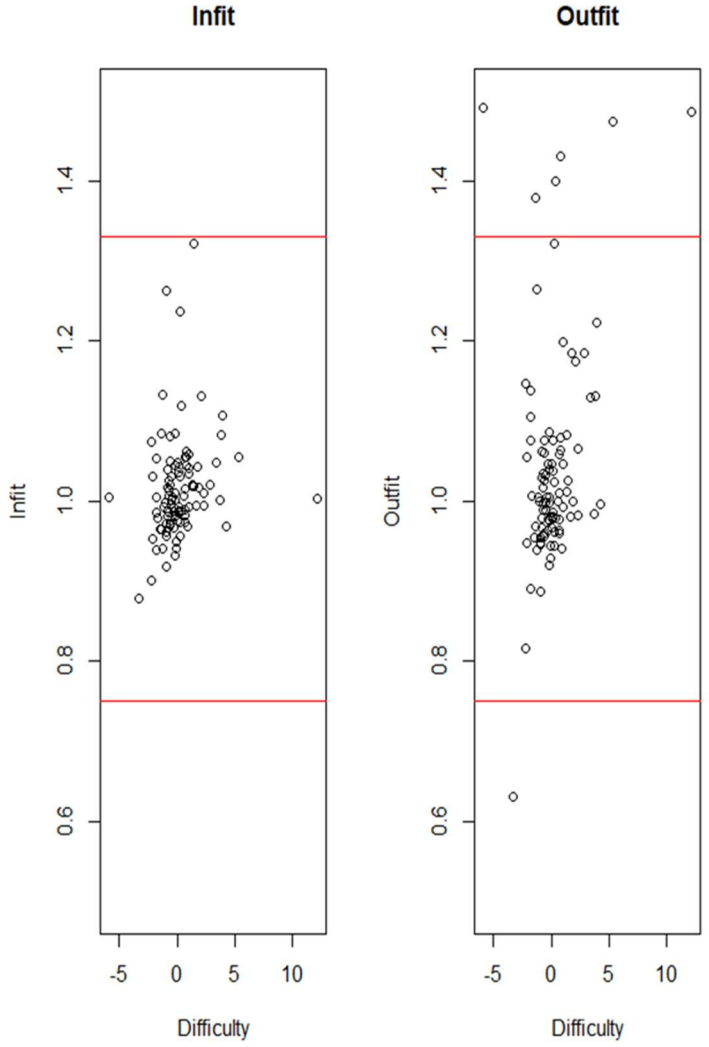


Abb. 19: Passende Infit- und Outfitwerte aus der Gesundheits- und Krankenpflege

Es lässt sich durch die ausgeprägte zentralistische Itemverortung um die 1 im Infit wie im Outfit auf eine überwiegend gute Itemmodellierung und Konformität in den erzielten Testergebnissen schließen. Es fallen aber auch atypische Lösungsmuster sowie vorhandene „Ausreißeritems“ als teilweise herausfordernde Items für die TN[221]auf (Anhang 24). Extreme „Abweicher“ sind in 10 Kategorien bzw. 7 Items im

[221] auch genannt Missfitwerte, d.h. es besteht eine große Devianz zwischen „beobachteten und erwarteten Werten“ (Hahnel 2018: 88). Ermittelte Item-Fit-Werte im Bereich > 1.33 sind zu korrigieren oder zu eliminieren. Ermittelte Item-Fit-Werte im unteren Bereich < 0.75 sollten nur dann aus dem Test entfernt werden, wenn das gesamte „Large-Scale-Assessment-Instrument“ reduziert werden soll (vgl. Bensch 2013: 122f).

Outfit zu erkennen (unter 0.75 und über 1.33). Ein Missfit liegt im Outfit sogar bei über 5 (Abb. 18). Es handelt sich um ein Item über die Erfassung des erhöhten Dekubitusrisikos. Im Infit zeigen sich dagegen alle Itemwerte nach der angeführten Daumenregel von Hahnel im empfohlenen Toleranzbereich (vgl. Hahnel 2018: 91).

Aber bei Berücksichtigung der p-Werte[222] zeigen sich im Outfit 18 betroffene Antwortkategorien und im Infit 12 betroffene Antwortkategorien, d.h. die Ergebnisse sind unter Annahme des Rasch-Modells unwahrscheinlich. Es handelt sich inhaltlich hauptsächlich um die Kompetenzen IK und DR aus den Themenkomplexen Beratung, Verhaltenseinschätzung und emotionaler Umgang. Für weitere Berechnungen zur Modellpassung der einzelnen Items wäre eine Elimination der einzelnen „Hotspots“ notwendig.

Die in „TEMA“ ermittelten Ergebnisse, einer vorhandenen hinreichenden Raschkonformität im Aufgabenpool, können nicht für das gesamte „Large-Scale-Assessment“ bestätigt werden.

Letztendlich zeigen die Testitems des technologiebasierten, computergestützten Kompetenztests zur Einschätzung von praktischer Pflegekompetenz, dass die Sonderform des Rasch-Modells bzw. des PCM eine mögliche pädagogische Unterstützung zur hochkomplexen Fähigkeitseinschätzung von Auszubildenden sowie Studierenden in der Gesundheits- und Krankenpflege sein kann.

Unter Berücksichtigung der Modellkonformität bei Betrachtung einer „spezifizierten“ Form des Raschmodells gilt die H_0-Hypothese in den durchgeführten Untersuchungen als nicht angenommen. Es liegt somit im Gesamttestscore keine „suffiziente“ Statistik vor, in der Modellpassung und Itempassung vorhanden sind und Leistungsunterschiede der TN erfasst werden. Die in Kapitel 1.2 gestellte statistische Fragestellung auf vorhandene Raschvalidität gilt somit als abgelehnt.

Auf den IRT-Ergebnissen aufbauend, wurden Dimensionsanalysen durchgeführt. Um das theoretisch entwickelte Kompetenzmodell, bestehend aus dem Faktor bzw. den Faktoren, auf seine Passung beurteilen zu können, wird unter Einsatz der empirisch gewonnenen TN-Daten ein multivariates Analyseverfahren, die Faktorenanalyse, gewählt. Zunächst wird der theoretische Hintergrund erläutert, danach werden die Umsetzung dargestellt und die ermittelten Ergebnisse bewertet.

[222] p-Werte sollen größer > oder gleich = 0,05 sein (vgl. Strobel 2015: 80).

16 Strukturgleichungsmodelle und Konfirmatorische Faktorenanalyse

Die Konfirmatorische Faktorenanalyse (CFA) ist ein multivariates Analyseverfahren[223] zur hypothesenprüfenden Funktion von theoretischen Modellen, auch als Strukturgleichungsmodell[224] bezeichnet. Es kann eingesetzt werden, wenn ein konstruiertes Modell auf seine Passung zu den erhobenen empirischen Daten geprüft oder ein konstruiertes Modell gegen andere, theoretisch entwickelte Modelle getestet werden soll. Dieses entwickelte Strukturgleichungsmodell im Rahmen der CFA wird zur Schätzung von Modellparametern und somit zur Überprüfung von Modellen eingesetzt. Das Ziel des Strukturmodells ist es, Zusammenhänge zwischen latenten Variablen abzubilden (vgl. Pospeschill 2010: 210ff.; vgl. Weiber, Mühlhaus 2014: 148f.).

Die einzelnen Faktoren werden vor der Analyse (a priori) der erhobenen TN-Daten bereits inhaltlich bestimmt. Eine Zuordnung der Variablen zu den Faktoren und deren Beziehungen zueinander und untereinander wird festgelegt. Die einzelnen beobachteten Variablen sollen den Faktor erklären. Bei der Auswertung wird das theoretische Modell (Modelle) mit dem empirischen Datensatz in Verbindung gebracht und auf seine Passung überprüft. Es wird anschließend entschieden, ob es angenommen oder abgelehnt wird. Darüber hinaus können auch die konkurrierenden Modelle mit dem Ursprungsmodell verglichen und überprüft werden. So wird die Entscheidung für das „beste“ theoretisch entwickelte Modell für den weiteren empirischen Gebrauch getroffen.

Beim klassischen Überprüfen im Rahmen der CFA wird die Anpassungsgüte (Fit-Wert) zwischen den erhobenen Teilnehmerdaten und dem theoretischen Modell geprüft. Es können aber auch weitere entwickelte, theoretische Alternativmodelle mit dem gleichen Datensatz überprüft werden. Die Unterscheidung innerhalb der konstruierten Modelle liegt in den unterschiedlichen theoretischen Annahmen.

Der **Ablauf einer CFA** verläuft in mehreren Phasen:

a. Modellspezifizierung
b. Datenbasisspezifizierung und Methode der Parameterschätzung und
c. Auswertung des vorliegenden Modells (vgl. Pospeschill 2010: 211).

Der **Phasenablauf** und weitere faktorielle Hintergründe sollen jetzt genauer aufgezeigt werden. Im ersten Schritt geht es um die **Modellspezifikation**. Alle Hypothesen werden als Gleichungen formuliert und in einem Pfaddiagramm dargestellt. Im Pfaddiagramm wird die Faktorenzahl bestimmt und die Zuordnung der Variablen zu den Faktoren festgelegt. Es werden speziell fixierte, einheitliche Symbole im Pfaddiagramm verwendet: Kreise für die latenten Variablen bzw. Faktoren, Quadrate für die beobachteten Items bzw. manifesten Variablen und verkleinerte Kreise oder auch ohne symbolische Erkennung als Fehlervariablen (Error-Variablen). Die Verbindungen zu den einzelnen Symbolen werden zum einen durch gezielte Pfeile bei

[223] hierbei werden gleichzeitig mehrere Variablen auf deren Abhängigkeiten oder Zusammenhänge analysiert (vgl. Backhaus et al. 2016: 589ff.).
[224] auch „structural equation model“ genannt

gerichteten Beziehungen[225] erzeugt, ungerichtete Beziehungen[226] durch geschwungene Doppelpfeile (vgl. Moosbrugger & Schermelleh-Engel 2012: 335ff.; vgl. Werner et al. 2016: 949).

Die Messfehler jeder Variablen sind unabhängig, es ist keine Korrelation zwischen den einzelnen Variablen vorhanden.

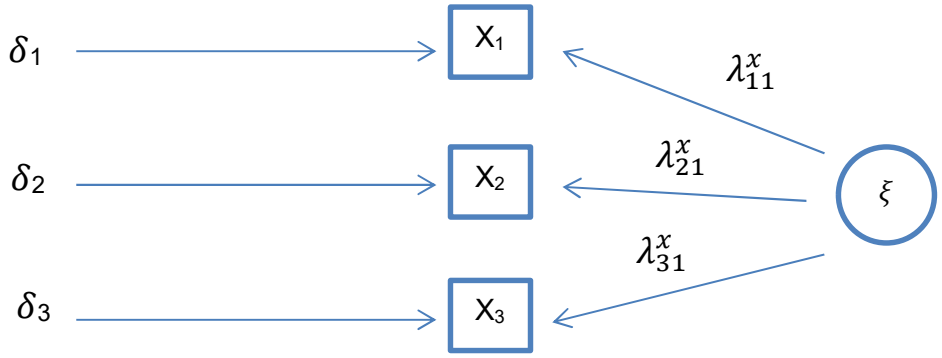


Abb. 20: Beispiel eines einfachen Messmodells (in Anlehnung an: Werner et al. 2016: 949; eigene Darstellung Ries)

Zur Erläuterung der oberen Abbildung soll an dieser Stelle erklärt sein, dass Xi (ξ) für den Faktor oder die latente Variable steht, x ist die jeweilige beobachtete Variable (= Indikatorvariable), und Delta (δ) steht für den Messfehler (Abb. 20).

Besteht eine Gemeinsamkeit zwischen den beobachteten Variablen, die sich auf das gleiche Konstrukt bzw. den vorhandenen Faktor beziehen, kann die Wirkung auf diesen zurückgeführt werden. Zu erkennen ist dies an einem hohen ξ. Die beobachteten Variablen beziehen sich somit ganzheitlich auf das entwickelte Konstrukt. Eine Kovarianz existiert nur deshalb bei den beobachteten Variablen, weil sie den Faktor erklären. Gleichzeitig bleibt bei jedem x ein δ vorhanden, das die lokale stochastische Unabhängigkeit aller beobachteten Variablen wiedergibt. Im Strukturgleichungsmodell bedeutet das, dass die beobachteten Variablen die Bezeichnung von manifesten Indikatorvariablen bekommen und den Faktor bzw. die latente Variable messen können. Ein passendes Messmodell ist entstanden. Sind mehrere Konstrukte bzw. Faktoren vorhanden, wird das im Messmodell angezeigt, wie im TEMA-Kompetenztest (siehe Kapitel 16.1/Abbildung). Besteht ein möglicher Zusammenhang der einzelnen Konstrukte innerhalb des Messmodells, kann dies durch geschwungene Doppelpfeile festgelegt werden (vgl. Werner et al. 2016: 950f.; vgl. Kaspar et al. 2016: 16).

Hier kann festgehalten werden, dass die CFA, wie auch die klassische Testtheorie, aus wahrem Wert und einem Fehlerwert zu verstehen ist (vgl. Döring, Bortz 2016b: 461).

[225] Jeder gezielte Pfeil bezeichnet ein partielles Regressionsgewicht.
[226] Diese geben Kovarianzen und Korrelationen an.

Die verschiedenen Messmodelle können alternativ auch als lineare Gleichungen aufgeführt werden. Bezogen auf das oben aufgeführte einfache Messmodell in Abb. 20 bedeutet dies (Gl. 15):

$$x_1 = \lambda_{11}^{x} \cdot \xi_1 + \delta_1$$

$$x_2 = \lambda_{21}^{x} \cdot \xi_1 + \delta_2$$

$$x_3 = \lambda_{31}^{x} \cdot \xi_1 + \delta_3$$

Die Faktorladungen λ (lambda-x) stehen dafür, wie eine manifeste Indikatorvariable x durch eine latente Variable bzw. ein Konstrukt ξ erklärt wird. Die angeführten Ladungen λ beinhalten zwei Indexwerte. Wert eins beschreibt die Variable, auf die der Effekt wirkt[227], und Wert zwei den ausgehenden Effekt[228] Xi. Als weitere Indikatoren gibt es die Angabe von Messfehlervarianzen θ^{δ} (theta-delta) (Abb. 21). Diese ergeben sich ausschließlich aus den Varianzen der δ-Variablen. Sie ermitteln sich nicht aus dem Faktor. Des Weiteren steht unter dem griechischen Buchstaben Phi (φ) die Angabe von internen Varianzen innerhalb eines Faktors bzw. einer latenten Variablen und Kovarianzen aus den Verbindungen zwischen den gesamten vorhandenen Faktorvariablen (vgl. Bortz, Schuster 2010: 385ff.; vgl. Moosbrugger, Kelava 2012: 336f.; vgl. Werner et al. 2016: 951).

Bisher lagen die Erläuterungen ausschließlich auf den latenten Faktorvariablen im Rahmen eines produzierten Messmodells. Bei der Entwicklung von komplexeren faktoriellen Strukturgleichungsmodellen wird zusätzlich zum entwickelten Messmodell noch ein Strukturmodell mit angeschlossen und in das gesamte Pfaddiagramm integriert. In diesem werden die Beziehungen zwischen den latenten Variablen bzw. Faktoren vertiefend konkretisiert. Die latente exogene Variable, Faktor oder auch Prädiktorvariable ξ (xi) genannt, bedingt eine Wirkung auf eine vorhandene latente, endogene Variable η (eta). Die vorhandene latente, endogene Variable kann als Kriteriumsvariable bei direkten Effekten in Erscheinung treten. Im Pfaddiagramm ist dies mit einem direkten geraden Pfeil von Prädiktor- zu Kriteriumsvariable gekennzeichnet. Man spricht im Zusammenhang mit dem entwickelten Pfaddiagramm von Prädiktor-Kriteriums-Beziehung. Ein indirekter Effekt wird angezeigt, wenn über eine dritte Variable der Effekt erzielt wird. Als Sonderform eines indirekt bewirkten Effektes können Prädiktor- und Kriteriumsvariablen auch zusammen als ein gemeinsamer Faktor auftreten. In den zuvor erläuterten Fällen wird von „Mediatorvariablen“ gesprochen (vgl. Werner et al. 2016: 952f.; vgl. Hussy, Schreier, Echterhoff 2013:156).

[227] Pfeilspitze zeigt auf die Variable.
[228] Pfeil führt vom Konstrukt weg.

Im Strukturmodell werden die gesamten Verbindungen der latenten Variablen vereint. Bei den endogenen Variablen werden die Messfehlervariablen mit ε (epsilon) angegeben. Das gibt auch hier wieder den unerklärbaren Anteil des jeweiligen Items (in diesem Fall dann y) an. Kommt es zu Effektparametern von ξ zu η wird dies mit γ (gamma) gekennzeichnet und bei nur intern endogenen Effekten mit β (beta) angegeben. Der nicht erklärbare Varianzanteil im Gesamtmodell kann über den Anteil innerhalb der einzelnen latenten, endogenen Variablen eingeschätzt werden. Sie wird als Modellgleichungsfehlervariable[229] mit dem griechischen Buchstaben ζ (zeta) bezeichnet. Die Varianzen werden eingeschätzt und bei der jeweiligen endogenen Variablen mit ψ (psi) angegeben.

Die rechnerische Aufbereitung kann alternativ zum Pfaddiagramm entweder in Einzelgleichungen (Gl. 15) oder in Matrixschreibweise (Abb. 21) erfolgen. Hier werden die einzelnen Gleichungen einschließlich der jeweiligen Verbindungen untereinander als ein jeweiliges Konstrukt aus den exogenen und eines aus den endogenen Variablen zusammengeführt. Somit wird durch die Matrixschreibweise eine verbesserte Übersichtlichkeit erreicht (vgl. Werner et al. 2016: 952f.; vgl. Bühner 2011: 390).

Die angefügte Abbildung (Abb. 21) soll die Zusammenhänge innerhalb des komplexen Strukturmessmodells bildhaft verdeutlichen.

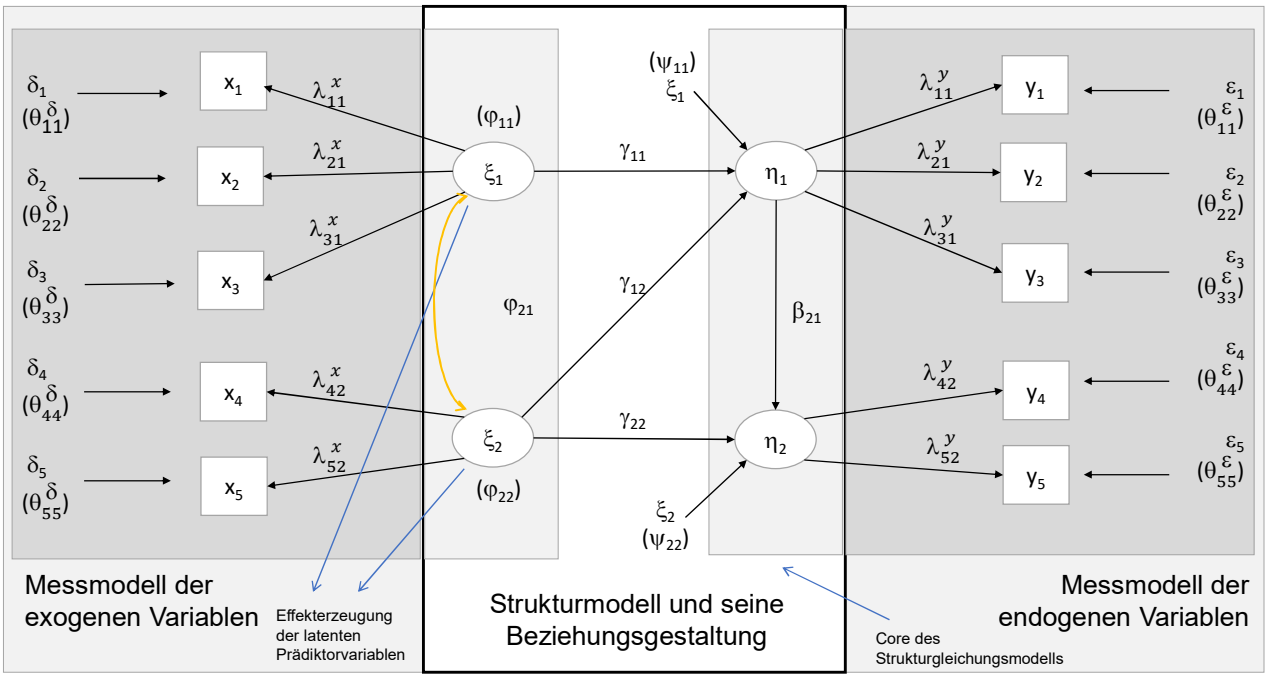


Abb. 21: Beispiel eines Strukturgleichungsmodells (entnommen aus: Werner et. al. 2016: 953 leicht modifiziert von Ries 2019)

[229] oder auch Residualvariable

Die Zusammenhänge zwischen der Beziehung von beobachteten Variablen bzw. Indikatorvariablen und den Faktoren haben stattgefunden. Die Ladungszusammenhänge sind definiert. Eine Modellspezifikation hat stattgefunden. Im nächsten Schritt soll die **Datenbasisspezifizierung und Methode der Parameterschätzung** bestimmt werden. Die Parameter des entwickelten Modells werden durch Bestimmung von Varianzen, Kovarianzen und Erwartungswerten genau ermittelt. Als Grundlage für die Berechnungen dienen die beobachteten Variablen. Diese müssen nachgewiesen werden, man spricht auch von der Identifikation des Problems. Die Anwendung gilt für jedes einzelne entwickelte Modell. Nach der Identifizierung kann das Modell auf seine Gültigkeit überprüft werden. Dabei werden die Abweichungen der Erwartungswerte und der vom Modell implizierten Varianzen und Kovarianzen der beobachtbaren Variablen von den empirischen geschätzten Größen minimiert. Dies geschieht in der Regel mit einem χ^2-Test. Die Modellpassung wird damit überprüft (vgl. Eid, Schmidt 2014: 352).

Soll eine Konfirmatorische Faktorenanalyse durchgeführt werden, müssen die manifesten Indikatorvariablen x in ihrem Urzustand vollständig vorliegen. Aus den daraus folgenden Kovarianzen wird eine übersichtliche empirische Kovarianzmatrix S erstellt (Zusammenhangsmaße von x). Wiederum aus den Verbindungen von x, unter Betrachtung der Parameterwerte z.B. λ, entsteht eine modelltheoretische Kovarianzmatrix $\hat{\Sigma}$ (Sigma-Dach). Die empirische Kovarianzmatrix S wird berechnet für die anschließenden Parameterschätzungen. Kovarianzmatrix S und Kovarianzmatrix $\hat{\Sigma}$ sollen nur geringe Abweichungen aufweisen. Die Modellidentifikation muss stattgefunden haben, damit die erwähnten Schätzungen der Modellparameter überhaupt durchgeführt werden können. Die Voraussetzung dafür ist aber immer, dass genügend empirische Informationen vorhanden sind, um die Parameter des Modells ermitteln zu können. Eine bestimmte vorhandene Metrik für die latenten Variablen muss nachweisbar sein (vgl. Moosbrugger, Kelava 2012:336f; vgl. Pospeschill 2010: 215; vgl. Werner et al. 2016: 955; vgl. Bühner 2011: 394ff.).

Die Bedingung der Identifikation ist erfolgt,

> „…wenn die Anzahl s der empirischen Varianzen und Kovarianzen der manifesten Indikatorvariablen größer ist als die Anzahl t der zu schätzenden Parameter…“ (Pospeschill 2010: 214).

Das Modell zeigt eine positive Anzahl von Freiheitsgraden (df). Als Gleichung (Gl. 16) dargestellt:

$$df = s - t > 0$$

Die Folge ist bei Erfüllung der genannten Voraussetzung, dass die Schätzung für den jeweils zu ermittelnden Parameter eindeutig ist (vgl. Moosbrugger, Kelava 2012: 336).

Bei Unteridentifizierung (*df* = -1) oder gerade Identifizierung (*df* = 0) eines Modells ist keine Modelltestung möglich (Gl. 16). Deshalb sollten pro Faktor mindestens drei manifeste Variablen vorhanden sein, um eine Modellüberprüfung durchführen zu können (vgl. Pospeschill 2010: 215; vgl. Werner et al. 2016: 964).

Die Varianzen der manifesten Indikatorvariablen x werden direkt berechnet, dagegen müssen die Varianzen der Faktoren bestimmt werden. Die Festlegung kann auf zwei Arten erfolgen. Zum einen kann der Faktor auf die Zahl 1 standardisiert werden oder als Alternative pro Faktor die Faktorladung einer Indikatorvariablen auf 1 fixiert werden. So ist die Varianz der latenten Variablen gleich der erklärten Varianz dieser Indikatorvariablen und übernimmt deren Skalierung (vgl. Werner et al. 2016: 955).

Die Schätzung von Parametern wird mit unterschiedlichen Methoden durchgeführt. Der gewünschte Effekt ist, dass die H_0-Hypothese (= Das Modell und die vorhandenen Daten passen zusammen.) erfüllt ist. Überwiegend wird dazu die Maximum-Likelihood-Methode (MLM)[230] verwendet. Als Voraussetzungen gelten hier Intervallskalierung und multivariate Normalverteilung. Die Anwendung der Schätzmethoden verfolgt das Ziel, dass Parameter des theoretischen Modells so bestimmt werden, dass sie mit den empirischen Varianzen und Kovarianzen sich möglichst gut darstellen und abbilden lassen. Die vom Modell implizierte, theoretische Kovarianzmatrix $\hat{\Sigma}$ unter Verwendung der geschätzten Parameter muss mit der empirischen Kovarianzmatrix S möglichst gut übereinstimmen, was mit verschiedenen Fit-Maßen ermittelt wird. Beispielsweise wird bei der MLM-Schätzung der ermittelte Likelihood-Wert als Maximalwert angegeben, um die empirische Kovarianzmatrix S zu erhalten. Der ermittelte Wert zeigt, selbst bei teils nicht vorhandenen Werten, eine hohe Güte und Genauigkeit (vgl. Moosbrugger & Schermelleh-Engel 2012: 336f; vgl. Werner et al. 2016: 956f.).

Als weitere Möglichkeit stehen Kleinstquadrate-Schätzverfahren zur Verfügung, wie beispielsweise die ULS-Methode[231]. Diesen wird im Rahmen dieser Arbeit nicht weiter nachgegangen, da sie für die weitere Auswertung keinen Einfluss haben und auch in den gesichteten Abstracts von „TEMA“ nicht verwendet wurden.

Als Ergänzung ist noch zu erwähnen, dass bei der Parameterfestlegung zwischen fixen[232] und freien[233] Parametern unterschieden wird (vgl. Pospeschill 2010: 215).

Der abschließende Schritt der CFA ist, dass eine **Modellevaluation** durchgeführt wird. Sie soll mit der Modellpassung aus dem „TEMA“-Projekt und den erhobenen TN-Daten aus der Gesundheits- und Krankenpflege erfolgen. Die praktische Ausführung ist in Kapitel 16.2 zu finden. Im Mittelpunkt der Evaluation steht hier die Betrachtung der Anpassungsgüte des Modells.

Die Güte innerhalb des Strukturgleichungsmodells zeigt sich des Weiteren über die Residuen. Sie bezeichnen die Abweichungen von Kovarianzmatrix S zur theoretischen Kovarianzmatrix $\hat{\Sigma}$. Je kürzer die Differenzen, desto genauer erklärt es das entwickelte Modell im Zusammenhang mit den verwendeten Daten. Statistisch ausgedrückt: S - $\hat{\Sigma}$ ergibt die einzelnen Residuen. Die ermittelten Residualwerte folgen

[230] Messung erfolgt überwiegend genau und ist für große Stichproben (bei ca.200-250 TN) geeignet.
[231] „Unweighted-Least-Squares“-Methode
[232] Hier wird eine Konstante „a priori“ als Parameter festgelegt.
[233] Hier wird kein Wert vor der Modellschätzung als Parameter festgelegt. Durch einen Anpassungsprozess (= Iteration) nach der Schätzung des Modells wird eine Annäherung zwischen beiden Kovarianzmatrixen (S und $\hat{\Sigma}$) versucht.

nach Werner et al. aus dem Einfluss von Stichprobenfehler und Modellfehler. Eine Bewertung der Verbindung zwischen den Variablen innerhalb der empirisch erhobenen Daten wird möglich (= Modell-Fits). Häufig werden auch standardisierte Residuen zu Hilfe genommen (vgl. Moosbrugger, Kelava 2012: 337; vgl. Werner et al. 2016: 957).

Eine gesamtheitliche, inferenzielle statistische Betrachtung vom Modell und den erhobenen Daten kann über den χ^2-Test erfolgen. Der erhaltene χ^2-Wert ist stichprobenabhängig und reagiert sehr sensibel, gerade bei größerem N. Das gewünschte Ergebnis ist hier: Die H_0-Hypothese nicht zu verwerfen. Wie auch bei den Residuen soll der globale χ^2-Wert im Ergebnis möglichst klein sein. Ist dies der Fall, können darauf aufbauend verschiedene deskriptive Gütemaße identifiziert werden, um Modell und erhobene Daten auf die weitere Passung zu überprüfen. Unter Berücksichtigung der empirisch erhobenen Daten können Parameter eingeschätzt und bewertet werden (vgl. Werner et al. 2016: 958; vgl. Moosbrugger & Schermelleh-Engel 2012: 337; vgl. Pospeschill 2010: 217f.).

Bezogen auf die Modellpassung wird bei Pospeschill und Moosbrugger & Schermelleh-Engel im Rahmen von deskriptiven Beurteilungsmaßen, wie beispielsweise dem RMSEA[234] oder dem SRMR[235], von Fit-Indizes oder deskriptiven Gütemaßen gesprochen. Werner et al. unterscheidet noch zwischen absoluten und relativen deskriptiven Gütekriterien. Bei den zu ermittelnden absoluten Gütemaßen werden die empirische Matrix S mit der modelltheoretischen Matrix $\hat{\Sigma}$ auf deren Übereinstimmung untersucht. Als absolute Maße gelten der RMSEA und der SRMR. Auf die beiden Maße soll hier noch kurz eingegangen werden (vgl. Pospeschill 2010: 220; vgl. Moosbrugger & Schermelleh-Engel 2012: 337; vgl. Werner et al. 2016: 967f).

Der **RMSEA** wird auch als Badness-of-Fit-Index, „Fehlermaß“ oder „Näherungswert“[236] der Passung eines Modells bezeichnet. Die zu beantwortende statistische Fragestellung lautet in diesem Falle: Wie stark weicht das erstellte Modell von den erhobenen Daten ab? Wie „schlecht“ beschreibt es die Daten? Als Ergebnis wird ein möglichst kleiner RMSEA-Index (Nahe 0 bzw. zwischen ≤ 0,05 und ≤0,06) unter Beachtung der Freiheitsgrade[237] und der Größe der Stichprobe[238] gewünscht. Bezogen auf die in dieser Arbeit erhobenen TN-Daten in der Gesundheits- und Krankenpflege liegt die erhobene Stichprobengröße bei einem N von 389. Dies würde nach Pospeschill bei einem N von > 250 mit einem RMSEA-Index von ≤ 0,06 von einer guten Passung der Daten zum Modell sprechen lassen (vgl. Pospeschill 2010: 220; vgl. Schröder 2010: 59; vgl. Werner et al. 2016: 967; vgl. Georg-Elias-Müller-Institut für Psychologie 2018; vgl. Bühner 2011: 424f.).

[234] Root Mean Square Error of Approximation
[235] Standardized Root Mean Square Residual
[236] in der Statistik auch als Approximation bezeichnet -> Näherung zwischen Modell und Daten (vgl. Werner et al. 2016: 967).
[237] Abkürzung: *df*
[238] Abkürzung: in RMSEA-Formel als n oder kann auch als N dargestellt werden (vgl. Backhaus et al. 2015: 148f.; vgl. Franken 2010: 107).

Die Formel zur Berechnung des RMSEA-Index lautet (Gl. 17):

$$RMSEA = \sqrt{\frac{\chi^2 - df}{n \cdot df}}$$

Ergänzend kann über die Genauigkeit der RMSEA-Schätzung eine Aussage mit der Bestimmung des Konfidenzintervalls getroffen werden. Gerade im Bereich von persönlichkeitsmerkmalsorientierten Forschungsvorhaben sollte der RMSEA-Wert seine Berücksichtigung finden (vgl. Bühner 2011: 425f.).

Im vorliegenden Kompetenztest in der Gesundheits- und Krankenpflege bestimmen die persönlichen Wahrnehmungs- und Erfahrungsmuster unter Einfluss der empathiegeleiteten Persönlichkeitsstruktur die Entscheidung für emotionale und kommunikationsorientierte Antwortentscheidungen. Deshalb sollte der RMSEA-Wert in der vorliegenden Forschungsevaluation errechnet, betrachtet und berücksichtigt werden.

Als Ergänzung zum RMSEA besteht die Möglichkeit, den **SRMR**[239]-Wert zu bestimmen. Der ermittelte Wert sollte sich auch hier der Zahl 0 (≤ 0,10) nähern. Dieser beschreibt das durchschnittliche Residuum innerhalb von standardisierten Variablen bzw. deren Residualmatrix. Die Komplexität des Modells wird im Gegensatz zum RMSEA beim SRMR bei der Ermittlung nicht berücksichtigt. Je größer die Stichprobe, desto mehr nähert sich der ermittelte SRMR-Wert 0 an, es liegt eine Abhängigkeit zur Stichprobe vor (vgl. Werner 2016: 967; vgl. Backhaus et al. 2015: 150; vgl. Bühner 2011: 424ff.).

Der SRMR-Wert wird mit der folgenden Formel (entnommen und Formelerläuterung ergänzt aus Bühner 2011: 426) bestimmt (Gl. 18):

$$SRMR = \sqrt{\sum_{i} \sum_{i<k} \frac{r_{ik}^2}{e}}$$

i = Anzahl der Items i
k = Anzahl der Items k
r_{ik} = Residualkorrelation zwischen den Items *i* und *k*
e = Anzahl der bekannten Größen aus der Varianz-/Kovarianzmatrix

Durch den erhaltenen SRMR-Wert wird ein weiteres Fit-Indiz für die Güte eines Modells bestimmt (vgl. Bühner 2011: 426f.).

[239] Standard-Root-Mean-Residual

Als mögliche relative, deskriptive Maße können der **CFI**[240] und der **NFI**[241] bestimmt werden. Sie werden auch als komparative[242] Fit-Indizes bezeichnet. Innerhalb der gesamten komparativen Fit-Indizes liegt ein Stufensystem von 1 bis 3 (aufsteigend) vor. Bei der Stufe 3 werden die meisten Parameter bzw. Informationen in das zu ermittelnde Fit-Indiz mit einbezogen. Dieses zeigt eine hohe Validität in Abhängigkeit des ermittelten bzw. getesteten Modells an. Deshalb wird der CFI-Wert (gehört der Stufe 3 an) ausführlicher dargestellt und im Kapitel 16.2 errechnet und anschließend evaluiert.

Auf den NFI-Wert (gehört in die Stufe 1) wird hier nicht weiter eingegangen, er wird aber bei Pospeschill und Bühner als mögliches deskriptives Indiz oder auch Gütekriterium erwähnt. Ein guter Modellfitwert sollte beim NFI bei ≥ 0,95 liegen.

Bei der Ermittlung des CFI-Wertes wird das untersuchte Modell mit dem Unabhängigkeitsmodell[243] bzw. dem Nullmodell verglichen. Die gesamten beobachteten Variablen sind im Unabhängigkeitsmodell als nicht korreliert anzusehen. Der Wert 1 wird von dem ermittelten Wert vom Ergebnis aus Zähler und Nenner subtrahiert. Je unterschiedlicher die beiden Modelle, desto stärker nähert sich der CFI-Wert dem Maximalwert 1.

Die einzusetzende Formel für die Ermittlung des CFI-Wertes nach Pospeschill 2010 lautet (Gl. 19):

$$CFI = 1 - \frac{\chi_P^2 - df_P}{\chi_N^2 - df_N}$$

Pospeschill und Moosbrugger & Schermelleh-Engel sprechen von einem guten Modellfitwert des CFI bei ≥ 0,97 (vgl. Bühner 2011: 427; vgl. Moosbrugger & Schermelleh-Engel 2012: 337f.; vgl. Werner et al. 2016: 968; vgl. Pospeschill 2010: 221).

Zusammenfassend lässt sich bezüglich der Modellevaluation sagen, dass für eine umfassende Beurteilung der Anpassungsgüte des entwickelten konfirmatorischen Faktorenmodells der $\chi 2$-Wert einschließlich p-Wert, CFI, RMSEA und SRMR ermittelt werden sollte.

Im nächsten Schritt werden zunächst im Zusammenhang auf das Ursprungsprojekt „TEMA“ die Hintergründe zur Berechnung der CFA erläutert. Danach erfolgen statistische Berechnungen mit den erhobenen TN-Daten aus den besuchten Hochschulen und Gesundheits- und Krankenpflegeschulen bzw. Bildungszentren für Gesundheit und Pflege einschließlich Interpretation der gewonnenen Ergebnisse.

[240] Comparative-Fit-Index
[241] Normal-Fit-Index nach Pospeschill (vgl. Pospeschill 2010: 220) oder Normed-Fit-Index nach Bühner (vgl. Bühner 2011: 424)
[242] vergleichend/Vergleich zwischen dem geprüften Modell und dem Unabhängigkeitsmodell
[243] auch Independence Model bezeichnet

16.1 Dimensionalität der Leistungsdaten

Im „TEMA“-Kompetenztest wurden drei verschiedene theoretische Modelle mit der konfirmatorischen Faktorenanalyse getestet, miteinander auf deren Modellpassung verglichen und deren priorisierende Modelldimensionalität bestimmt. Die Überprüfung von vermuteten Zusammenhängen zwischen Variablen wurde erreicht und nachgewiesen. Bevor dies geschehen konnte, wurden 9 Item-Pakete[244] konstruiert, die in den jeweiligen Modellen abgebildet wurden (vgl. Döring et al. 2016: 118ff.). Die entwickelten Modelle waren ein Generalfaktormodell - „Client-Directed Care Model“[245], ein theoretisch entwickeltes dreidimensionales Subdomänenkompetenzmodell - „Subdomain Model“[246] sowie ein weiteres dreidimensionales pflegerisches Handlungsfeld- bzw. Settingmodell - „Testlet Model“[247]. Das „Subdomain Model“ ist das ursprünglich entwickelte heuristische Modell, wie zu Beginn der Arbeit im theoretischen Teil (siehe Kapitel 4.2, Abb. 1) erwähnt. Das Generalfaktormodell, mit nur einem vorhandenen Faktor, wurde gegen die anderen beiden Modelle getestet (vgl. Wittmann, Kaspar, Döring 2017: 197; vgl. Kaspar et al. 2016: 15f.).

Die aufgeführte Abbildung (Abb. 22) soll die einzelnen alternativen Faktorenmodelle aus dem TEMA-Projekt nochmals verdeutlichen.

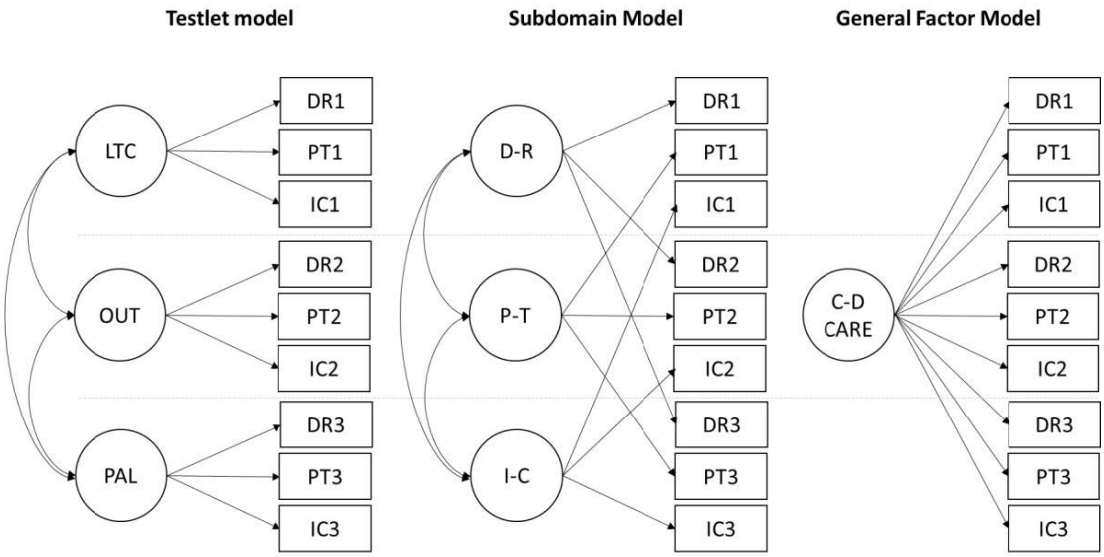


Abb. 22: Entwickelte Faktorenmodelle im Rahmen des „TEMA“-Kompetenzprojektes (entnommen aus: Kaspar et al. 2016: 16)

[244] Die 9 Item-Pakete wurden aus dem gesamten entwickelten Test-Item-Pool zusammengestellt. Die „Schnürung“ der Testpakete richtete sich nach dem heuristischen Kompetenzmodell aus dem Teilkompetenzbereich der unmittelbar bewohner-/klientenorientierten Kompetenzen: Diagnostisch-Reflexive Kompetenz (= DR1, DR2, DR3), Praktisch-Technische Kompetenz (= PT1, PT2, PT3) und Interaktiv-Kommunikative Kompetenz (= IK1, IK2, IK3) (vgl. Kapitel 4.2).

[245] oder auch „C-D Care Model“ genannt

[246] bestehend aus den entwickelten Kompetenzbereichen: Diagnose + Reflexion (DR), Praxis + Technik (PT), Interaktion + Kommunikation (IK); jede Teilkompetenz stellt sich als separate Abbildung dar.

[247] Die drei beruflichen Handlungsfelder bzw. Settings beinhalten Langzeitpflege/long term care (LTC), Ambulante Pflege/outpatient care (OUT) und Palliative Pflege/palliative care (PAL).

Im Ergebnis zeigte sich beim „TEMA“-Kompetenztest, dass sich das eindimensionale Modell als angebracht gegenüber den beiden anderen dreidimensionalen Modellen ergab (vgl. Wittmann, Kaspar, Döring 2017: 197 u. 198; vgl. Kaspar et al. 2016: 16).

Auf dieser zuvor erläuterten Grundlage (Abb. 22) werden durch Verwendung der gewonnenen TN-Daten aus der Gesundheits- und Krankenpflege eine Überprüfung der Modellpassung durchgeführt und die entwickelten Strukturgleichungsmodelle aus dem Ursprungsprojekt „TEMA“ gegeneinander verglichen. Die Ergebnisse werden im Wesentlichen mit dem Package „lavaan“ der Statistiksoftware R erzeugt.

Bei der Durchführung einer CFA wird im ersten Schritt eine Modellspezifizierung, wie in Kapitel 16 beschrieben, durchgeführt. Ein weiteres Strukturgleichungsmodell wurde nicht entwickelt, und bestehende Modelle wurden auch nicht verändert. Es bestätigten sich nach Rekonstruktion die identischen, bereits verwendeten Modellpassungen aus dem „TEMA“-Projekt für die weitere Überprüfung von Zusammenhängen der beobachteten Variablen (vgl. „TEMA“-Projektgruppe 2014; vgl. Kaspar et al. 2016: 16). Bei den mehrdimensionalen Modellen wurde von der Annahme der gegenseitigen Verbindung der jeweils drei Faktoren zueinander (gebogener Pfeil) ausgegangen und deshalb mit dem Wert 1 fixiert, da die praktische Pflegekompetenz zum Ende der dreijährigen Ausbildung in ihrer Gesamtkomplexität gemessen und erfasst wird (vgl. Pospeschill 2010: 213).

16.2 Ergebnisse in R innerhalb der Datenanalyse

Es werden die gleichen, a priori gewählten, 62 Items aus dem Ursprungsprojekt „TEMA“ auf ihre Modellpassung untersucht. Es zeigt sich in den vielfältigen literarischen Veröffentlichungen, dass unterschiedliche Angaben zu den verwendeten Testitems benannt werden, was im Rahmen der Gesamtbetrachtung der CFA zu leichter Verwirrung bei der eigenen Aufbereitung führt. Die Forscherin hat sich an die Itemanzahl aus einem unveröffentlichten Statistikskript der Projektgruppe „TEMA“ gehalten (vgl. Wittmann, Kaspar, Döring 2017: 197; vgl. Kaspar 2018; vgl. Döring et al. 2016a: 255).

Die anfänglichen Versuche, mit dem Statistikprogramm AMOS zu arbeiten, zeigen sich als nicht praktikabel, teils sogar umständlich für die Erstellung von Strukturmodellen, ebenso für die Ergebnissicherung zur Modellpassung bzw. zur Dimensionalitätserfassung. Deshalb wurde die Alternative mit „lavaan“ bevorzugt. So zeigte sich der Vorteil, dass nach einer aufwendigen und statistisch begleitenden Skripterstellung die gesamten gewünschten, komplexen Daten in der empirischen Breite zur Verfügung standen (Anhang 25). Als Grundlage dafür wurde das R Skript „Burnout“ von Döring u. Bortz herangezogen (vgl. Döring, Bortz 2016: 960ff.).

Die folgende Tabelle zeigt den essenziellen Teil der ermittelten Ergebnisse bei Betrachtung jedes einzelnen Strukturmodells in der Übersicht auf.

Tab. 10: Übersicht der drei verglichenen Messmodelle mit den zentralen entnommenen Ergebnissen aus R „lavaan“

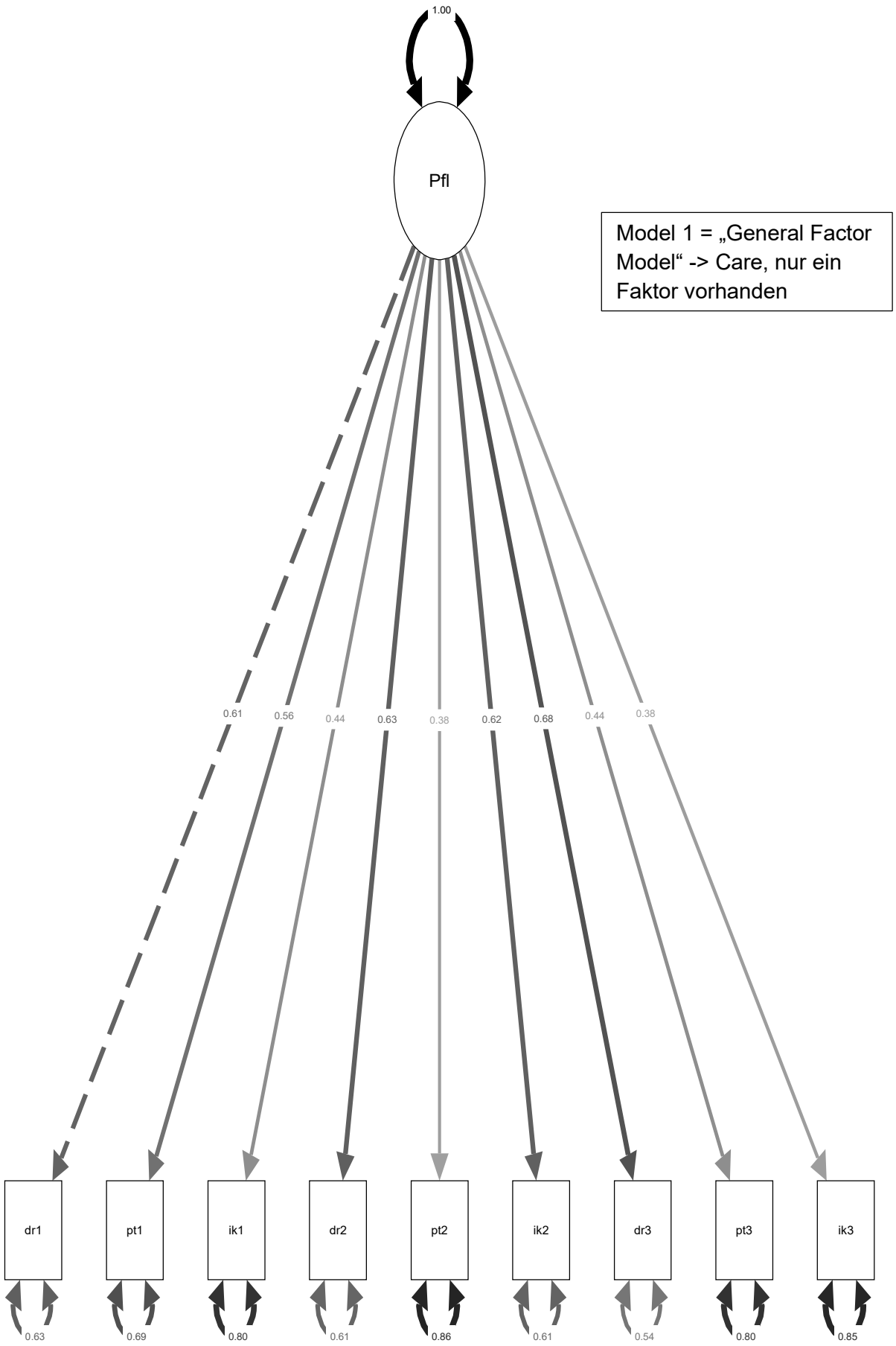
	Pflege - 1D	Subdomain - 3D „Heuristisches Modell“	Testlet - 3D „Setting Modell“
Number of free parameters	18	21	21
LogLikelihood (LL) user (H_0)	176.535	177.242	180.131
Model Fit Test Statistik	χ^2=32.152, df=27, p=0.226	χ^2=30.738, df=24, p=0.161	χ^2=24.961, df=24, p=0.408
CFI	0.992	0.989	0.998
NFI	0.951	0.953	0.962
RMSEA	0.022	0.027	0.010
SRMR	0.031	0.031	0.028
AIC	-317.070	-312.485	-318.262
BIC	-245.726	-229.249	-235.027

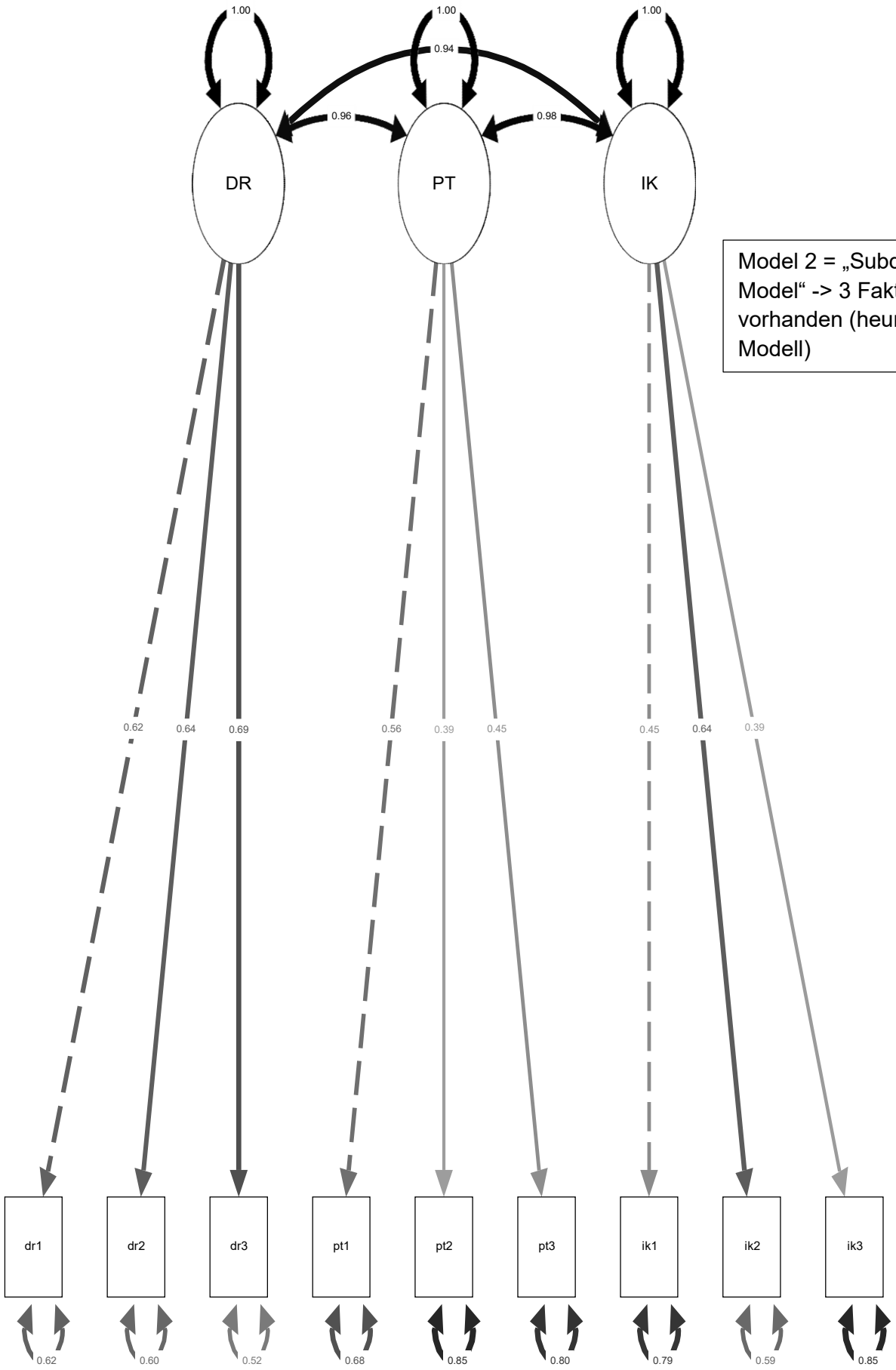
Jedes einzelne Strukturmodell zeigt sich, für sich betrachtet, in der Zuordnung nach seinen Dimensionalitätsparametern als durchaus geeignet und für Modelltestungen verwendbar (Tab. 9). Im CFI, NFI, RMSEA, SRMR und AIC dominiert leicht das Mod3, während im BIC das Mod1 überwiegt und der AIC an Position 2 folgt. Die „A priori“- Modelle ergeben sich als nicht negativ signifikant und keines muss verworfen werden.

Hier ist anzumerken, dass in den vorhandenen „TEMA“-Unterlagen die numerischen Beziehungen der einzelnen Faktoren zueinander nicht zu erkennen waren. Wie von Moosbrugger & Schermelleh-Engel vorgeschlagen, wurden deshalb die Faktorladungen in Mod1, Mod2 und Mod3 auf den Standardwert 1 fixiert (vgl. Moosbrugger & Schermelleh-Engel 2012: 335).

Die folgenden Abbildungen (Abb. 23, 24, 25) zeigen zunächst solo uni die drei gegeneinander getesteten Strukturgleichungsmodelle[248] als erzeugte Plots in „lavaan“ auf.

[248] Mod1 (Pflege), Mod2 (heuristisches Modell DR, PT, IK) und Mod3 (Setting)


1.00
Pfl
Model 1 = „General Factor Model“ -> Care, nur ein Faktor vorhanden
0.61
0.56
0.44
0.63
0.38
0.62
0.68
0.44
0.38
dr1
pt1
ik1
dr2
pt2
ik2
dr3
pt3
ik3
0.63
0.69
0.80
0.61
0.86
0.61
0.54
0.80
0.85



Model 2 = „Subdomain Model“ -> 3 Faktoren vorhanden (heuristisches Modell)

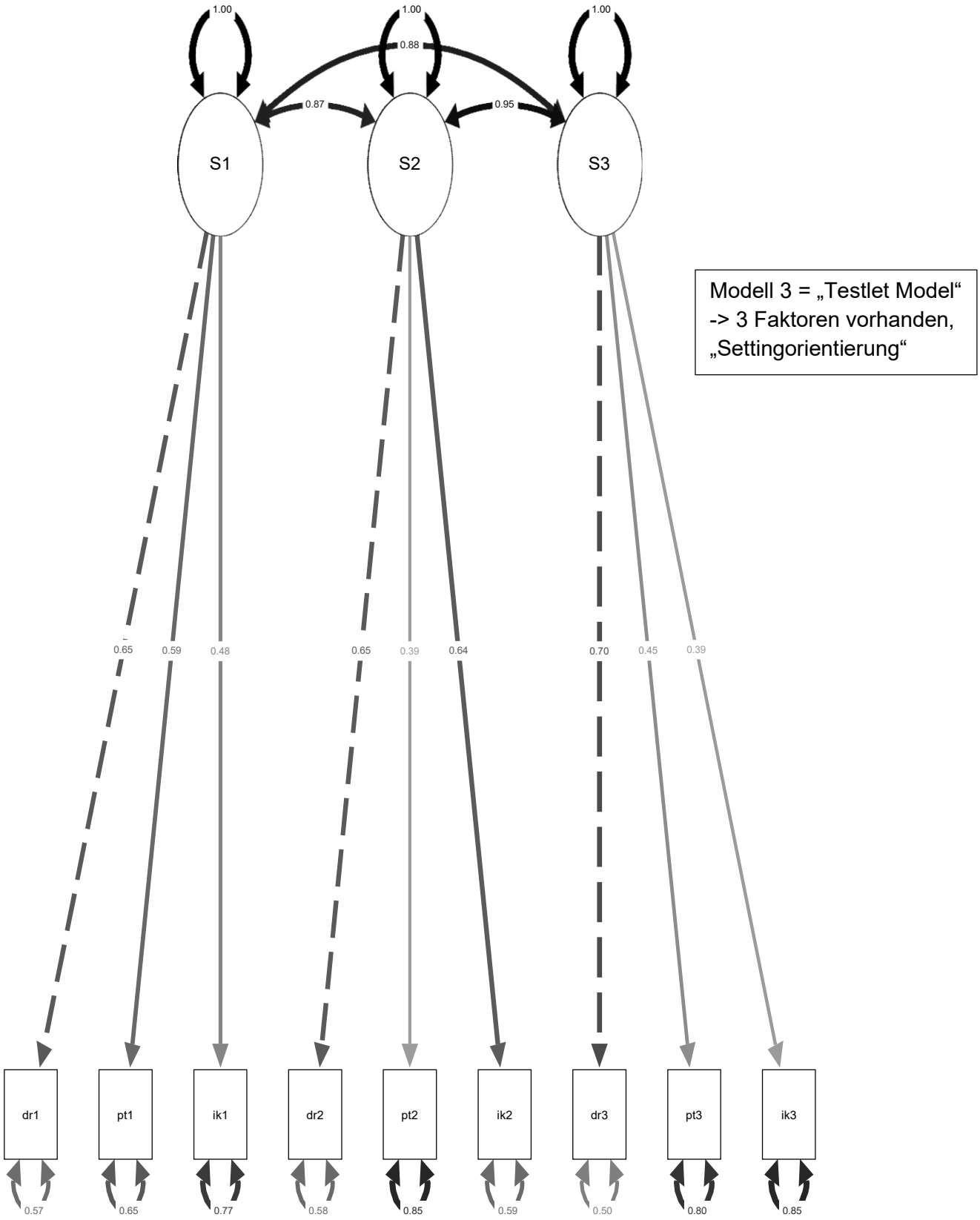


Abb. 23, 24, 25: Ergebnisse der drei Strukturgleichungsmodelle aus der CFA als Plots erzeugt mit R

Die einzelnen Pflegekompetenz-Strukturgleichungsmodelle werden mit ihrem Faktor bzw. ihren jeweiligen Faktoren, den dazugehörigen Variablen und entsprechenden Fehlervarianzen dargestellt. In der Einzelbetrachtung lassen sich die vorhandenen Maßzahlen so bewerten, dass das mehrdimensionale Mod3 auch hier hervorzuheben ist und sich besser als Mod1 präsentiert. Mod2 stellt sich wiederum schlechter dar als Mod1. Die Koeffizienten in allen Modellen sind sehr ähnlich. In Mod2 und Mod3 werden extrem hohe Korrelationen zwischen den 3 latenten Variablen angezeigt. Das bedeutet, dass diese Facetten nicht unabhängig sind.

Dies ist so zu verstehen, dass im Modellvergleich festgestellt werden kann, dass Mod2 und Mod3 auf einen Faktor zurückzuführen sind.

Tab. 11: Modellvergleichstabelle I (Anova Mod1 zu Mod2)

Modellvergleich	Mod1	Mod2
Chi² (= χ^2) Difference Test	1.4142	
Df Difference	3	
p-Wert	0.7022	

Tab. 12: Modellvergleichstabelle II (Anova Mod1 zu Mod3)

Modellvergleich	Mod1	Mod3
Chi² (= χ^2) Difference Test	7.1917	
Df Difference	3	
p-Wert	0.06603	

Tab. 13: Modellvergleichstabelle III (Anova Mod2 zu Mod3)

Modellvergleich	Mod2	Mod3
Chi² (= χ^2) Difference Test	-5.7774	
Df Difference	0	
p-Wert		

Keine Modelltestung möglich, da Df diff bei 0 liegt!

Es stellt sich anhand des durchgeführten Befehls „Anova“ in R heraus, dass das Mod1 als das „einfachere“ Modell zu bevorzugen ist, da weniger Faktoren notwendig sind, um die Ausprägungen der Variablen zu erklären. Der χ^2- Wert in Modellvergleichstabelle I (Tab. 11) liefert einen kleinen und somit guten Modellfit unter Berücksichtigung der Freiheitsgrade (vgl. Moosbrugger & Schermelleh-Engel 2012: 337).

In Modellvergleichstabelle III (Tab. 13) kann keine Modellvergleichstestung zwischen Mod2 und Mod3 stattfinden, da kein positiver Df diff vorhanden. Das ursprünglich entwickelte heuristische Modell aus „TEMA“, Mod2, spricht hier nicht für das „ideale“ Modell der Wahl. Zusammengefasst gilt das Mod1 als bestätigt![249] Das Teilergebnis der Dimensionalitätsanalyse für die Gesundheits- und Krankenpflege kann somit ebenfalls mit dem Ursprungsprojekt „TEMA“ gleichziehen.

Bezogen auf die in Kapitel 1.2 gestellte Forschungsfrage nach dem passenden Messmodell und der anzustrebenden H_0-Hypothese, dass eine gute Anpassung von den empirisch erhobenen Daten zum theoretisch konstruierten Modell bestehen sollte, kann aufgrund der durchgeführten statistischen Erhebungen bestätigt werden (vgl. Werner et al. 2016: 956): Das reduzierte Mod1 mit seiner eindimensionalen Konstruktion und seinen 9 Item-Paketen lässt sich als durchaus repräsentativ und erklärend für die ermittelten TN-Daten verstehen. Der Faktor „Care“ bzw. „praktische Pflegekompetenz“ kann mit den vorliegenden Items überwiegend gut erfasst werden.

[249] Anmerkung: Selbst bei Modell-Fit Einschätzung der „klassischen“ Goodness-of-Fit-Maße (= Goodness-of-Fit-Index/**GFI** und Adjusted-Goodness-of-Fit-Index/**AGFI**) mit empfohlenen Schwellenwerten von 0.9 und darüber, ohne Berücksichtigung der Stichprobengröße, bestätigen sich die bereits aufgeführten und ermittelten Dimensionalitätsanalysen (vgl. Weiber, Mühlhaus 2014: 210ff.). Die Interpretation der beiden Index-Werte wurde aufgrund des reduzierten N im Projekt der Gesundheits- und Krankenpflege in die Ergebnissicherung mit eingeschlossen. Der GFI und AGFI liegt bei allen drei Modellen über 0.9, d.h., die empirischen Varianzen und Kovarianzen werden durch die vorhandenen Modelle passend abgebildet (vgl. Weiber, Mühlhaus 2014: 211f.).

17 Abschließender Kommentar zum „Empirischen Teil 2“ - Statistische Aufbereitung und Datenanalyse

Die personenbezogenen Daten weisen in der Mehrzahl bei den TN ein reduziertes Interesse am Lesen auf, aber bei gleichzeitig täglicher Verwendung des Internets. Eine große Anzahl an TN erlebt die praktische Pflegetätigkeit als positiv und Beruf mit Spaß und einer Form von Erfüllung (siehe Kapitel 14.1, Tab. 3). Der Gedanke einer Überprüfung der beruflichen Lesekompetenz wäre für weitere Kompetenztestungen bei den TN mit einzubringen und sicher von Interesse im Kreise der Gesundheits- und Krankenpflege. Gleichzeitig lässt sich interpretieren, dass die emotionalen Kompetenzen im pflegerischen Alltag als nicht zu unterschätzend zu bewerten sind. Kaspar u. Hartig weisen in ihrem Fachartikel auf die beeinflussende Bedeutung von Auszubildenden mit emotionalen Zuständen im Praxisalltag hin. In den pflegerischen Praxisszenarien einschließlich Testitems wurde dies berücksichtigt. Die weitere Reaktionalität des Patienten sowie des gesamten pflegerischen Anschlussdialoges kann mit dem technologieorientierten Kompetenzmessinstrument aber nicht abgedeckt und weiterverfolgt werden (vgl. Kaspar, Hartig 2016: 115; vgl. Wittmann, Kaspar, Döring 2017: 200).

In der statistischen Auswertung zeigt sich das Kompetenzmessinstrument nach aufwendiger Neukodierung der Rohdaten bzw. des Rohdatensatzes auf die jeweiligen Skalenniveaus und der Übertragung in SPSS zur weiteren Verwendung in R als empirisch valides Messinstrument mit anspruchsvollen, aber gut entwickelten Testitems. Die Auswertungen mit der IRT zeigen teils hohe Itemschwierigkeiten mit teils folgernden reduzierten Personenfähigkeiten. Die Raschvalidität gilt aber aufgrund der fehlenden lokalen stochastischen Unabhängigkeit in den globalen Modellgeltungstests als nicht bestätigt. In der CFA stellt sich das Mod1 als durchaus handhabbar und passend heraus. Das Mod3 mit seiner settingorientierten Passung wäre bei weiterer Überarbeitung bzw. Ausbau des Kompetenzmessinstruments, aufgrund seiner guten Ergebnisse in den Dimensionsanalysen, wie u.a. im CFI und RMSEA, nicht vollständig zu vernachlässigen. Das ursprünglich entwickelte Mod2 mit seinem heuristischen Gedanken steht mit seinem empirischen Modellansatz hinten an.

Im vorliegenden statistischen Teilergebnis können in den durchgeführten Untersuchungen die Endergebnisse und Interpretationen aus „TEMA“, bezogen auf das Kompetenzmessinstrument und die Modellwahl nur teilweise bestätigt werden (vgl. Wittmann, Kaspar, Döring 2017: 199f.; vgl. Döring et al. 2016b: 129f.).

IV Resümee

Nach der durchgeführten, vollständigen Datenanalyse werden die in Kapitel 1.2 gestellten Forschungsfragen beantwortet. Darüber hinaus schließen sich Reflektionen für neue Forschungsvorhaben und wissenschaftlich-methodische Weiterarbeit an.

Als Bonus ergeben sich zusätzliche Vorschläge für fortführende pflegewissenschaftlich-pflegepädagogische Projekte.

18 Zusammenfassung der Endergebnisse im durchgeführten Projekt einschließlich weiterer Anregungen

Die Übertragbarkeit des technologiebasierten, computergestützten Kompetenztests in die Gesundheits- und Krankenpflege und darüber hinaus in die Gesundheits- und Kinderkrankenpflege ist möglich. Der Einsatz zur Überprüfung der praktischen Pflegekompetenz ist sinnvoll und bietet sich an (siehe komplette Analyse und Evaluation in den Kapiteln II und III). Es werden die praktischen Facetten für den gesamten TN-Kreis ausreichend abgebildet. Für die TN aus der Gesundheits- und Kinderkrankenpflege ist eine zusätzliche Erweiterung und Spezifizierung auf „Säugling und Kind“ notwendig. Aufgrund der Ausrichtung auf die zukünftige generalistische Ausbildung bzw. den generalistischen Studiengang Pflege ist das Messinstrument als durchaus pädagogisch wertvolle Unterstützung zum Ende der dreijährigen Ausbildung bzw. des Studiums zu sehen und in die Lehrpläne mit aufzunehmen. Ebenso ist eine Ergänzung für den klassischen Klinikalltag im Setting der „Akutversorgung“ absolut wünschenswert und angebracht (vgl. Döring et al. 2016b: 130; vgl. Döring et al. 2014: 244f.).

In den Subfragestellungen zeigen sich in den Gesamtergebnissen der Gesundheits- und Krankenpflege annähernde Parallelen zum Ursprungsprojekt „TEMA“. Hinzuweisen ist hier nochmals auf die erhobene vergleichbare TN-Zahl (siehe Kapitel 7.1.3, 9.3 und 11), die ermittelten Antworthäufigkeiten der einzelnen Testitems im Kompetenztest in SPSS mit Hinweisen der Unterscheidung in „TEMA“ (siehe Kapitel 14.2), die positive Beantwortung der „CosMed“-Items der TN mit der Tendenz zu noch besseren Endergebnissen als in „TEMA“ (siehe Kapitel 14.2) und der Bevorzugung des eindimensionalen Modells bei der CFA (siehe Kapitel 16.2 und 17). Bei einer weiteren Itemmodellierung innerhalb des Kompetenztests sollte die Selektion der „CosMed“-Items in die jeweiligen Pflegesettings mit aufgenommen und mit der IRT auf Eignung überprüft werden.

Zu der Subfragestellung auf Hintergründe und Verknüpfungspunkte im Kompetenztest sei auf eine mögliche und notwendige Entfaltung von Pflegekompetenz, der Entwicklung eines geeigneten Instrumentes für die Messung von praktischer Pflegekompetenz, einer Form von aufzufindender Kompetenzdiagnostik für die tägliche Pflegepraxis hingewiesen (siehe Kapitel 1.1). Erstmalig umgesetzt in der groß angelegten ASCOT-Studie (siehe Kapitel 4 u. 4.1), die auf breiter Basis für diverse praktische Berufsbilder grundlegende Kompetenzprofile entwickelt hat, mit einer Schwerpunktsetzung in der Altenpflege in Form des „TEMA“-Testinstruments (siehe

Kapitel 4.2). Aus dem Projekt „CosMed“ wurden weitere 12 evaluierte Items an den Pflegekompetenztest mit angedockt (siehe Kapitel 4.3). Eine hohe Relevanz ist in der Gesundheits- und Krankenpflege mit der zukünftigen generalistischen Ausbildung zuzuschreiben (siehe Kapitel 2 u. 3; vgl. Fachkommission 2019 - Rahmenausbildungspläne). Die Notwendigkeit zeigt sich bei allen Stakeholdern von hohem Belang!

In der Subfragestellung auf Handhabung bzw. Praktikabilität des „TEMA“-Messinstrumentes ist auf einen hohen Aufwand beim Aufbau, der Durchführung und dem Abbau hinzuweisen (siehe Kapitel 10). Für weitere Anwendungen im Feld wäre eine EDV-technische Vereinfachung sinnvoll, indem beispielsweise die Möglichkeit der Einwahl über ein öffentliches Netzwerk mit dem eigenen Notebook erfolgen könnte. Dies würde den monetären und zeitlichen Aufwand deutlich absenken und vereinfachen. Allerdings ist an der jeweiligen Schuleinrichtung eine funktionsfähige und stabile Netzwerkdimension unter Beachtung des passenden Typs[250] notwendig mit unbegrenzter Einwahl an Notebooks. Eine abgekoppelte Datenerhebung unter verstärktem Einbezug der Lehrkräfte ist unter den veränderten technischen Bedingungen einfacher möglich, birgt aber auch Risiken im regulären wissenschaftlichen Forschungsprozess.

Die letzte Subfragestellung bezieht sich auf die Feststellung von praktischer Pflegekompetenz anhand der wissenschaftlich modellierten Testitems. Hier ist zum einen auf die noch folgenden statistischen Fragestellungen einschließlich Antwortkonstrukt hinzuweisen, und zum anderen zeigt der „TEMA“-Test durch seine Videosequenzen mit darauf aufbauenden empirisch gehaltvollen Testitems eine curricular und beruflich, generalistisch orientierte, geeignete Bandbreite für die praktische Pflegekompetenz an. Der Einbezug der sozial-emotionalen Momente zeigt sich als durchaus komplex im entwickelten Scoring und der Auswertung. Der Schwerpunkt liegt in der Abfrage von Wissen bei vorliegendem geschlossenem Antwortsystem im täglichen pflegerischen Praxisalltag. Der Bereich des bewohner- und klientenbezogenen Kompetenzbereichs bestätigt sich in der gesamten vorliegenden Arbeit mit dem Schwerpunkt Gesundheits- und Krankenpflege einschließlich in „TEMA“ als absolut notwendig in den vorliegenden Settings (siehe Kapitel 4.2, 4.3 u. 5). Das pflegerische Kompetenzprofil sollte auf weitere Testitems zu den Themen Wissenschaft, Akutversorgung, Arbeiten in Teams, Motorik und Sensibilität ausgeweitet werden. Ebenso ist das zu Beginn von „TEMA“ entwickelte „heuristische Modell“ mit den organisationsbezogenen und selbstbezogenen Kompetenzen in den beruflichen Kompetenztest erweitert mit aufzunehmen (vgl. Wittmann, Kaspar, Döring 2017: 189 u. 199f.; vgl. Döring et al. 2014: 251ff.). Folglich ist mit dem pflegerischen „TEMA“-Kompetenzmessinstrument eine grundlegende Möglichkeit über die entwickelten Testitems geschaffen worden, eine individuelle Einschätzung des TN zur Vorbereitung auf die praktische Prüfung zu ermöglichen und darüber hinaus eine Hilfestellung für Lehrkräfte zur personenbezogenen Förderung einer notwendigen komple-

[250] wie LAN (Local Area Networks), WLAN (Wireless Local Area Networks) oder VPN (Virtual Private Networks)

xen praktischen Pflegekompetenz zu leisten. Ein weiterer empirischer Ausbau an relevanten Testitems ist deshalb unbedingt weiterzuverfolgen!

Die statistische Fragestellung zu Häufigkeiten im laufenden Projekt ist für das „TEMA“-Kompetenzmessinstrument im Endergebnis vergleichbar mit dem Ursprungsprojekt in der Altenpflege (siehe Kapitel 14.2). Im verwendeten „CosMed“-Itempool ergibt sich sogar noch ein positiveres Ergebnis (siehe ebenfalls Kapitel 14.2). Aus datenschutzrechtlichen Gründen wurde dabei auf spezifizierte Gruppenzuweisungen pro Bildungseinrichtung verzichtet! Die wissenschaftliche Aufbereitung von weiteren Testitems aus dem beruflichen Feld der „Medizinischen Fachangestellten“ wäre lohnenswert, da die bereits vorhandenen Items im Feld getestet und ausgewertet wurden. Weitere berufliche Vergleichsstudien eröffnen so einen Blick auf gemeinsame Felder der einzelnen Ausbildungsberufe.

In der folgenden Fragestellung zur Eignung der Item-Response-Theorie als Auswertungsverfahren bildet das metrische Latent-Trait-Modell eine empirisch verlässliche, abzubildende Einschätzung für die vorliegenden Items und vorhandenen TN-Fähigkeiten ab. Die Passung der modellierten Testitems konnte wissenschaftlich fundiert und vielfältig empirisch überprüft werden mit Feststellung von einzelnen „Hotspots“ (siehe Kapitel 15.4). Bei der Durchführung von globalen Modellgeltungstests auf Signifikanz wurde das Gesamtmodell abgelehnt. Eine vorhandene Raschvalidität konnte nicht ermittelt werden, mit der Folgerung einer Ablehnung der H_0-Hypothese (siehe Kapitel 15.4 u. 17).

Die abschließende statistische Fragestellung zur CFA und des passenden Messmodells kann mit der Dimensionalität des „einfaktoriellen“ Kompetenzmessmodells oder der auch „General Factor Model“ bestätigt werden (siehe Kapitel 16.2 u. 17). Eine Zurückführung auf den latenten Faktor „Client-Directed-Care“ wird in den mehrdimensionalen Modellen verifiziert (siehe Kapitel 16.2).

Als Fazit der statistischen Erhebungen zum „Large-Scale-Assessment“ gilt, dass das Messen von praktischer Pflegekompetenz durch die Befunde in der IRT und CFA möglich ist, trotz der nicht erfüllten statistischen Strenge nach Rasch. Eine Selektion von Items aufgrund zu niedriger p-Werte müsste durchgeführt werden. Die vorhandene TN-Zahl ist als geeignet zu bewerten. Die Folge von vorhandenen Schätzproblemen ist deshalb in der durchgeführten CFA auf ein Minimum reduziert (vgl. Bühner 2011: 431f.). Das Modell ist durch überwiegend passable Items gekennzeichnet mit der Beziehung auf die vorliegenden latenten Variablen bzw. deren Faktor/en. Eine Identifizierung des Mod 1 ist erreicht!

Abschließend sei hier noch auf die zu Beginn der Arbeit gestellten Zielsetzungen der Forscherin verwiesen. In der vorliegenden Arbeit konnten nach ausführlicher Diagnostik und Analyse des Kompetenzmessinstruments und seiner TN alle erreicht werden. Die evidenzbasierte Übertragung (siehe Kapitel 2 bis 19), das breite TN-Feld im fast gesamten süddeutschen Raum (siehe Kapitel 9.2), die Validitätsprüfung (siehe Kapitel 12 bis 19) mit gleichzeitiger Eignung des „TEMA“-Kompetenzmessinstrumentes (siehe Kapitel 7 bis 19) und die Verbreitung von der Idee eines

möglichen „Hilfsmittels“ für Bildungseinrichtungen bzw. deren Auszubildenden im technologiebasierten, computergestützten Bereich zur Vorbereitung auf die praktischen Abschlussprüfungen in der Gesundheits- und Krankenpflege konnte umfassend erarbeitet werden (siehe Kapitel 9 bis 11).

Das entwickelte pflegerische Kompetenzmessinstrument aus dem Ursprungsprojekt „TEMA“ zeigt sich in der Gesamtbetrachtung als Herausforderung für die TN, aber als empirisch überprüftes Messinstrument in der beruflichen Bildung im Themenfeld „Pflege“!

19 Impulse und Ausblick für weitere Forschungsvorhaben

Der Kompetenzbegriff stellt sich als sehr heterogen dar (siehe Kapitel 3.1). Als Ergänzung sei hier die empirische Studie von Dütthorn erwähnt, die als eine Kernkompetenz „Pflegerische Beziehungen gestalten“ für notwendig im beruflichen Alltag betrachtet. Eine didaktische Ausgestaltung ist auf spezifische „Lehr- und Lernarrangements“ unter Berücksichtigung der individuell anzutreffenden, problemorientierten Fallsituationen zu beziehen. Der Focus richtet sich nicht nur auf die rein kognitiven Fähigkeiten, sondern auch auf die kommunikativ-emotionalen Fähigkeiten mit der Prägung von persönlichen Steuerungsmechanismen (vgl. Dütthorn 2014: 140f.; vgl. Dütthorn 2016: 10ff.). So können über den gesamten pflegerischen Ausbildungsverlauf multiple Kompetenzen geweckt, kontinuierlich ausgebaut und verwirklicht werden (vgl. Hundenborn 2019: 10; vgl. Darmann-Finck, Einig 2019: 91ff.). Die Verbindung wird im entwickelten „TEMA“-Kompetenzmessinstrument vorgefunden und entspricht der vorangestellten Denkweise. Deshalb sind Ausweitungen in der Instrumentenentwicklung für weitere Projektvorhaben als durchaus sinnig und notwendig zu betrachten und voranzutreiben!

Erweiternd sollten von wissenschaftlicher Seite zukünftige Richtlinien für ein notwendiges „Kompetenzverständnis“ in der Pflege einschließlich eines empirischen „Kompetenzkonstrukts“ für die zukünftige generalistische Pflegeausbildung unter Einbezug des Stuttgarter Modells entworfen und aufgestellt werden (vgl. Brater 2016: 197ff.; vgl. Holoch 2007: 379ff.). Curriculare Kompetenzprägungen finden sich bei Hundenborn und Darmann-Finck, Einig als Anregungen für die Umsetzung in der jeweiligen Bildungseinrichtung (vgl. Hundenborn 2019: 5ff.; vgl. Darmann-Finck, Einig 2019: 85ff.).

Wie in Kapitel 14.1 erwähnt, sind, über die Fragestellungen des Projektes hinaus, weitere Erkenntnisse aus den personenbezogenen Daten, wie beispielsweise einem linearen Zusammenhang von höherer Schulbildung und Lösungswahrscheinlichkeit der Items, durch den statistischen Einsatz der Rangkorrelation von pädagogischem Interesse (vgl. Kreienbrock, Pigeot, Ahrens 2012: 395).

Ein Folgeprojekt mit der erweiterten Modellierung von Testitems aus den noch fehlendenden Kompetenzbereichen des „Heuristischen Modells“, organisationsbezogener Kompetenzbereich und selbstbezogener Kompetenzbereich, wäre absolut wünschenswert. Bei den Itemmodellierungen ist auf prägnante und klare Antwortoptionen in einem passenden Umfang zu achten, die bei „Speedtestings“ die kognitive Auffassung erleichtern. Eine generelle Erweiterung auf die Bereiche der „Akutversorgung“ mit Videosequenzen der Ersten Hilfe, Führung von Kollegen in Krisensituationen usw. und „Neugeborene, Kinder und Jugendliche“ unter den Gesichtspunkten „Kommunikation und Empathie“, gepaart mit pflegewissenschaftlichen Aspekten, aufgrund der generalistischen Ausbildung ab 2020, ist als ein „Muss“ zu verstehen! So würde ein komplett entwickeltes, technologiebasiertes Messinstrument für eine individuelle und grundsätzliche Beurteilung von beruflicher, praktischer Pflegekompetenz zur Verfügung stehen! Zu berücksichtigen ist aber, dass durch die Erweiterung der fehlenden Kompetenzbereiche, die Bearbeitungszeit des Kompe-

tenztests nicht parallel dazu noch weiter erhöht wird, da eine konzentrierte Bearbeitung nach über 120 Minuten nicht mehr vollständig gewährleistet sein kann. (vgl. Smith 2012: 175; vgl. Porst 2014: 19ff.; vgl. Döring, Bortz 2016: 430ff.).

Da Begrenzungen in den statistischen Analysen gegeben sind, wäre eine Möglichkeit der weiteren Spezifikation, die Erweiterung durch den ursprünglichen Datensatz aus „TEMA“ noch ergiebiger. Es könnten beispielsweise „Matchings“[251] durchgeführt werden. Als weitere Fragestellungen ist die Fortentwicklung der WLE-Reliabilität von 0.868 zu nennen (siehe Kapitel 15.4), sowie die Dimensionalitätsanalyse in Bezug zu Mod3 (siehe Kapitel 16.2 u. 17) als interessant zu bewerten.

Die Fokussierung der hier durchgeführten Kompetenztestungen ausschließlich auf die pflegerischen Aspekte stellt eine gewisse Einseitigkeit dar! Eine Erweiterung sollte in den grundlegenden Bereichen der mathematischen Kompetenz und Lesekompetenz vorgenommen werden, trotz der überwiegend in Deutschland geborenen TN! Beispielsweise ist für den Bereich der „Akutversorgung“ eines Patienten unter Verabreichung von Infusionslösungen mit vorheriger Berechnung der Tropfengeschwindigkeit/Min. und dem Lesen des dazugehörigen Beipackzettels zum korrekten Gebrauch, die Ausprägung der genannten Kompetenzen nicht wegzudenken. Auch die weitere Anwerbung von Pflegekräften aus dem Ausland stärkt das Vorhandensein der zuvor genannten Kompetenzen. Diese Feststellungen unterstützt auch Knappich in seinen Ausführungen zur Ermittlung des mathematischen Textverständnisses mit einer notwendigen und umfassenden Kompetenzermittlung (vgl. Knappich 2017: 173).

Es wäre anzustreben, dass das entwickelte „TEMA“-Testinstrument an den einzelnen Bildungseinrichtungen, als „Probelauf“ zunächst einmal im gesamten süddeutschen Raum, zur Unterstützung für die Vorbereitungen auf die praktischen Abschlussprüfungen in der Pflege den Lehrkräften sowie Auszubildenden zur Verfügung steht. Dazu sollten ein erweiterter, austauschbarer Itempool oder auch weitere Testhefte entwickelt und vorhandene weiterentwickelt werden. So könnte ein Wiedererkennungseffekt vermieden werden. Weitere praktische Skills sind aufgrund des absoluten Defizits der TN bei Item A48 im Kompetenztest unbedingt einzuüben, umgesetzt durch konkrete, videogestützte Fallszenarien, mit anschließend gekoppelten klinischen Praxisanleitungen bzw. Praxisbegleitungen zum Thema: Injektion (wie z.B. s.c. oder i.m.[252]) (siehe Kapitel 14.2). Zusätzlich sollten in der erweiterten Itementwicklung Fragestellungen mit offenem Antwortformat modelliert werden, mit der Möglichkeit einer noch individuelleren Beurteilung und Kompetenzeinschätzung des einzelnen TN. Die methodische Datenanalyse muss dann neu betrachtet werden, ebenso das notwendige Scoring der einzelnen Aufgaben.

Für eine Verbesserung der Personenfähigkeiten im Kompetenztest wäre eine Lösung mit der Form des adaptiven Testens zu erarbeiten, mit dem Hintergrund der

[251] Paarbildung; als individuelles Matching oder Häufigkeitsmatching (vgl. Kreienbrock, Pigeot, Ahrens 2012: 199f.)
[252] intramuskulär

IRT. Es werden nach einem gesteuerten System die Aufgaben dem jeweiligen TN vorgelegt. Pospeschill erläutert dazu:

> „Beim **adaptiven Testen** orientiert sich die Auswahl aus einem Itempool bzw. die Vorgabe einzelner Items am spezifischen Antwortverhalten einzelner Probanden. Zumeist erfolgt dies durch eine Anpassung der Itemauswahl hinsichtlich ihrer Schwierigkeit an die Fähigkeit des Probanden. Die angepasste Vorgabe von Items soll dabei mehr diagnostischen Nutzen bringen (möglichst exakte Informationen zur individuellen Merkmalsausprägung) und gleichzeitig testökonomisch sein (durch Vorgabe einer reduzierten Anzahl von Items).“ (Pospeschill 2010: 153 u. 154).

Auch über die spezielle Form der „Selbstadaptation“[253] im Test sollte nachgedacht werden, da bisher allgemein nur vereinzelte Kompetenzmessinstrumente für den pflegerischen Kontext zur Verfügung stehen, die meist über das Versuchsstadium nicht hinausgehen (vgl. Darmann-Finck, Reuschenbach 2013: 28).

Als Gewinn für die TN und Forscher wäre eine computergestützte, technologisch entwickelte Möglichkeit der sofortigen Rückmeldung der Ergebnisse im Kompetenztest[254] und darüber hinaus eine Gesamtrückmeldung zur Häufigkeitsverteilung der Items des jeweiligen Kurses bzw. der Kohorte anzusehen. So wäre eine leichtere Auswertung der TN-Daten bei gleichzeitig reduziertem zeitlichem Aufwand erreicht!

Der dauerhafte Erwerb von pflegerischem Fachwissen und die Einschätzung von notwendigem, professionellem Pflegehandeln bedürfen eines kontinuierlichen Trainings, angestoßen von Seiten der Lehrenden und vertieft und in Selbstlernphasen weitergeführt von den Lernenden.[255] Es sind aus diesem Grund inhaltlich für die zukünftige Ausrichtung des „modernen“ Pflegeberufs verstärkte Lehrvermittlung in pflegewissenschaftlichen Aspekten gefragt, die auch die akademische Herangehensweise an Pflegeprobleme als neuen Weg für die tägliche Pflege ermöglichen. Das könnte beispielsweise über entwickelte Lerntutorials umgesetzt werden, die den Auszubildenden in den jeweils zu erwerbenden Kompetenzbereichen einschließlich deren Kompetenzschwerpunkten[256] zur Verfügung gestellt werden. Die Vorbereitung für die Lehreinheit und deren Basisvermittlung der gesetzlichen und curricularen Kompetenzen würde dann zunächst zu Hause stattfinden. Die Evaluation einschließlich der Entwicklung einer gemeinsamen Lösungsstrategie zur Bewältigung des betrachteten Pflegeszenarios findet im theoretischen wie praktischen Unterricht statt. Einen positiven Anstoß für eine zukunftsweisende, kompetenzfördernde Lernstrategie ist beispielsweise „Dominiks Pflegeuniversum“[257], gedacht als „Ausschnittslernen“ für den wissenschaftlichen, praktischen Pflegealltag, allerdings nicht

[253] Der TN wählt selbst die weiterführenden Aufgaben aus.

[254] Zusätzlich, zur weiteren Förderung des Lernerwerbs, könnten Vertiefungsvideos zum Wissenserwerb im Anschluss an jedes beantwortete Item entwickelt werden. Neue Informationen werden so vertieft und erlernt. Zwei Vertiefungsvideos wurden als mögliche Idee von der Autorin entwickelt und können bei Interesse angefragt werden.

[255] beispielsweise in „Pflegelaboren“ der Bildungseinrichtungen

[256] Formulierung bezogen auf die neue Ausbildungs- und Prüfungsverordnung – Pflegeberufegesetz ab 2020

[257] zu finden auf YouTube z.B. das Tutorial „URL: https://www.youtube.com/watch?v=6B-3zOZ60R4 [Stand: 19.05.2019]“ über Notwendigkeit von Mundpflege im Pflegealltag

im Frage-Antwort-Modus als Überprüfung von Wissen im Rahmen von erworbenen Lerninhalten und Handeln in verschiedenen Settings, sondern in Form von interessenweckender Wissensvermittlung und der Entwicklung von Verständnis für empiriegeleitetes Wissen in der Pflegepraxis. Hier wird die persönliche Einschätzung von praktischen pflegerischen Kompetenzen nicht mit berücksichtigt, sondern weiterer praktischer Wissenserwerb „nur“ geweckt (vgl. Bruch, Dichter 2018: 329ff.).

Eine Weiterführung des Forschungsprojektes sollte deshalb mit dem Ansatz begonnen werden, dass eine Erleichterung bei der Durchführung des Kompetenztests ermöglicht wird, und das unter dem Focus der Bereitstellung und Handhabung mit Bezug auf die neuesten computergestützten, technischen Entwicklungen. Der Einbezug von Computerspezialisten ist für die Weiterentwicklung sicher von Vorteil und zu begrüßen! Eine zukunftsweisende Möglichkeit bietet dabei durchaus der Einsatz von virtuellen Realitäten[258]. Dies ermöglicht dem TN das Eindenken, Mitdenken, Handeln und Miterleben in Patient-Pflege-Situationen auch unter „fast“ realen Bedingungen. Das Eintauchen in die virtuelle Welt ruft beim Anwender bzw. TN „Echtheit“ im täglichen Pflegealltag hervor. So spielen auch Kommunikation und Empathie eine verstärkte, tragende Rolle in den technologie- bzw. computergestützten Handlungsszenarien und nicht nur das Vorbereiten, Durchführen und Nachbereiten von Pflegetechniken bzw. Pflegesituationen. Als Alternative könnte auch mit „realen Videovignetten“ nach dem Vorbild des Konzeptes der TU Dresden gearbeitet werden[259]. Der Vorteil läge hier in der unmittelbaren Besprechung der gesehenen Fallsequenzen mit Herausarbeitung einer Strategie einer kompetenten, individuellen, pflegerischen Versorgung des zu betreuenden Patienten. Modellierte Items wären mündlich von der Lehrkraft mit einzubringen und mit den Auszubildenden an Lösungsansätzen zu arbeiten und zu diskutieren. Eine empirisch quantitative Aufbereitung wäre so allerdings nicht möglich. Deshalb können in diesem Zusammenhang die genannten „Realitäten“ nur als ergänzende oder vertiefende Vorbereitungsmöglichkeit zu einer standardisierten Überprüfung auf die praktische Prüfung in der Pflege gesehen werden (vgl. König, Lindwedel-Reime 2019; vgl. Altmeppen 2019: 175ff.; vgl. Schlegel, Geering, Weber 2020: 57ff.).

Es ist deshalb einer zukünftigen, praktischen, kompetenzorientierten Ausbildung in der Gesundheits- und Krankenpflege, Gesundheits- und Kinderkrankenpflege und Altenpflege eindringlich zu wünschen, dass sie Mindeststandards für eine „gute“ Pflege am Bett festlegt und als positive dauerhafte Folge eine zufriedenstellende Versorgung und Wertschätzung für Patient und Personal entsteht. Eine professionelle pflegerische Kombination aus fachlich-technischem, kommunikativem und empathischem Standing ist dafür unabdingbar und notwendig. Es darf nicht nur das monetäre, kurzfristig denkende Konstrukt einer grenzenlosen Gewinnorientierung im Mittelpunkt stehen, bei der der Faktor Mensch sich als größtes Hindernis darstellt (vgl. Allert 2019: 36ff.)!

[258] abgekürzt VR
[259] Im Mittelpunkt steht die Entwicklung von „hermeneutischer Fallkompetenz“ in der Lehrerbildung.

Es gilt deshalb, Politik, Berufsverbände, Pflegekammern, öffentliche Einrichtungen, Lehrende und Lernende darauf aufmerksam zu machen, einen positiven Wandel und ein verbessertes Image für die Pflege zu erzielen, um Menschen mit den notwendigen Kompetenzen professionell ausbilden zu können. Dafür ist die Entwicklung von geeigneten „Hilfsmitteln“ unter Beachtung der zukünftigen generalistischen Ausrichtung gefragt! Es ist zu fordern, dass der BMBF weitere Gelder für die berufliche Bildung investiert, um eine zukunftsfähige Ausrichtung für eine komplexe Ausbildung bzw. ein hochkomplexes Studium in Deutschland zu gewährleisten.

Der Einsatz von computergestützten, technologiebasierten Kompetenzmessinstrumenten in der beruflichen Bildung von Pflegeberufen sollte den Lehrenden sowie Auszubildenden nicht genügen, um ein persönliches Kompetenzprofil für die tägliche Arbeit in der Pflege zu beschreiben, sondern sie können helfen, eine grundsätzliche Einschätzung von benötigten Pflegekompetenzmerkmalen vorzunehmen. Das bedeutet aber nicht, dass der Blick auf die notwendige Entwicklung von einem subjektiven und umgebungsorientierten Pflegekompetenzprofil abgedeckt ist. Der Einsatz von standardisierter Kompetenzdiagnostik ist für die angeführten Kompetenzbereiche in Form von quantitativen Messverfahren nur schwer möglich bzw. umzusetzen (vgl. Darmann-Finck, Grissmann 2011: 202; vgl. Sauter, Staudt 2016: 20f.).

Abschließend ist noch anzumerken, dass das entwickelte „TEMA“-Messinstrument auch als ein mögliches Einschätzungskriterium für Neueinstellungen von Pflegekräften in den videographisch erfassten Settings oder als internes klinisches Weiterbildungsangebot für examinierte Pflegekräfte genutzt werden kann. Auch internationale berufliche Vergleichsstudien, wie in PISA, wären wissenschaftlich zu untersuchen und zu prüfen. Der Weg der Kompetenzdiagnostik ist in Zukunft mit dem Gedanken und Ziel einer zukunftsorientierten praktischen Pflege für alle Shareholder und Stakeholder unbedingt weiterzuverfolgen!

Abschlussgedanken

„Eine neue Art von Denken ist notwendig, wenn die Menschheit weiterleben will.“ (Albert Einstein)

„Es kommt darauf an, dass ihr nicht so leicht aufgebt!“ (Stephen Hawking)

„Diese Welt braucht nicht uns, aber wir brauchen diese Welt!“ (Simone Ries, 2018)

Literaturverzeichnis[260]

Alexandrowicz, Rainer (2013): R in 10 Schritten. 1. Aufl. Stuttgart: UTB GmbH; facultas.wuv.

Allert, Rochus (2019): Spiritueller Impuls. Christliche Spiritualität, die Ökonomie und das Geld. In: *PflegeLeben. Pflege-Werte- Zukunft - das Magazin für Mitglieder des Katholischen Pflegeverbandes e.V.*, S. 36–39.

Altmeppen, Sandra (2019): Wie kann hermeneutische Fallkompetenz gelingen? In: *PADUA* 14 (3), S. 175–180. DOI: 10.1024/1861-6186/a000489.

Anderson, D. R.; Burnham, K. P.; White, G. C. (1998): Comparison of Akaike information criterion and consistent Akaike information criterion for model selection and statistical inference from capture-recapture studies. In: *Journal of Applied Statistics* 25 (2), S. 263–282. DOI: 10.1080/02664769823250.

ASCOT-Forschergruppe (2014): Projekt TEMA: Entwicklung und Erprobung von technologieorientierten Messinstrumenten zur Feststellung beruflicher Handlungskompetenz in der Pflege älterer Menschen. gefördert vom Bundesministerium für Bildung und Forschung, 2014. Online verfügbar unter - nicht veröffentlicht -, zuletzt geprüft am 04.03.2019.

Bachmann, Monika; Schlegel, Claudia (2018): «Die Prüfung ist anspruchsvoll, aber beliebt». Ein Gespräch über die Vermittlung von Kompetenzen in der Pflegeausbildung. Monika Bachmann interviewt Claudia Schlegel. In: *PADUA* 13 (5), S. 349–351. DOI: 10.1024/1861-6186/a000461.

Backhaus, Klaus; Erichson, Bernd; Plinke, Wulff; Weiber, Rolf (2016): Multivariate Analysemethoden. Eine anwendungsorientierte Einführung. 14., überarbeitete und aktualisierte Auflage. Berlin, Heidelberg: Springer Gabler. Online verfügbar unter http://dx.doi.org/10.1007/978-3-662-46076-4.

Backhaus, Klaus; Erichson, Bernd; Weiber, Rolf (2015): Fortgeschrittene multivariate Analysemethoden. Eine anwendungsorientierte Einführung. 3., überarbeitete und aktualisierte Auflage. Berlin, Heidelberg: Springer Gabler (Lehrbuch).

Baethge, Martin; Seeber, Susan (2016): Die gemeinsame theoretische und methodische Basis der ASCOT-Projekte. In: Klaus Beck, Margarete Landenberger und Fritz Oser (Hg.): Technologiebasierte Kompetenzmessung in der beruflichen Bildung. Ergebnisse aus der BMBF-Förderinitiative ASCOT. 1. Auflage. Bielefeld: W. Bertelsmann Verlag GmbH & Co. KG (Wirtschaft - Beruf - Ethik, 32), S. 15–31.

Baethge-Kinsky, Volker; Baethge, Martin; Lischewski, Julia (2016): Bedingungen beruflicher Kompetenzentwicklung: institutionelle und individuelle Kontextfaktoren (SiKoFak). In: Klaus Beck, Margarete Landenberger und Fritz Oser (Hg.): Technologiebasierte Kompetenzmessung in der beruflichen Bildung. Ergebnisse aus der BMBF-Förderinitiative ASCOT. 1. Auflage. Bielefeld: W. Bertelsmann Verlag GmbH & Co. KG (Wirtschaft - Beruf - Ethik, 32), S. 265–293.

Bandura, Albert (1976): Die Analyse von Modellierungsprozessen. In: Albert Bandura und Hainer Kober (Hg.): Lernen am Modell. Ansätze zu einer sozial-kognitiven Lerntheorie. 1. Aufl. Stuttgart: Klett, S. 9–68.

Bandura, Albert; Kober, Hainer (Hg.) (1976): Lernen am Modell. Ansätze zu einer sozial-kognitiven Lerntheorie. 1. Aufl. Stuttgart: Klett.

Bandura, Albert; Verres, Rolf; Kober, Hainer (1979): Sozial-kognitive Lerntheorie. 1. Aufl. Stuttgart: Klett-Cotta (Konzepte der Humanwissenschaften).

Bartholomeyczik, Sabine (2001a): Professionelle Kompetenz in der Pflege. Teil I. In: *Pflege aktuell. Fachzeitschrift des Deutschen Berufsverbandes für Pflegeberufe* (5), S. 284–287.

Bartholomeyczik, Sabine (2001b): Professionelle Kompetenz in der Pflege. Teil II. In: *Pflege aktuell. Fachzeitschrift des Deutschen Berufsverbandes für Pflegeberufe* (6), S. 344–347.

Bartholomeyczik, Sabine (2001c): Professionelle Kompetenz in der Pflege. Teil III. In: *Pflege aktuell. Fachzeitschrift des Deutschen Berufsverbandes für Pflegeberufe* (7-8), S. 412–414.

Beck, Klaus; Landenberger, Margarete; Oser, Fritz (2016): ASCOT- Ein Forschungsprogramm zu Kompetenzerwerb und Kompetenzmessung in Ausbildungsberufen. In: Klaus Beck, Margarete Landenberger und Fritz Oser (Hg.): Technologiebasierte Kompetenzmessung in der beruflichen Bildung. Ergebnisse aus der BMBF-Förderinitiative ASCOT. 1. Auflage. Bielefeld: W. Bertelsmann Verlag GmbH & Co. KG (Wirtschaft - Beruf - Ethik, 32), S. 9–13.

[260] Die gesamte Literaturverwaltung erfolgte mit dem Computerprogramm „CITAVI 6“.

Beck, Klaus; Landenberger, Margarete; Oser, Fritz (Hg.) (2016): Technologiebasierte Kompetenzmessung in der beruflichen Bildung. Ergebnisse aus der BMBF-Förderinitiative ASCOT. 1. Auflage. Bielefeld: W. Bertelsmann Verlag GmbH & Co. KG (Wirtschaft - Beruf - Ethik, 32). Online verfügbar unter https://www.content-select.com/index.php?id=bib_view&ean=9783763954605.

Beckmann, Jan P. (2015): Wilhelm von Ockham. Orig.-Ausg., 3. Aufl., unveränd. Nachdruck. München: Beck (C. H. Beck Paperback, 533).

Benner, Patricia E.; Wengenroth, Martin; Staudacher, Diana (Hg.) (2012): Stufen zur Pflegekompetenz. From Novice to Expert. 2., vollst. überarb. und erg. Aufl. Bern: Huber (Programmbereich Pflege). Online verfügbar unter http://www.socialnet.de/rezensionen/isbn.php?isbn=978-3-456-85020-7.

Bensch, Sandra (2018b): „Dafür haben wir doch keine Zeit“ – Wissenszirkulation wider die Umstände. In: *PADUA* 13 (5), S. 321–327. DOI: 10.1024/1861-6186/a000457.

Bensch, Sandra (2018a): Krankenhauspflege: Eine Pflege ohne Plan? In: *PADUA* 13 (5), S. 315–320. DOI: 10.1024/1861-6186/a000456.

Bensch, Sandra (2013): Konstruktvalidität der Module "Mobilität" und "Kognitive und kommunikative Fähigkeiten" des Neuen Begutachtungsassessments zur Feststellung von Pflegebedürftigkeit. Zugl: Vallendar, Philos.-Theol. Hochsch., Diss., 2011. 1. Aufl. Hungen: Hpsmedia (Reihe Pflegewissenschaft).

Bohnse-Rohmann, Mathias; Burchert, Heiko (Hg.) (2011): Neue Bildungskonzepte für das Gesundheitswesen. Bonn, Bielefeld: Bundesinstitut für Berufsbildung; Bertelsmann (Berichte zur beruflichen Bildung).

Bortz, Jürgen; Lienert, Gustav A. (2003): Kurzgefasste Statistik für die klinische Forschung. Leitfaden für die verteilungsfreie Analyse kleiner Stichproben. 2., aktualisierte und bearb. Aufl. Heidelberg: Springer (Springer-Lehrbuch).

Bortz, Jürgen; Schuster, Christof (2010): Statistik für Human- und Sozialwissenschaftler. 7., vollständig überarbeitete und erweiterte Auflage. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg (Springer-Lehrbuch). Online verfügbar unter http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10448295.

Brandenburg, Hermann; Panfil, Eva-Maria; Mayer, Herbert (Hg.) (2007): Pflegewissenschaft 2. Lehr- und Arbeitsbuch zur Einführung in die Methoden der Pflegeforschung. Unter Mitarbeit von Martina Boger, Anneke de Jong, Heleen Prakke, Martin Schnell und Peter Weschler. 1. Aufl. Bern: Hans Huber.

Brater, Michael (2016): Was sind "Kompetenzen" und wieso können sie für Pflegende wichtig sein? What are "Competences" and why could they be important for Nurses? In: *Pflege & Gesellschaft* 21 (3), 197-213, zuletzt geprüft am 03.04.2019.

Brockhaus (Hg.) (1999): Brockhaus-Enzyklopädie. 19., völlig neu bearb. Aufl. Mannheim: F.A. Brockhaus GmbH (12).

Bruch, Dominik; Dichter, Martin N. (2018): Wissenszirkulation kann so einfach sein. In: *PADUA* 13 (5), S. 329–335. DOI: 10.1024/1861-6186/a000458.

Brühl, Albert (2016): Testtheorie und Instrumentenentwicklung. Zweiter Tag 2016. Philosophisch-Theologische-Hochschule. Lehrstuhl für Statistik und standardisierte Verfahren der Pflegeforschung. Vallendar, 16.11.2016.

Bühner, Markus (2011): Einführung in die Test- und Fragebogenkonstruktion. 3.Auflage. München: PEARSON.

Bundesministerium der Justiz und für Verbraucherschutz; Bundesamt für Justiz (10.11.2003a): Ausbildungs- und Prüfungsverordnung für die Berufe in der Krankenpflege. (KrPflAPrV), vom aufgehoben. Fundstelle: Internet. In: *Gesetze im Internet*. Online verfügbar unter www.gesetze-im-internet.de, zuletzt geprüft am 22.01.2019.

Bundesministerium der Justiz und für Verbraucherschutz; Bundesamt für Justiz (16.07.2003b): Gesetz über die Berufe in der Krankenpflege (Krankenpflegegesetz). KrPflG, vom aufgehoben. Fundstelle: Internet, BGBl. I S.1442; BGBl. I S.2581. Online verfügbar unter www.gesetze-im-Internet.de, zuletzt geprüft am 22.01.2019.

Bundesministerium für Bildung und Forschung (2012): Berufliche Kompetenzen sichtbar machen. Die Forschungsinitiative ASCOT. Bildung. Ideen zünden!, zuletzt geprüft am 24.02.2019.

Bundesministerium für Bildung und Forschung (2014): Impulse für die Ausbildung der Zukunft. Innovative Kompetenzmessung in einer dynamischen Arbeitswelt. Bonn, zuletzt geprüft am 22.02.2019.

Bundesministerium für Bildung und Forschung (2015): Technologiebasierte Kompetenzmessung in der beruflichen Bildung (ASCOT). Ergebnisse und Bedeutung für Politik und Praxis, zuletzt geprüft am 22.02.2019.

Bundesministerium für Gesundheit (06.12.2018): Glossar - Pflegeberufegesetz. PflBG. Fundstelle: Internet. Online verfügbar unter https://www.bundesgesundheitsministerium.de/service/begriffe-von-a-z/p/pflegeberufegesetz.html, zuletzt geprüft am 01.02.2019.

Bundesministerium für Gesundheit; Bundesministerium für Familie, Senioren, Frauen und Jugend (2017): Gesetz zur Reform der Pflegeberufe - Pflegeberufereformgesetz. PflBRefG, vom 17.07.2017. Fundstelle: Internet. In: *Bundesgesetzblatt online* (Teil I Nr.49), S. 2581–2614. Online verfügbar unter http://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl117s2581.pdf, zuletzt geprüft am 22.01.2019.

Bundesministerium für Gesundheit; Bundesministerium für Familie, Senioren, Frauen und Jugend (2018): Ausbildungs- und Prüfungsverordnung für die Pflegeberufe (Pflegeberufe-Ausbildungs- und Prüfungsverordnung). PflAPrV, vom 02.10.2018. Fundstelle: Internet. In: *Bundesgesetzblatt online* 2018 (Teil I Nr.34), S. 1572–1621. Online verfügbar unter http://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl118s1572.pdf, zuletzt geprüft am 22.01.2019.

Campbell, Donald T. (1957): Factors relevant to the validity of experiments in social settings. In: *Psychological bulletin* 54 (4), S. 297–312.

Chomsky, Noam (Hg.) (1970): Aspekte der Syntax-Theorie. Berlin: Akademie.

Chomsky, Noam (2017): Sprache und Geist. Mit einem Anhang: Linguistik und Politik. 12. Auflage. Frankfurt am Main: Suhrkamp (Suhrkamp-Taschenbuch Wissenschaft, 19).

Darmann-Finck, Ingrid (2011): Kompetenzmessung im Berufsfeld Pflege zwischen Subjektivität und Irrelevanz. In: Direktorium des Instituts für Public Health und Pflegeforschung (Hg.): IM FOKUS. Kompetenzmessung in der Pflegeausbildung. ipp info 06 (09). Bremen: Druckerei Peter von Kölln, ohne Seitenangabe.

Darmann-Finck, Ingrid; Einig, Claudia (2019): Curriculumentwicklung für interprofessionelles Lernen, Lehren und Arbeiten. In: Michael Ewers, Elise Paradis und Doreen Herinek (Hg.): Interprofessionell Lernen, Lehren und Arbeiten. Gesundheits- und Sozialprofessionen auf dem Weg zu kooperativer Praxis. Weinheim: Juventa Verlag ein Imprint der Julius Beltz, S. 85–101.

Darmann-Finck, Ingrid; Glissmann, Gerlinde (2011): Kompetenzdiagnostik im Berufsfeld Pflege. In: *Pflege* 24 (3), S. 195–204. DOI: 10.1024/1012-5302/a000120.

Darmann-Finck, Ingrid; Reuschenbach, Bernd (2013): Entwicklungsstand der Kompetenzmessung im Berufsfeld Pflege. Current state of competence assessment in nursing. In: *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen (ZEFQ)* (107), S. 23–29. Online verfügbar unter www.sciencedirect.com.

Deutsche Gesellschaft für Pflegewissenschaft e.V. (2016): Ethikkodex Pflegeforschung der Deutschen Gesellschaft für Pflegewissenschaft. Unter Mitarbeit von Renate Stemmer und Sabine Bartholomeyczik. Deutsche Gesellschaft für Pflegewissenschaft. Online verfügbar unter https://dg-pflegewissenschaft.de/wp-content/uploads/2017/05/Ethikkodex-Pflegeforschung-DGP-Logo-2017-05-25.pdf, zuletzt geprüft am 12.04.2019.

Deutsches Institut für Internationale Pädagogische Forschung - DIPF (2017): Entwicklung und Erprobung von technologieorientierten Messinstrumenten. Unter Mitarbeit von Johannes Hartig. Liebniz-Institut für Bildungsforschung und Bildungsinformation. Online verfügbar unter https://www.dipf.de/de/forschung/projektarchiv/tema, zuletzt aktualisiert am 04.08.2017, zuletzt geprüft am 01.04.2019.

Dietze, Katharina (2011): Schlüsselkompetenzen in der beruflichen Ausbildung zur Gesundheits- und Krankenpflege - Eine empirische Analyse. In: Mathias Bohnse-Rohmann und Heiko Burchert (Hg.): Neue Bildungskonzepte für das Gesundheitswesen. Bonn, Bielefeld: Bundesinstitut für Berufsbildung; Bertelsmann (Berichte zur beruflichen Bildung), S. 133–147.

Dietzen, Agnes; Nickolaus, Reinhold; Rammstedt, Beatrice; Weiß, Reinhold (Hg.) (2016): Kompetenzorientierung. Berufliche Kompetenzen entwickeln, messen und anerkennen. Bundesinstitut für Berufsbildung. 1. Auflage. Bielefeld: W. Bertelsmann Verlag GmbH & Co. KG (Berichte zur beruflichen Bildung). Online verfügbar unter http://www.content-select.com/index.php?id= bib_view&ean= 978376 3956982.

Direktorium des Instituts für Public Health und Pflegeforschung (Hg.) (2011): IM FOKUS. Kompetenzmessung in der Pflegeausbildung. *ipp info* 06 (09). Bremen: Druckerei Peter von Kölln.

Docherty, Vincent J. (2005): Langenscheidt Power Dictionary Englisch. Englisch-Deutsch: Deutsch-Englisch. völlige neubearb., 8. [Dr.]. Berlin, München: Langenscheidt.

Döring, Nicola; Bortz, Jürgen (2016a): Qualitätskriterien in der empirischen Sozialforschung. In: Nicola Döring und Bortz, Jürgen (2016b) (Hg.): Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. Unter Mitarbeit von Sandra Pöschl. 5. vollständig überarbeitete, aktualisierte und erweiterte Auflage. Berlin, Heidelberg: Springer (Springer-Lehrbuch), S. 81–119.

Döring, Nicola; Bortz, Jürgen (2016b) (Hg.): Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. Unter Mitarbeit von Sandra Pöschl. 5. vollständig überarbeitete, aktualisierte und erweiterte Auflage. Berlin, Heidelberg: Springer (Springer-Lehrbuch). Online verfügbar unter http://dx.doi.org/10.1007/978-3-642-41089-5.

Döring, Nicola; Bortz, Jürgen (2016c): Forschungs- und Wissenschaftsethik. In: Nicola Döring und Bortz, Jürgen (2016b) (Hg.): Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. Unter Mitarbeit von Sandra Pöschl. 5. vollständig überarbeitete, aktualisierte und erweiterte Auflage. Berlin, Heidelberg: Springer (Springer-Lehrbuch), S. 121–139.

Döring, Ottmar; Möllers, Michaela; Schöpf, Nicolas (2014): Qualitätssicherung durch Kompetenzmessung. Technologiebasierte Kompetenzerfassung am Beispiel der Altenpflege. In: *Wirtschaft & Beruf. Zeitschrift für berufliche Bildung* 66 (3), S. 34–43. Online verfügbar unter https://www.researchgate.net/publication/289779637_Qualitatssicherung_durch_Kompetenzmessung_Technologiebasierte_Kompetenzerfassung_am_Beispiel_der_Altenpflege, zuletzt geprüft am 04.03.2019.

Döring, Ottmar; Weyland, Ulrike; Wittmann, Eveline; Nauerth, Annette; Hartig, Johannes; Kaspar, Roman; Möllers, Michaela; Rechenbach, Simone; Simon, Julia; Worofka, Iberé; Kraus, Kristina (2016b): Technologiebasierte Messung beruflicher Handlungskompetenz in der Pflege älterer Menschen: Kompetenzmodellierung und Testverfahrensentwicklung. In: Agnes Dietzen, Reinhold Nickolaus, Beatrice Rammstedt und Reinhold Weiß (Hg.): Kompetenzorientierung. Berufliche Kompetenzen entwickeln, messen und anerkennen. 1. Auflage. Bielefeld: W. Bertelsmann Verlag GmbH & Co. KG (Berichte zur beruflichen Bildung), S. 117–132.

Döring, Ottmar; Wittmann, Eveline; Hartig, Johannes; Weyland, Ulrike; Nauerth, Annette; Kaspar, Roman; Möllers, Michaela; Rechenbach, Simone; Simon, Julia; Worofka, Iberé (2014): Verbundprojekt: Entwicklung und Erprobung von technologieorientierten Messinstrumenten zur Feststellung der beruflichen Handlungskompetenz in der Pflege älterer Menschen (TEMA). Abschlussbericht. Forschungsinstitut Betriebliche Bildung (f-bb) gGmbH; Otto-Friedrich-Universität Bamberg, Sozial- und wirtschaftswissenschaftliche Fakultät; Deutsches Institut für Internationale Pädagogische Forschung; FH Bielefeld, Lehreinheit Pflege und Gesundheit. bisher unveröffentlicht - mit freundlicher Unterstützung des TEMA-Projektteams zur Einsicht für das Projekt in der Gesundheits- und Krankenpflege zur Verfügung gestellt!, zuletzt geprüft am 28.01.2019.

Döring, Ottmar; Wittmann, Eveline; Weyland, Ulrike; Nauerth, Annette; Hartig, Johannes; Kaspar, Roman; Möllers, Michaela; Rechenbach, Simone; Simon, Julia; Worofka, Iberé; Kraus, Kristina (2016a): Technologiebasierte Messung von beruflichen Kompetenzen für die Pflege älterer Menschen: berufsfachliche Kompetenzen, allgemeine Kompetenzen und Kontextfaktoren (TEMA). In: Klaus Beck, Margarete Landenberger und Fritz Oser (Hg.): Technologiebasierte Kompetenzmessung in der beruflichen Bildung. Ergebnisse aus der BMBF-Förderinitiative ASCOT. 1. Auflage. Bielefeld: W. Bertelsmann Verlag GmbH & Co. KG (Wirtschaft - Beruf - Ethik, 32), S. 243–264, zuletzt geprüft am 07.12.2018.

Dudenredaktion (Hg.) (2015): Der Duden - Das Fremdwörterbuch - Band 5. In zwölf Bänden; das Standardwerk zur deutschen Sprache. Unter Mitarbeit von Kathrin Kunkel-Razum. 11., vollst. überarb. und aktualisierte Aufl. 12 Bände. Berlin: Dudenverlag. Online verfügbar unter http://www.vlb.de/GetBlob.aspx?strDisposition=a&strIsbn=9783411040612.

Dütthorn, Nadin (2014): Pflegespezifische Kompetenzen im Europäischen Bildungsraum. Eine empirische Studie in den Ländern Schottland, Schweiz und Deutschland. Dissertation. Göttingen: V & R unipress GmbH; Universitätsverlag Osnabrück (Pflegewissenschaft und Pflegebildung, Band 8).

Dütthorn, Nadin (2016): Pflegespezifische Kompetenzen im europäischen Bildungsraum. Spagat zwischen bildungspolitischer Zielsetzung und berufspädagogischer Handlungspraxis. 3. Berlinder Pflegekonferenz am 08.11.2016, zuletzt geprüft am 05.08.2019.

Eid, Michael; Schmidt, Katharina (2014): Testtheorie und Testkonstruktion. Göttingen: Hogrefe (Bachelorstudium Psychologie).

Epskamp, Sascha; Stuber, Simon; Nak, Jason; Veenman, Myrthe (2019): Package 'semPlot'. Path Diagrams and Visual Analysis of Various SEM Packages Output. R package version 1.1.1. Online verfügbar unter https://cran.r-project.org/web/packages/semPlot/semPlot.pdf.

Erpenbeck, John; Rosenstiel, Lutz von; Grote, Sven; Sauter, Werner (Hg.) (2017): Handbuch Kompetenzmessung. Erkennen, verstehen und bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis. Fachverlag für Wirtschafts- und Steuerrecht Schäffer. 3., überarbeitete und erweiterte Auflage. Stuttgart: Schäffer-Poeschel Verlag.

Evers, Thomas (2011): Zur Konstruktion von Curricula in Gesundheitsberufen. In: Mathias Bohnse-Rohmann und Heiko Burchert (Hg.): Neue Bildungskonzepte für das Gesundheitswesen. Bonn, Bielefeld: Bundesinstitut für Berufsbildung; Bertelsmann (Berichte zur beruflichen Bildung), S. 49–62.

Ewers, Michael; Paradis, Elise; Herinek, Doreen (Hg.) (2019): Interprofessionell Lernen, Lehren und Arbeiten. Gesundheits- und Sozialprofessionen auf dem Weg zu kooperativer Praxis. Juventa Verlag. Weinheim: Juventa Verlag ein Imprint der Julius Beltz.

Fachkommission (2019): Rahmenpläne der Fachkommission nach §53 PflBG. Rahmenlehrpläne für den theoretischen und praktischen Unterricht. Rahmenausbildungspläne für die praktische Ausbildung. Unter Mitarbeit von Rainer Ammende, Frank Arens, Ingrid Darmann-Finck, Roswitha Ertl-Schmuck, Brigitte v. Germeten-Ortmann, Gertrud Hundenborn et al. Online verfügbar unter https://www.bibb.de/de/86562.php, zuletzt geprüft am 01.08.2019.

Franken, Georg (2010): Konstruktvalidität der Subskala "Kognitive und kommunikative Fähigkeiten" des Neuen Begutachtungsassessments zur Feststellung von Pflegebedürftigkeit (NBA). Masterarbeit. Philosophisch-Theologische-Hochschule, Vallendar, zuletzt geprüft am 28.01.2019.

Frei Blatter, Vreni u. Ochsner Oberarzbacher, Liliane (2008): Der Einsatz von Simulationspatienten (SP) in der Pflegeausbildung eingebettet in die Skillslab-Methode. In: Gerda Nussbaumer und Christine von Reibnitz (Hg.): Innovatives Lehren und Lernen. Konzepte für die Aus- und Weiterbildung von Pflege- und Gesundheitsberufen. Bern: Hans Huber (Pflegepädagogik), S. 113–133, zuletzt geprüft am 10.12.2018.

Freitag, Walburga; Hartmann, Ernst; Loroff, Claudia; Stamm-Riemer, Ida; Völk, Daniel; Buhr, Regina (Hg.) (2011): Gestaltungsfeld Anrechnung. Hochschulische und berufliche Bildung im Wandel. Münster: Waxmann.

Funke, Joachim; Reuschenbach, Bernd (2011): Einsatz technischer Mittel in der psychologischen Diagnostik. In: Lutz F. Hornke, Manfred Amelang und Martin Kersting (Hg.): Leistungs-, Intelligenz- und Verhaltensdiagnostik. [Vollst. Neuausg.]. Göttingen: Hogrefe (Enzyklopädie der Psychologie, Themenbereich B Methodologie und Methoden; Ser. 2 Psychologische Diagnostik; Bd. 3), S. 595–631.

Georg-Elias-Müller-Institut für Psychologie (2018): CFA. Konfirmatorische Faktorenanalyse - Confirmatory Factor Analysis (CFA). Unter Mitarbeit von Uwe Mattler. Hg. v. Georg-August-Universität Göttingen. Georg-Elias-Müller-Institut für Psychologie. Göttingen. Online verfügbar unter https://www.psych.uni-goettingen.de/de/it/team/zezula/courses/multivariate-verfahren-2011/verfahren/cfa, zuletzt geprüft am 05.11.2018.

Gerlach, Anke (2017): Zur Frage der professionellen Identität akademisch Qualifizierter in der Pflege. Ergebnisse einer empirischen Studie. In: Tobias Sander und Sarah Dangendorf (Hg.): Akademisierung der Pflege. Berufliche Identitäten und Professionalisierungspotenziale im Vergleich zu Sozialer Arbeit und Frühpädagogik. 1. Auflage. Weinheim, Basel: Beltz Juventa, S. 60–82.

Hahnel, Carolin (2018): Einführung in die Item Response Theory mit R. Workshop-Power Point Präsentation. Frühjahrsakademie zu Methoden der empirischen Bildungsforschung. DIPF Frankfurt. Institut zur Qualitätsentwicklung im Bildungswesen. Berlin, 15.03.2018, zuletzt geprüft am 22.08.2018.

Hähnel, Martin (2015): Das Ethos der Ethik. Zur Anthropologie der Tugend. Wiesbaden: Springer Fachmedien Wiesbaden.

Hatzinger, Reinhold; Hornik, Kurt; Nagel, Herbert; Maier, Marco J. (2014): R. Einführung durch angewandte Statistik. 2., aktualisierte Aufl. Hallbergmoos: PEARSON (Always learning).

Hermeneit, Elke (2008): Was testet PISA? Norderstedt: Grin Verlag.

Holoch, Elisabeth (2007): Integrative Pflegeausbildung: Das Stuttgarter Modell. Kompetenzerwerb und Transferlernen: Schwerpunkte und Herausforderungen im Revisionsprozess. In: *Die Zeitschrift für Pflegewissenschaft - PRINTERNET* 9 (6), 379-386, zuletzt geprüft am 03.04.2019.

Hornke, Lutz F.; Amelang, Manfred; Kersting, Martin (Hg.) (2011): Leistungs-, Intelligenz- und Verhaltensdiagnostik. [Vollst. Neuausg.]. Göttingen: Hogrefe (Enzyklopädie der Psychologie, Themenbereich B Methodologie und Methoden; Ser. 2 Psychologische Diagnostik; Bd. 3).

Hülsken-Gießler, Manfred; Brinker-Meyendrisch, Elfriede; Keogh, Johann; Muths, Sabine; Sieger, Margot; Stemmer, Renate; Stöcker, Gertrud; Walter, Anja (2010): Kerncurriculum Pflegewissenschaft für pflegebezogene Studiengänge - eine Initiative zur Weiterentwicklung der hochschulischen Pflegebildung in Deutschland. In: *Pflege & Gesellschaft* 15. (3), S. 216–236.

Hundenborn, Gertrud (2016): Fallorientierte Didaktik in der Pflege. Grundlagen und Beispiele für Ausbildung und Prüfung. Print-on-Demand Ausgabe der 1. Auflage. München: Urban & Fischer in Elsevier.

Hundenborn, Gertrud (2019): Pflegeausbildung kompetenzorientiert gestalten. Kompetenzverständnis und Kompetenzsystematik im Kontext der Pflegeberufereform. In: *PflegeLeben. Pflege-Werte- Zukunft - das Magazin für Mitglieder des Katholischen Pflegeverbandes e.V.* (1), S. 5–11.

Hussy, Walter (2013): Psychologie als eine empirische Wissenschaft. In: Walter Hussy, Gerald Echterhoff und Margrit Schreier (Hg.): Forschungsmethoden in Psychologie und Sozialwissenschaften für Bachelor. 2., überarbeitete Auflage. Berlin, Heidelberg: Springer, S. 1–49.

Hussy, Walter; Echterhoff, Gerald; Schreier, Margrit (Hg.) (2013): Forschungsmethoden in Psychologie und Sozialwissenschaften für Bachelor. 2., überarbeitete Auflage. Berlin, Heidelberg: Springer.

Irribarra, David Torres & Freund, Rebecca (2016): Package 'WrightMap'. IRT Item-Person Map with 'ConQuest' Integration. R package version 1.2.1. Online verfügbar unter https://cran.r-project.org/web/packages/WrightMap/WrightMap.pdf.

Kaiser, Franz-Josef (1985): Fallstudie. In: Dieter Lenzen (Hg.): Enzyklopädie Erziehungswissenschaft. Band 4: Methoden und Medien der Erziehung und des Unterrichts, Bd. 4. 12 Bände. Stuttgart: Klett-Cotta (4), S. 440–444.

Kälble, Karl (2017): Zur Professionalisierung der Pflege in Deutschland. Stand und Perspektive. In: Tobias Sander und Sarah Dangendorf (Hg.): Akademisierung der Pflege. Berufliche Identitäten und Professionalisierungspotenziale im Vergleich zu Sozialer Arbeit und Frühpädagogik. 1. Auflage. Weinheim, Basel: Beltz Juventa, S. 27–58.

Kaspar, Roman (05.05.2018): TEMA Parcels CFA - Skript. Zur Verfügung gestellt vom Forschungsprojekt "TEMA". nicht veröffentlicht.

Kaspar, Roman; Döring, Ottmar; Wittmann, Eveline; Hartig, Johannes; Weyland, Ulrike; Nauerth, Annette; Möllers, Michaela; Rechenbach, Simone; Simon, Julia; Worofka, Iberé (2016): Competencies in Geriatric Nursing. Empirical Evidence from a Computer-Based Large-Scale Assessment Calibration Study. In: *Vocations and Learning* 9 (2), 185–206 bzw. 1-33. DOI: 10.1007/s12186-015-9147-y.

Kaspar, Roman; Hartig, Johannes (2016): Emotional competencies in geriatric nursing: empirical evidence from a computer based large scale assessment calibration study. In: *Advances in health sciences education : theory and practice* 21 (1), S. 105–119. DOI: 10.1007/s10459-015-9616-y.

Kerngruppe Curriculum (2006): Integrative Pflegeausbildung: Das Stuttgarter Modell ©. Pflegeberuflicher und pädagogischer Begründungsrahmen. 1. Aufl. Braunschweig: Winklers (Integrative Pflegeausbildung, das Stuttgarter Modell).

Knappich, Thomas (2017): Messung des mathematischen Textverständnisses von Altenpflegeschülern zu Beginn der Pflegeausbildung. Eine empirische Untersuchung an privaten und öffentlichen Altenpflegeschulen. Inaugural-Dissertation. Online verfügbar unter https://kidoks.bsz-bw.de/frontdoor/deliver/index/docId/1020/file/Knappich_Dissertation_final.pdf.

Knigge-Demal, Barbara; Eylmann, Constanze (2006): Kompetenzorientierte Prüfungsgestaltung Teil 1 - Anhand von Fallbeispielen - Berichte aus Lehre und Forschung. Hg. v. Fachhochschule Bielefeld, Fachbereich Pflege und Gesundheit. Bielefeld (18).

Knigge-Demal, Barbara; Eylmann, Constanze; Hundenborn, Gertrud (2013): Anforderungs- und Qualifikationsrahmen für den Beschäftigungsbereich der Pflege und persönlichen Assistenz älterer Menschen. im Rahmen des Projektes "Erprobung des Entwurfs eines Qualifikationsrahmens für den Beschäftigungsbereich der Pflege, Unterstützung und Betreuung älterer Menschen". Hg. v. Fachhochschule Bielefeld und Deutsches Institut für angewandte Pflegeforschung e.V. Online verfügbar unter https://www.dip.de/fileadmin/data/pdf/projekte/01Anforderungs_und_Qualifikationsrahmen_09_2013.pdf, zuletzt geprüft am 03.03.2019.

Koller, Ingrid; Alexandrowicz, Rainer; Hatzinger, Reinhold (2012): Das Rasch-Modell in der Praxis. Eine Einführung mit eRm. 1. Aufl. Wien: facultas.wuv (utb-studi-e-book, 3786). Online verfügbar unter http://www.utb-studi-e-book.de/9783838537863.

Koller, Martina M. (2014): Statistik für Pflege- und Gesundheitsberufe. Wien: Facultas (Manual).

König, Peter; Lindwedel-Reime, Ulrike (2019): Multiprofessionelles Skills-Lab. Hochschule Furtwangen. Online verfügbar unter https://www.hs-furtwangen.de/fakultaeten/gesundheit-sicherheit-gesellschaft/labore/multiprofessionelles-skills-lab/.

Kreienbrock, Lothar; Pigeot, Iris; Ahrens, Wolfgang (2012): Epidemiologische Methoden. 5. Aufl. Heidelberg: Spektrum Akademischer Verlag. Online verfügbar unter http://dx.doi.org/10.1007/978-3-8274-2334-4.

Kunina-Habenicht, Olga (16.06.2018b): Einführung in die Konstruktion von Kompetenztests und Datenauswertung in SPSS: Teil 2. Workshop auf der Sommerakademie 2018 an der PH Ludwigsburg. Graduiertenakademie Pädagogische Hochschulen Karlsruhe. PH Ludwigsburg, 16.06.2018b. Online verfügbar unter http://www.ph-bw.de/graph, zuletzt geprüft am 15.05.2019.

Kunina-Habenicht, Olga (15.06.2018a): Einführung in die Konstruktion von Kompetenztests und Datenauswertung in SPSS: Teil1. Workshop auf der Sommerakademie 2018 an der PH Ludwigsburg. Graduiertenakademie Pädagogische Hochschulen Karlsruhe. PH Ludwigsburg, 15.06.2018a. Online verfügbar unter http://www.ph-bw.de/graph, zuletzt geprüft am 20.08.2018.

Lamprianou, Iasonas (2020): Question for my Disputation and Dissertation in Germany. Gemeinsamer Mailaustausch am 27.06.2020.

Lamprianou, Iasonas (2020): Applying the Rasch Model in Social Sciences Using R and BlueSky Statistics. New York and London: Taylor & Francis.

Lenzen, Dieter (Hg.) (1985): Enzyklopädie Erziehungswissenschaft. Band 4: Methoden und Medien der Erziehung und des Unterrichts. 12 Bände. Stuttgart: Klett-Cotta (4).

Linacre, John Michael (1994): Sample Size and Item Calibration (or Person Measure) Stability. In: Journal of Rasch Measurement Transactions 7 (4), S. 328. Online verfügbar unter http://www.rasch.org/rmt/rmt74m.htm, zuletzt geprüft am 05.10.2020.

Mai, Markus (2010): Das Sturzrisiko von Patienten im Krankenhaus. Entwicklung eines konstruktvaliden Sturzrisikoeinschätzungsinstruments unter dem Einsatz von Modellen aus dem Bereich der probabilistischen Testtheorie. 1. Aufl. München: Dr. Hut.

Mair, Patrick; Hatzinger, Reinhold; Maier, Marco J.; Rusch, Thomas; Debelak, Rudolf (2020): Package ‚eRm'. Extended Rasch Modeling. R package version 1.0-1. Online verfügbar unter https://cran.r-project.org/web/packages/eRm/eRm.pdf

Masters, Geoff N. (1982): A rasch model for partial credit scoring. In: *Psychometrika* 47 (2), S. 149–174. DOI: 10.1007/BF02296272.

Minks; Karl-Heinz (2011): Lebenslanges Lernen und Durchlässigkeiet - demographische und sozioökonomische Herausforderungen. In: Walburga Freitag, Ernst Hartmann, Claudia Loroff, Ida Stamm-Riemer, Daniel Völk und Regina Buhr (Hg.): Gestaltungsfeld Anrechnung. Hochschulische und berufliche Bildung im Wandel. Münster: Waxmann, S. 21–34.

Monteverde, Settimio (Hg.) (2012): Handbuch Pflegeethik. Ethisch denken und handeln in den Praxisfeldern der Pflege. 1. Aufl. Stuttgart: Kohlhammer (Pflegepraxis).

Monteverde, Settimio; Kesselring, Annemarie (2012): Forschung in der Pflege und Ethik. In: Settimio Monteverde (Hg.): Handbuch Pflegeethik. Ethisch denken und handeln in den Praxisfeldern der Pflege. 1. Aufl. Stuttgart: Kohlhammer (Pflegepraxis), S. 138–155.

Moosbrugger, Helfried (2012): Item-Response-Theorie (IRT). In: Helfried Moosbrugger und Augustin Kelava (Hg.): Testtheorie und Fragebogenkonstruktion. 2., aktualisierte und überarbeitete Auflage. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg (Springer-Lehrbuch), S. 228–271.

Moosbrugger, Helfried; Kelava, Augustin (Hg.) (2012): Testtheorie und Fragebogenkonstruktion. 2., aktualisierte und überarbeitete Auflage. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg (Springer-Lehrbuch). Online verfügbar unter http://dx.doi.org/10.1007/978-3-642-20072-4.

Moosbrugger, Helfried; Schermelleh-Engel, Karin (2012): Exploratorische (EFA) und Konfirmatorische Faktorenanalyse (CFA). In: Helfried Moosbrugger und Augustin Kelava (Hg.): Testtheorie und Fragebogenkonstruktion. 2., aktualisierte und überarbeitete Auflage. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg (Springer-Lehrbuch), S. 326–343.

Muraki, Eiji (1992): A Generalized Partial Credit Model: Application of an EM Algorithm. In: *Applied Psychological Measurement* 16 (2), 159-176. DOI: 10.1002/j.2333-8504.1992.tb01436.x.

Muthén, Linda K.; Muthén, Bengt O. (1998-2012): Mplus. Statistical Analysis With Latent Variables. User´s Guide. Seventh Edition. Los Angeles: CA: Muthén & Muthén, zuletzt geprüft am 15.05.2019.

Nauerth, Annette; Walkenhorst, Ursula (2009): Kompetenzenwicklung im Gesundheits- und Sozialbereich. In: Ursula Walkenhorst, Annette Nauerth, Inge Bergmann-Tyacke und Kordula Marzinzik (Hg.): Kompetenzentwicklung im Gesundheits- und Sozialbereich. Bielefeld: UVW, Webler (Hochschulwesen Wissenschaft und Praxis (Bielefeld, Germany), N.F., 17), S. 9–22.

Nickolaus, Reinhold; Walker, Felix (2016): Kompetenzorientierung in der beruflichen Bildung. In: Agnes Dietzen, Reinhold Nickolaus, Beatrice Rammstedt und Reinhold Weiß (Hg.): Kompetenzorientierung. Berufliche Kompetenzen entwickeln, messen und anerkennen. 1. Auflage. Bielefeld: W. Bertelsmann Verlag GmbH & Co. KG (Berichte zur beruflichen Bildung), S. 7–27.

Nohl, Martina (2009): Entwicklung von Übergangskompetenz in der Laufbahnberatung. Zugl.: Darmstadt, Techn. Univ., Diss., 2009. 1. Aufl. Berlin: Köster.

Nussbaumer, Gerda; Reibnitz, Christine von (Hg.) (2008): Innovatives Lehren und Lernen. Konzepte für die Aus- und Weiterbildung von Pflege- und Gesundheitsberufen. Bern: Hans Huber (Pflegepädagogik).

OECD (2018): PISA - Internationale Schulleistungsstudie der OECD. Online verfügbar unter http://www.oecd.org/berlin/themen/pisa-studie/, zuletzt geprüft am 13.01.2019.

Partchev, Ivailo; Maris, Gunter; Hattori, Tamaki (2017): Package 'irtoys'. A Collection of Functions Related to Item Response Theory (IRT). Online verfügbar unter https://cran.r-project.org/web/packages/irtoys/irtoys.pdf.

Porst, Rolf (2014): Fragebogen. Ein Arbeitsbuch. 4., erweiterte Auflage. Wiesbaden: Springer VS (Lehrbuch).

Pospeschill, Markus (2010): Testtheorie, Testkonstruktion, Testevaluation. Mit 77 Fragen zur Wiederholung. München, Basel: Ernst Reinhardt Verlag (utb.de-Bachelor-Bibliothek, 3431).

R Core Team; Bivand, Roger, Carey, Vincent J.; DebRoy, Saikat; Eglen, Stephen; Guha, Rajarshi; Herbrandt, Swetlana; Lewin-Koh, Nicholas; Myatt, Mark; Pfaff, Ben; Quistorff, Brian; Warmerdam, Frank; Weigand, Stephan; Free Software Foundation, Inc. (2018): Package 'foreign'. Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase', …. R package version 0.8-71. Online verfügbar unter https://cran.r-project.org/web/packages/foreign/foreign.pdf.

Raven, Uwe (1989): Professionelle Sozialisation und Moralentwicklung. Zum Berufsethos von Medizinern. Wiesbaden: Deutscher Universitätsverlag. Online verfügbar unter http://dx.doi.org/10.1007/978-3-663-14566-0.

Reibnitz, Christine v. (2008): Kompetenzentwicklung und -förderung in Gesundheitsberufen. In: Gerda Nussbaumer und Christine von Reibnitz (Hg.): Innovatives Lehren und Lernen. Konzepte für die Aus- und Weiterbildung von Pflege- und Gesundheitsberufen. Bern: Hans Huber (Pflegepädagogik), S. 31–66, zuletzt geprüft am 10.12.2018.

Reuschenbach, Bernd (2008): Einfluss von Expertise auf Problemlösen und Planen im komplexen Handlungsfeld Pflege. Zugl.: Heidelberg, Univ., Diss. Berlin: Logos-Verl.

Ries, Simone (2014b): Lehrergesundheit an Gesundheits- und Krankenpflegeschulen. Masterarbeit. Philosophisch-Theologische-Hochschule, Vallendar.

Ries, Simone (2014a): Forschungspraktikumsbericht. Lehrergesundheit. Philosophisch-Theologische-Hochschule, Vallendar.

Rizopoulos, Dimitris (2018): Package 'ltm'. Latent Trait Models under IRT. R package version 1.1-1. Online verfügbar unter https://cran.r-project.org/web/packages/ltm/ltm.pdf.

Robitzsch, Alexander; Kiefer, Thomas; Wu, Margaret (2019): Package 'TAM'. Test Analysis Modules. R package version 3.1-45. Online verfügbar unter https://cran.r-project.org/web/packages/TAM/TAM.pdf.

Rosseel, Yves; Oberski, Daniel; Byrnes, Jarrett; Vanbrabant, Leonard; Savalei, Victoria; Merkle, Ed; Hallquist, Michael; Rhemtulla, Mijke; Katsikatsou, Myrsini; Barendse, Mariska; Chow, Michael; Jorgensen, Terrence D. (2018): Package 'lavaan'. Latent Variable Analysis. R package version 0.6-3. Online verfügbar unter https://cran.r-project.org/web/packages/lavaan/lavaan.pdf.

Rost, Jürgen (1996): Lehrbuch Testtheorie, Testkonstruktion. 1. Aufl. Bern: Huber (Aus dem Programm Huber). Online verfügbar unter http://www.sub.uni-hamburg.de/ebook/ebook.php?act=b&cid=498.

Rost, Jürgen (1999): Was ist aus dem Rasch-Modell geworden? In: *Psychologische Rundschau* 50 (3), S. 140–156.

Rost, Jürgen (2004): Lehrbuch Testtheorie - Testkonstruktion. 2., vollst. überarb. und erw. Aufl. Bern: Hans Huber (Psychologie Lehrbuch).

Sander, Tobias; Dangendorf, Sarah (Hg.) (2017): Akademisierung der Pflege. Berufliche Identitäten und Professionalisierungspotenziale im Vergleich zu Sozialer Arbeit und Frühpädagogik. 1. Auflage. Weinheim, Basel: Beltz Juventa.

Sauter, Werner; Staudt, Anne-Kathrin (2016): Kompetenzmessung in der Praxis. Mitarbeiterpotenziale erfassen und analysieren. 1. Aufl. 2016. Wiesbaden: Springer Gabler (essentials). Online verfügbar unter http://gbv.eblib.com/patron/FullRecord.aspx?p=4332301.

Schermelleh-Engel, Karin; Werner, Christina S. (2012): Methoden der Reliabilitätsbestimmung. In: Helfried Moosbrugger und Augustin Kelava (Hg.): Testtheorie und Fragebogenkonstruktion. 2., aktualisierte und überarbeitete Auflage. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg (Springer-Lehrbuch), S. 120–141.

Schiffer, Peter (2018): Begutachtung Ethikkommission [DGP]. Formblatt.

Schlegel, Claudia (2018): OSCE - Kompetenzorientiert Prüfen in der Pflegeausbildung. Einführung und Umsetzung von OSCE-Stationen. Mit über 40 Beispielen von OSCE Aufgabenstellungen und Beurteilungskriterien. Berlin, Heidelberg: Springer. Online verfügbar unter https://ebookcentral.proquest.com /lib/gbv/detail.action?docID=5223891.

Schlegel, Claudia; Geering, Alain; Weber, Uwe (2020): Virtuelle Realität verbessert die Wirklichkeit. Ein Gewinn für Lernsettings der Pflegeaus- und Weiterbildung. In: *Pflege Zeitschrift* 73 (1-2), S. 57–60.

Schnell, Martin W. (2007): Pflegeforschungsethik. In: Hermann Brandenburg, Eva-Maria Panfil und Herbert Mayer (Hg.): Pflegewissenschaft 2. Lehr- und Arbeitsbuch zur Einführung in die Methoden der Pflegeforschung. Unter Mitarbeit von Martina Boger, Anneke de Jong, Heleen Prakke, Martin Schnell und Peter Weschler. 1. Aufl. Bern: Hans Huber, S. 165–175.

Schön, Jasmin (2011): Kooperative Pflegeausbildung: Altenpflege und Gesundheits- und Krankenpflege - Ein Beitrag zur Curriculumentwicklung an einer beruflichen Schule in Baden-Württemberg. In: Mathias Bohnse-Rohmann und Heiko Burchert (Hg.): Neue Bildungskonzepte für das Gesundheitswesen. Bonn, Bielefeld: Bundesinstitut für Berufsbildung; Bertelsmann (Berichte zur beruflichen Bildung), S. 79–97.

Schröder, Martina (2010): Konstruktvalidität der Subskala Mobilität des Neuen Begutachtungsassessments für Pflegebedürftigkeit (NBA). Masterarbeit. Philosophisch-Theologische-Hochschule, Vallendar, zuletzt geprüft am 05.11.2018.

Seeber, Susan; Schumann, Matthias; Ketschau, Thilo J.; Rüter, Theresa; Kleinhans, Janne (2016): Modellierung und Messung von Fachkompetenzen Medizinischer Fachangestellter (CoSMed). In: Klaus Beck, Margarete Landenberger und Fritz Oser (Hg.): Technologiebasierte Kompetenzmessung in der beruflichen Bildung. Ergebnisse aus der BMBF-Förderinitiative ASCOT. 1. Auflage. Bielefeld: W. Bertelsmann Verlag GmbH & Co. KG (Wirtschaft - Beruf - Ethik, 32), S. 205–223.

Seifried, Jürgen; Faßhauer, Uwe; Seeber, Susan (Hg.) (2014): Jahrbuch der berufs- und wirtschaftspädagogischen Forschung 2014. Leverkusen: Budrich, Barbara (Schriftenreihe der Sektion Berufs- und Wirtschafts-pädagogik der Deutschen Gesellschaft für Erziehungswissenschaft (DGfE)).

Simon, Julia; Rechenbach, Simone; Kaspar, Roman; Möllers, Michaela; Worofka, Iberé (2015): Entwicklung eines heuristischen Kompetenzmodells für die Pflege älterer Menschen: Strukturelle und inhaltliche Analysen unter Berücksichtigung domänenspezifischer Anforderungen. Hg. v. Ulrike Weyland, Marisa Kaufhold, Annette Nauerth und Elke Rosowski. Online verfügbar unter www.bwpat.de/spezial10/simon_etal_gesundheitsbereich-2015.pdf, zuletzt geprüft am 24.01.2019.

Smith, Sarah A. (2012): Nurse Competence: A Concept Analysis. In: *International Journal of Nursing Knowledge: The Official Journal of NANDA International* 23 (3), S. 172–182.

Sozialministerium Baden-Württemberg; Landesarbeitsgemeinschaft der Lehrerinnen und Lehrer für Pflegeberufe Baden-Württemberg e.V. (2004): Vorläufiger Landeslehrplan Baden-Württemberg für die Ausbildung zur Gesundheits- und Krankenpflegerin oder zum Gesundheits- und Krankenpfleger und zur Gesundheits- und Kinderkrankenpflegerin oder zum Gesundheits- und Kinderkrankenpfleger. o.A.: o.A.

Spiegelhalter, David J.; Best, Nicola G.; Carlin, Bradley P.; v.d. Linde, Angelika (2002): Bayesian Measures of Model Complexity and Fit (with Discussion). In: *Journal of the Royal Statistical Society* 64 (4), S. 583–639.

Strobel, Carolin (2020): Globaler Waldtest. Gemeinsamer Mailaustausch am 01.07.2020.

Strobl, Carolin (2015): Das Rasch-Modell. Eine verständliche Einführung für Studium und Praxis. Hg. v. Bernhard Kittel, Claudia Fantapié Altobelli, Carolin Strobl, Gerhard Tutz, Holger Hinz, Ingwer Borg, et al. Mering: Rainer Hampp Verlag (Sozialwissenschaftliche Forschungsmethoden). Online verfügbar unter https://www.wiso-net.de/document/EBOK,AEBO__9783957101501117.

„TEMA"-Projektgruppe (06.03.2014): Testaufgaben zu Setting 1,2 und 3 und Testaufgaben von COSMED als Vorlage für die Hauptstudie. unveröffentlicht.

Ulrike Weyland; Karin Reiber; Bundesinstitut für Berufsbildung (Hg.) (2017): Entwicklungen und Perspektiven in den Gesundheitsberufen - aktuelle Handlungs- und Forschungsfelder. 1. Aufl. Bielefeld: W. Bertelsmann Verlag (Berichte zur beruflichen Bildung). Online verfügbar unter https://ebookcentral.proquest.com/lib/gbv/detail.action?docID=5107605.

Walkenhorst, Ursula; Nauerth, Annette; Bergmann-Tyacke, Inge; Marzinzik, Kordula (Hg.) (2009): Kompetenzentwicklung im Gesundheits- und Sozialbereich. Bielefeld: UVW, Webler (Hochschulwesen Wissenschaft und Praxis (Bielefeld, Germany), N.F., 17), zuletzt geprüft am 06.01.2019.

Weiber, Rolf; Mühlhaus, Daniel (2014): Strukturgleichungsmodellierung. Eine anwendungsorientierte Einführung in die Kausalanalyse mit Hilfe von AMOS, SmartPLS und SPSS. 2., erw. und korrigierte Aufl. Berlin [u.a.]: Springer Gabler (Springer-Lehrbuch).

Weidner, Frank (2011): Professionelle Pflegepraxis und Gesundheitsförderung. Eine empirische Untersuchung über Voraussetzungen und Perspektiven des beruflichen Handelns in der Krankenpflege. Zugl.: Osnabrück, Univ., Diss., 1994. 3., unverän. Aufl. Frankfurt am Main: Mabuse-Verl. (Mabuse-Verlag Wissenschaft, 22).

Weiß, Reinhold (2016): Vorwort. In: Agnes Dietzen, Reinhold Nickolaus, Beatrice Rammstedt und Reinhold Weiß (Hg.): Kompetenzorientierung. Berufliche Kompetenzen entwickeln, messen und anerkennen. 1. Auflage. Bielefeld: W. Bertelsmann Verlag GmbH & Co. KG (Berichte zur beruflichen Bildung), S. 5–6.

Werner, Christina S.; Schermelleh-Engel, Karin; Gerhard, Carla; Gäde, Jana C. (2016): Strukturgleichungsmodelle. In: Nicola Döring und Bortz, Jürgen (2016b) (Hg.): Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. Unter Mitarbeit von Sandra Pöschl. 5. vollständig überarbeitete, aktualisierte und erweiterte Auflage. Berlin, Heidelberg: Springer (Springer-Lehrbuch), S. 945–973.

Wickham, Hadley; Miller, Evan; RStudio (2019): Package 'haven'. Import and Export 'SPSS', 'Stata' and 'SAS' Files. R package version 2.1.0. Online verfügbar unter https://cran.r-project.org/web/ packages/haven/haven.pdf.

Wickham, Hardley; Hester, Jim Hester; Francois, Romain; R Core Team; RStudio; Jylänki, Jukka; Jørgensen, Mikkel (2018): Package 'readr'. Read Rectangular Text Data. Online verfügbar unter https://cran.r-project.org/web/packages/readr/readr.pdf.

Winther, Esther (2010): Kompetenzmessung in der beruflichen Bildung. Bielefeld: W. Bertelsmann Verlag. Online verfügbar unter http://www.content-select.com/index.php?id=bib_view&ean=9783763947270.

Wittmann, Eveline; Kaspar, Roman; Döring, Ottmar (2017): Berufsfachliche Handlungskompetenz in der unmittelbar klientenbezogenen Pflege: eindimensional oder mehrdimensional? In: Ulrike Weyland, Karin Reiber und Bundesinstitut für Berufsbildung (Hg.): Entwicklungen und Perspektiven in den Gesundheitsberufen - aktuelle Handlungs- und Forschungsfelder. 1. Aufl. Bielefeld: W. Bertelsmann Verlag (Berichte zur beruflichen Bildung), S. 185–204, zuletzt geprüft am 04.10.2018.

Wittmann, Eveline; Weyland, Ulrike; Kaspar, Roman; Döring, Ottmar; Hartig, Johannes; Nauerth, Annette; Rechenbach, Simone; Möllers, Michaela; Simon, Julia; Worofka, Iberé (2015): Betriebliche Ausbildungsmerkmale und berufsfachliche Handlungskompetenz in der Altenpflege. In: *Zeitschrift für Berufs - und Wirtschaftspädagogik* 111 (3). Online verfügbar unter https://www.researchgate.net/ publication/288503665_Betriebliche_Ausbildungsmerkmale_und_berufsfachliche_Handlungskompetenz_in_der_Altenpflege, zuletzt geprüft am 25.01.2019.

Wittmann, Eveline; Weyland, Ulrike; Nauerth, Annette; Döring, Ottmar; Rechenbach, Simone; Simon, Julia; Worofka, Iberé (2014): Kompetenzerfassung in der Pflege älterer Menschen - Theoretische und domänenspezifische Anforderungen der Aufgabenmodellierung. In: Jürgen Seifried, Uwe Faßhauer

und Susan Seeber (Hg.): Jahrbuch der berufs- und wirtschaftspädagogischen Forschung 2014. Leverkusen: Budrich, Barbara (Schriftenreihe der Sektion Berufs- und Wirtschafts-pädagogik der Deutschen Gesellschaft für Erziehungswissenschaft (DGfE)), S. 53–66. Online verfügbar unter https://www.pedocs.de/volltexte/2014/9725/pdf/Seifried_JB_berufs_wirtschaftsp_Forschung_2014_Wittmann_ua_Kompetenzerfassung_.pdf, zuletzt geprüft am 25.01.2019.

Wollschläger, Daniel (2017): Grundlagen der Datenanalyse mit R. Eine anwendungsorientierte Einführung. 4., überarbeitete und erweiterte Auflage. Berlin: Springer Spektrum (Statistik und ihre Anwendungen). Online verfügbar unter http://dx.doi.org/10.1007/978-3-662-53670-4.

Abkürzungsverzeichnis

Abb.	Abbildung
BfG	Bundesministerium für Gesundheit
BMBF	Bundesministerium für Bildung und Forschung
BMFSFJ	Bundesministerium für Familie, Senioren, Frauen und Jugend
bzw.	beziehungsweise
ca.	circa
d. h.	das heißt
Dr.	Doktor/Doktorin
DRK	Deutsches Rotes Kreuz
dt.	deutsch
EDV	Elektronische Datenverarbeitung
et al.	et alii/und andere
f.	und folgende Seite
ff.	und weitere folgende Seiten
gGmbH	gemeinnützige Gesellschaft mit beschränkter Haftung
GmbH	Gesellschaft mit beschränkter Haftung
Gl.	Gleichung
h	Uhr
H_0	Nullhypothese
H_1	Gegenhypothese
Hg.	Herausgeber
IT	Informationstechnik
o.S.	ohne Seite
Prof.	Professor/Professorin
RLP	Rheinland-Pfalz

SHA	Schwäbisch Hall
Tab.	Tabelle
TN	Teilnehmer/Teilnehmern
u.	und
u.a.	unter anderem
usw.	und so weiter
vgl.	vergleiche
z.B.	zum Beispiel

Hinweis: Weitere Abkürzungen sind im Text erklärt!

Abbildungsverzeichnis

Abb. 1: Gesamtbetrachtung des „heuristischen“ Kompetenzmodells mit Bezug auf die Pflegepraxis 34
Abb. 2: Pflegesettings in „TEMA“ 36
Abb. 3: „Startseite“ in ein Pflegesetting 37
Abb. 4: „Beispieltestblatt“ aus dem pflegerischen und handlungsorientierten Kompetenztest 37
Abb. 5: Anzahl der teilgenommenen Bildungseinrichtungen 52
Abb. 6: Übersicht zur Verteilung der beantworteten Items aller TN aus SPSS 70
Abb. 7: Gesamtpunktzahlen der TN unter Scoringbedingungen in SPSS 71
Abb. 8: Verteilungsdiagramm der korrekten Antworten aus „CoSMed“ 73
Abb. 9: Korrelationen zwischen latenter Variable und ihren manifesten Variablen 77
Abb. 10: Erklärungen zum Aufbau einer ICC in der PTT aus den verwendeten Testitems in der Gesundheits- und Krankenpflege 78
Abb. 11: Spezifizierte ICC zur Iteminformation und Personenfähigkeit 85
Abb. 12: IC-Funktionen unter Berücksichtigung des 2 PL-Modells 87
Abb. 13: Beispiel eines Partial-Credit-Modells mit drei Schwellenparametern 90
Abb. 14: Übersicht aller dichotomen und ordinalen bzw. polytomen Items im Kompetenztest in der Gesundheits- und Krankenpflege 94
Abb. 15: Übersicht der Itemschwierigkeiten der polytomen Testitems 95
Abb. 16: Dichotomes Itembeispiel A41 aus dem Bereich DR 100
Abb. 17: Polytomes Itembeispiel A82 aus dem Bereich DR 101
Abb. 18: Gesamtübersicht aller Item-Fit-Werte in der Gesundheits- und Krankenpflege 102
Abb. 19: Passende Infit- und Outfitwerte aus der Gesundheits- und Krankenpflege 103
Abb. 20: Beispiel eines einfachen Messmodells 106
Abb. 21: Beispiel eines Strukturgleichungsmodells 108
Abb. 22: Entwickelte Faktorenmodelle im Rahmen des „TEMA“-Kompetenzprojektes 114
Abb. 23, 24, 25: Ergebnisse der drei Strukturgleichungsmodelle aus der CFA als Plots erzeugt mit R 119

Tabellenverzeichnis

Tab. 1: Verteilung der entwickelten Items einschließlich entwickelter Videovignetten im unmittelbar bewohner- und klientenorientierten Kompetenzbereich 35
Tab. 2: Teilnehmerauswahl 54
Tab. 3: Personenbezogene Daten der TN 65
Tab. 4: Zwei Itemauswertungsbeispiele zur Häufigkeitsverteilung in SPSS 69
Tab. 5: Beantwortungsmuster „CoSMed“ in SPSS 72
Tab. 6: Auszug aus den ermittelten Ergebnissen von Itemschwierigkeiten aus TAM 92
Tab. 7: Auszug aus Diskriminationsermittlungen in TAM 93
Tab. 8: Modellvergleich des PCM und GPCM in der Gesundheits- und Krankenpflege für das Antwortverhalten 96
Tab. 9: Ergebnisse aus der globalen Modelltesttatistik in „TAM“ 98
Tab. 10: Übersicht der drei verglichenen Messmodelle mit den zentralen entnommenen Ergebnissen aus R „lavaan“ 116
Tab. 11: Modellvergleichstabelle I (Anova Mod1 zu Mod2) 120
Tab. 12: Modellvergleichstabelle II (Anova Mod1 zu Mod3) 120
Tab. 13: Modellvergleichstabelle III (Anova Mod2 zu Mod3) 120

Mathematisches Formelverzeichnis

Gl. 1: Sample Size

Gl. 2: Cronbachs Alpha

Gl. 3: Rasch-Modell unter IC-Funktion

Gl. 4: Likelihood-Funktion

Gl. 5: LogLikelihood-Funktion

Gl. 6: AIC

Gl. 7: BIC

Gl. 8: CAIC

Gl. 9: Goodness of Fit-Test

Gl. 10: Iteminformationsfunktion

Gl. 11: Testinformationswert

Gl.12: 2 PL-Modell mit Diskriminationsparameter

Gl. 13: Logarithmierter Quotient

Gl. 14: Partial-Credit-Modell

Gl. 15: Lineare Gleichungen

Gl. 16: Freiheitsgrade

Gl. 17: RMSEA

Gl. 18: SRMR

Gl. 19: CFI-Wert

Anhang

Anhang 1: Standardisiertes Informationsschreiben für interessierte TN 150
Anhang 2: Einverständniserklärung und Datenschutzinformation 151
Anhang 3: Personenbezogener Datenerfassungsbogen (auf Grundlage von „SiKoFak“)..... 152
Anhang 4: Forschungstagebuch „Kompetenzmessung“ (Nachweis)..... 153
Anhang 5: „Projektanfrage“ an das Forschungsinstitut für Berufliche Bildung in Nürnberg 154
Anhang 6: Original „TEMA“ Server (nach Aufbau)..... 156
Anhang 7: Forschungsequipment für die Feldphase (auf- und abgebaut) 158
Anhang 8: Anschreiben „Namensfreigabe“ der Bildungseinrichtung 160
Anhang 9: Übersicht der teilgenommenen Bildungseinrichtungen 162
Anhang 10: Zusatzadapter und Einwahl in ein zweites Netzwerk 164
Anhang 11: Vorbereitete Notebooks für die TN und Beispiele von Forschungsfeldern..... 165
Anhang 12: Screenshots aus Originaldatensätzen in Microsoft „EXCEL“ 168
Anhang 13: Screenshots aus Originaldatensätze nach Recodierung in „SPSS“... 169
Anhang 14: Erstellte Punkteliste - Scoring der einzelnen Items in Anlehnung an „TEMA“ 170
Anhang 15: Ergebnisse zu Itemhäufigkeitsverteilungen in SPSS (ergänzend zu Kapitel 14.2) 174
Anhang 16: Veränderungen des Cronbachs Alpha und Item-Trennschärfe..... 199
Anhang 17: Item-Mittelwerte in der Gesundheits- und Krankenpflege aus SPSS . 201
Anhang 18: Entwickeltes R-Skript für die IRT-Berechnungen..... 203
Anhang 19: Itemparameter unter der Bedingung des PCM nach Ausgabe in R ... 209
Anhang 20: Itemparameter unter der Bedingung des GPCM in R 214
Anhang 21: Diskriminationsparameter im GPCM = Mod2 in “TAM” 216
Anhang 22: Entwickeltes R-Skript in „eRm“ u.a. zur globalen Modelgeltung..... 218
Anhang 23: Itemcharakteristische Funktionen im 1 PL-Modell in „TAM“ und „irtoys“ mit beobachteten und erwarteten Lösungshäufigkeiten (weitere Beispiele)..... 220
Anhang 24: Infit- und Outfitwerte einschließlich p-Werten 224
Anhang 25: Entwickeltes R-Skript für die CFA-Berechnungen 229

Anhang 1: Standardisiertes Informationsschreiben für interessierte TN

Ludwigsburg im Oktober 2017

Liebe Schülerinnen und Schüler, liebe Studierende,

im Rahmen meines Dissertationsvorhabens an der Philosophisch-Theologischen Hochschule in Vallendar suche ich Gesundheits- und Krankenpflegeschulen und Hochschulen, die mich gerne bei meiner wissenschaftlichen Studie unterstützen würden.

Gerne informiere ich Sie über meine **computergestützte Befragung zur Einschätzung der beruflichen Kompetenz in der Gesundheits- und Krankenpflege**. Innerhalb der Testung geht es um alltägliche Handlungssituationen, die von Ihnen als Teilnehmer/in eingeschätzt werden sollen.

Wichtig für Sie als Teilnehmer/in: Die Befragung ist keine schriftliche Prüfung mit Benotung, Einzelbewertung oder konkreter Abfrage von Lernstoff!!! Alles bleibt absolut anonym! Das „Testinstrument“ soll in erster Linie von Ihnen erprobt und erfahren werden.

Was heißt das?

- Sie sind im 3. Ausbildungsjahr bzw. 6. oder 7. Semester im Bachelorstudiengang Pflege DUAL;
- Sie haben Lust Ihr erworbenes praktisches Fachwissen einzusetzen und zu überprüfen (anonym) und gleichzeitig Ihre Meinung in die Pflegewelt mit einzubringen;
- Sie möchten Pflegeforschung im praktischen Kontext erfahren;
- Sie suchen eine Möglichkeit der praktischen Prüfungsvorbereitung.

Dann erfüllen Sie alle Voraussetzungen für eine Teilnahme!

Wichtig für Sie als Teilnehmer/in:

- Die Teilnahme ist absolut freiwillig, es kann jederzeit von der Teilnahme zurückgetreten werden.
- Die Dauer der computergestützten Testung beträgt insgesamt mit Aufklärung über die Testung ca. 2 Zeitstunden.
- Es gibt keine Noten oder individuelle Rückmeldungen.
- Der "Test" zeigt Videoszenen aus der täglichen, praktischen Arbeit - mit verschiedenen Antwortmöglichkeiten - somit gut von Ihnen zu bewältigen.
- Zwei Fachleute begleiten Sie unterstützend durch die gesamte "Testungsphase".

Erbetene Vorbereitungen für Sie als Teilnehmer/in:

Es wäre super, wenn Sie ein Earphone oder Kopfhörer mitbringen könnten. So werden die weiteren Teilnehmer/innen durch das Ablaufen der Videosequenzen nicht gestört! Des Weiteren noch einen Kugelschreiber zum Unterschreiben der Einverständnis- und Datenschutzerklärung zur anonymisierten Verwertung der erhobenen Daten zu pflegewissenschaftlichen Zwecken.
Nach erfolgreichem Abschluss der Dissertation wird Ihrer Schule bzw. Hochschule diese auf Wunsch per PDF zur Verfügung gestellt.

Vielen Dank für Ihre hilfreiche Unterstützung! Ich würde mich wirklich sehr freuen, Sie kennen zu lernen! Bei Fragen oder Anregungen stehe ich Ihnen jederzeit zur Verfügung!

Herzliche Grüße

Simone Ries

Promovendin an der PTHV Vallendar
Betreuer: Herr Prof. Dr. Albert Brühl - Lehrstuhl für Statistik und standardisierte Verfahren der Pflegeforschung
---
Simone Ries, MScN, Wissenschaftliche Mitarbeiterin, Evangelische Hochschule Ludwigsburg, Paulusweg 6, 71638 Ludwigsburg, Tel: 07141-9745-283, s.ries@eh-ludwigsburg.de, www.eh-ludwigsburg.de

Anhang 2: Einverständniserklärung und Datenschutzinformation

Promovendin: Frau Simone Brigitte Ries, MScN
Promotionsbetreuer: Herr Prof. Dr. Albert Brühl, Philosophisch-Theologische Hochschule Vallendar
Paulusweg 6
71638 Ludwigsburg
Tel.: 07141-9745-283
Handy: 0170-2374873
E-Mail: s.ries@eh-ludwigsburg.de

_____

Nachname, Vorname

Einverständniserklärung und Datenschutzinformation

Hiermit erkläre ich mich einverstanden, dass Frau Ries im Rahmen ihrer Promotionsarbeit an der Philosophisch-Theologischen Hochschule Vallendar (Fakultät: Pflegewissenschaft) die erhobenen computergestützten beruflichen Kompetenzmessungsdaten zu wissenschaftlichen Zwecken verwenden darf. Das vorläufige Thema der Forschungsstudie lautet: „Berufliche computergestützte Handlungskompetenzmessung in der Gesundheits- und Krankenpflege in Anlehnung an das Forschungsprojekt TEMA“.

Ich wurde von Frau Ries über Wesen, Bedeutung und Tragweite der Studie vollständig aufgeklärt. Ich kann jederzeit weitere Fragen zu Inhalten der Studie stellen.

Die Teilnahme an der Studie ist absolut freiwillig. Die erhobenen Daten werden nicht an die jeweilige Schuleinrichtung oder andere Dritte des Unternehmens weitergegeben.

Jederzeit und ohne Angabe von Gründen kann ich meine Zustimmung widerrufen, ohne dass ein Nachteil für mich entsteht. Zu meinem Widerruf schreibe ich eine Mail an die oben stehende E-Mailadresse mit Angabe meines Benutzernamens. Die erhobenen Daten werden daraufhin umgehend gelöscht.

Mir ist bekannt, dass meine allgemeinen und personenbezogenen Daten anonym und ausschließlich für pflegewissenschaftliche Forschungszwecke verwendet, gespeichert, quantitativ ausgewertet und veröffentlicht werden dürfen (Nach der Auswertung entstehen u.a. anonymisierte Gruppenergebnisse.). Es entstehen für mich keine Kosten, noch werde ich finanziell für die Teilnahme an der Studie entschädigt.

Ich erkläre hiermit meine freiwillige Teilnahme. Eine Kopie der Einwilligungserklärung werde ich erhalten.

Nach Beendigung der Gesamtstudie bzw. Promotion wird die schriftliche Ausarbeitung gerne zur Verfügung gestellt.

_____

Ort, Datum Unterschrift

Anhang 3: Personenbezogener Datenerfassungsbogen (auf Grundlage von „SiKoFak“)

Personenbezogene Daten zur Kompetenzmessung in der Gesundheits- und Krankenpflege - Frau Ries

Bitte **unterstreichen** Sie das Merkmal, welches nach Ihrer Meinung am ehesten für Sie zutrifft oder ergänzen Sie mit einem weiteren Schrägstrich (/)!

Soziodemographische Merkmale

1. Geschlecht: männlich / weiblich
2. Geburtsland: Deutschland / Europa / außerhalb der EU
3. Lebensalter: unter 20 / 20-30 / 30-40 / über 40
4. Höchster Schulabschluss: Hauptschule mit Zusatzabschluss (entspricht dem mittleren Bildungsabschluss / Realschule / Gymnasium
5. Familienstand: ledig / verheiratet / verwitwet / alleinerziehend / keine Angabe
6. Familiäre Pflege (Versorgung eines Angehörigen zu Hause): findet statt / fand statt / trifft nicht zu
7. Lesen von Büchern oder Fachzeitschriften: nie / einmal pro Jahr / einmal im Monat / einmal in der Woche / häufiger als einmal pro Woche
8. Computer- oder Internetnutzung: nie / weniger als zwei Stunden pro Woche / mehr als zwei Stunden pro Woche

Ausbildungsbezogene Hintergrundmerkmale

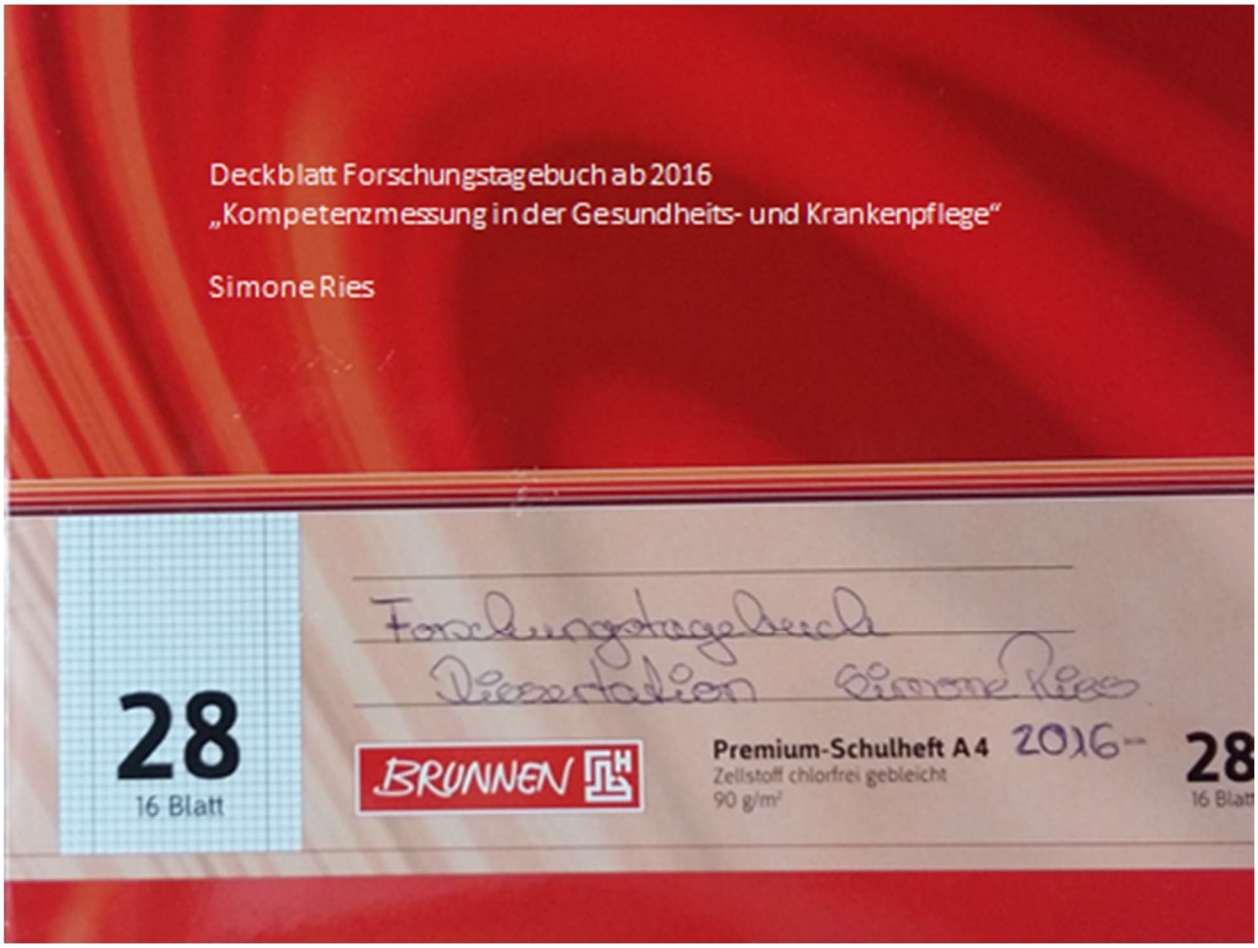
1. Übergang von Schule zu Ausbildung bzw. Studium der Pflege: direkt / erst später
2. Ausbildungszeit bzw. Studienzeit: verkürzt / drei Jahre / 8. Semester / 9. Semester / andere Studien- bzw. Ausbildungszeit
3. Primärer Grund für die Berufswahl: weil der Beruf Spaß macht / weil der Beruf eine sichere Berufsperspektive versprach / weil ich mich für den Beruf für geeignet gehalten habe / weil mir meine Eltern, Verwandte, Freunde, PartnerIn dazu geraten haben bzw. hat / weil ich für diesen Beruf später respektiert werde / weil andere Menschen zu begleiten, einfach eine große Erfüllung bringt / anderer Grund: _____

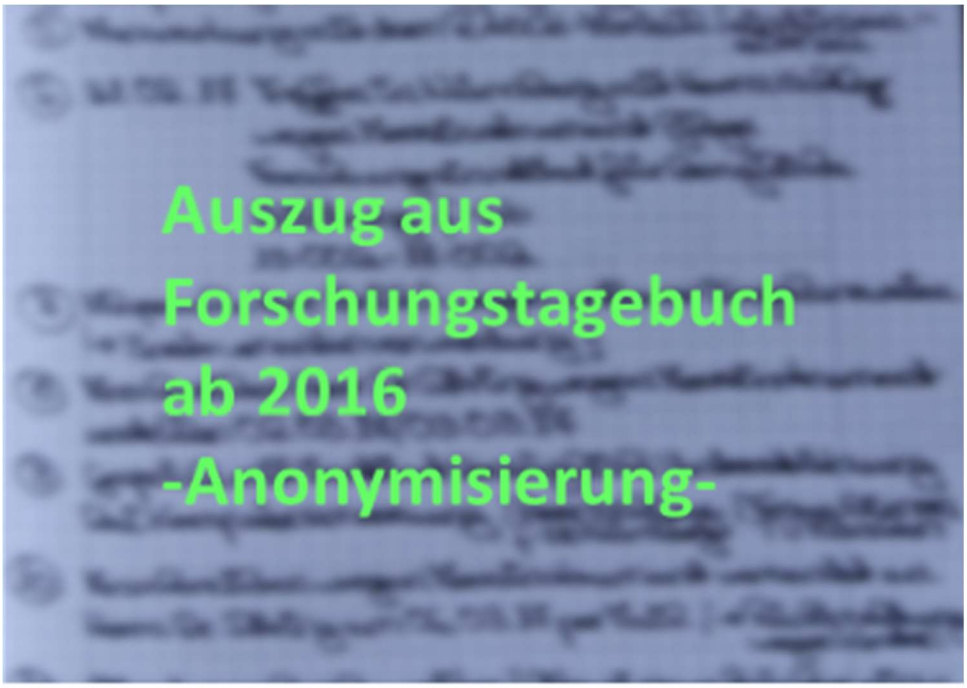
Besonderes Qualitätsmerkmal der beruflichen Gesundheits- und Krankenpflegeausbildung bzw. des Studiums der Pflege

1. Betriebliches Interaktionsklima
2. Unterstützung durch Ausbilder
3. Unterstützung durch Kollegen
4. Betriebliche Gestaltungsqualität
5. Betriebliche Instruktionsqualität
6. Betriebliche Lernprozesskontrolle
7. Strukturqualität betrieblicher Arbeitsaufgaben
8. Aufgabenvielfalt, Fähigkeitsentsprechung
9. Betriebliche Bedeutsamkeit
10. Inhaltliche Autonomie der betrieblichen Aufgabenbearbeitung
11. Zeitliche Dispositionschancen der betrieblichen Aufgabenbearbeitung
12. Förderung des Lernklimas im familiären Umfeld
13. Anderes Qualitätsmerkmal (bitte aufführen):

Vielen Dank für die Bearbeitung und viel Erfolg für die Zukunft!
Simone Ries, MScN

Anhang 4: Forschungstagebuch „Kompetenzmessung“ (Nachweis)





© Ries - eigenes Bildmaterial

Anhang 5: „Projektanfrage“ an das Forschungsinstitut für Berufliche Bildung in Nürnberg

Vorhaben:

Dissertation im Bereich: Kompetenzmessung in der beruflichen Gesundheits- und Krankenpflegeausbildung in Anlehnung an „TEMA“ (vorläufiger Titel der Arbeit!)

Hintergrund: Die wachsenden beruflichen, organisationsbezogenen und persönlichen Anforderungen im pflegerischen Arbeitsumfeld, gleichzeitig immer weniger zeitliche Ressourcen und ein vorhandener Fachkräftemangel verlangen nach einem Handlungsprofil für die klinischen und ambulanten Pflegeeinrichtungen und im Besonderen für die Pflegekräfte.

Eine gute Einschätzungsmöglichkeit der beruflichen Pflegekompetenzen zeigte das bereits für die Altenpflege modellierte domänenspezifische Kompetenzmodell, aufgeteilt in sechs Fragmente aus dem Projekt „TEMA“. Die Umsetzung im Feld erfolgte bei den Probanden in computergestützter Form. Das pflegerische Handeln wurde an mehreren Handlungssituationen quantitativ gemessen und ausgewertet.

Das verwendete Kompetenzmodell soll bei der anstehenden Dissertation nicht im Umfeld der Altenpflege, sondern in der Gesundheits- und Krankenpflege eingesetzt werden. Im Fachartikel: Technologiebasierte Messung beruflicher Handlungskompetenz in der Pflege älterer Menschen: Kompetenzmodellierung und Testverfahrensentwicklung[261] wird auf die mögliche Übertragung hingewiesen, da die Handlungsfelder auch in der Gesundheits- und Krankenpflege vorzufinden sind.

Vorstellungen, Vorschläge und Anregungen für die anstehende „Feldarbeit“ im Rahmen des Dissertationsvorhabens:

- Einsatz des computergestützten Messinstrumentes (unmittelbar bewohner-/klientenbezogener Kompetenzbereich) vom Projekt „TEMA“[262] (Beim gemeinsamen Telefonat von Herrn Dr. Döring und Frau Ries, stand im Raum, dass der Server für zwei Monate kostenlos für den Probandeneinsatz zur Verfügung gestellt werden könnte. [?])
- Einsatzzeitraum des Messinstrumentes/Feldzugang: voraussichtlich Juli-August 2017 (Examenszeit an den Schulen bei Abschluss der Gesundheits- und Krankenpflegeausbildung Ende September 2017); ein Pretest sollte an zwei Schulen im Vorfeld (Mitte/Ende Juni 2017) stattfinden

[261] Döring, O. et al. (2016): Technologiebasierte Messung beruflicher Handlungskompetenz in der Pflege älterer Menschen: Kompetenzmodellierung und Testverfahrensentwicklung. In: Dietzen, A. et al. (Hrsg.): Kompetenzorientierung. Berufliche Kompetenzen entwickeln, messen und anerkennen. Bielefeld: Bertelsmann, S. 117-132.

[262] Döring, O. et al. (2016): Technologiebasierte Messung von beruflichen Kompetenzen für die Pflege älterer Menschen: berufsfachliche Kompetenzen, allgemeine Kompetenzen und Kontextfaktoren (TEMA). In: Beck, K.; Landenberger, M.; Oser, F. (Hrsg.): Technologiebasierte Kompetenzmessung in der beruflichen Bildung. Ergebnisse aus der BMBF-Förderinitiative ASCOT. Bielefeld: Bertelsmann, S. 243-264.

- Einrichtungen: Einsatz des computergestützten Messinstrumentes in ca. 15-18 Bildungszentren für Gesundheit und Pflege bzw. Gesundheits- und Krankenpflegeschulen in Süddeutschland - Schwerpunkt Bayern, Baden-Württemberg (ein Gesamt-N von ca. 300-350 Probanden soll erreicht werden); eine Anreise an den Arbeitsort von Frau Ries (Ludwigsburg) wäre zwar für die Durchführung erleichternd, könnte sich aber als Hemmnis oder Ablehnung von Seiten der Probanden darstellen.
- Ein „IT-Spezialist“ wurde empfohlen wegen dem Serveraufbau an den jeweiligen Schulen (technische Betreuung). Könnte dies Frau Ries nach Einweisung durch den „IT“-Spezialisten auch selbst übernehmen (unter Berücksichtigung der rechtlichen Kriterien[?])?
- Die Laptops (ca. 20 Stück) werden von Frau Ries zur Verfügung gestellt und am jeweiligen Einsatzort für den „IT“-Spezialisten und die Probanden vorbereitet einschließlich der dortigen Organisation des Raumes zur Testdurchführung. Eine Übersichtsliste mit den kooperierenden „Test“-Schulen und dem jeweils vorbereiteten Raum in den Einrichtungen werden an Herrn Dr. Döring und den „IT“-Spezialisten per Mail oder per Post übermittelt.
- Es sind vor Beginn des Feldzugangs rechtliche Rahmenbedingungen mit Frau Ries zu besprechen, wie z.B. Regressforderungen, usw. Dies wurde bereits im stattgefundenen Telefonat von Herrn Dr. Döring angesprochen.
- Eine Kostenkalkulation wird vom „F-BB“ erstellt und an Frau Ries geschickt. Es wäre für die weiteren Planungen wünschenswert, wenn diese **bis spätestens Ende März 2017** vorliegen könnte. Anmerkung: Zu berücksichtigen ist, dass Frau Ries die Kosten aus eigener Tasche tragen muss.
- Eine Erweiterung/Ergänzung des Tests ist von Frau Ries per Paper and Pencil -> Teilkompetenz/Selbstbezogener Kompetenzbereich - Gesundheitserhaltungskompetenz denkbar.

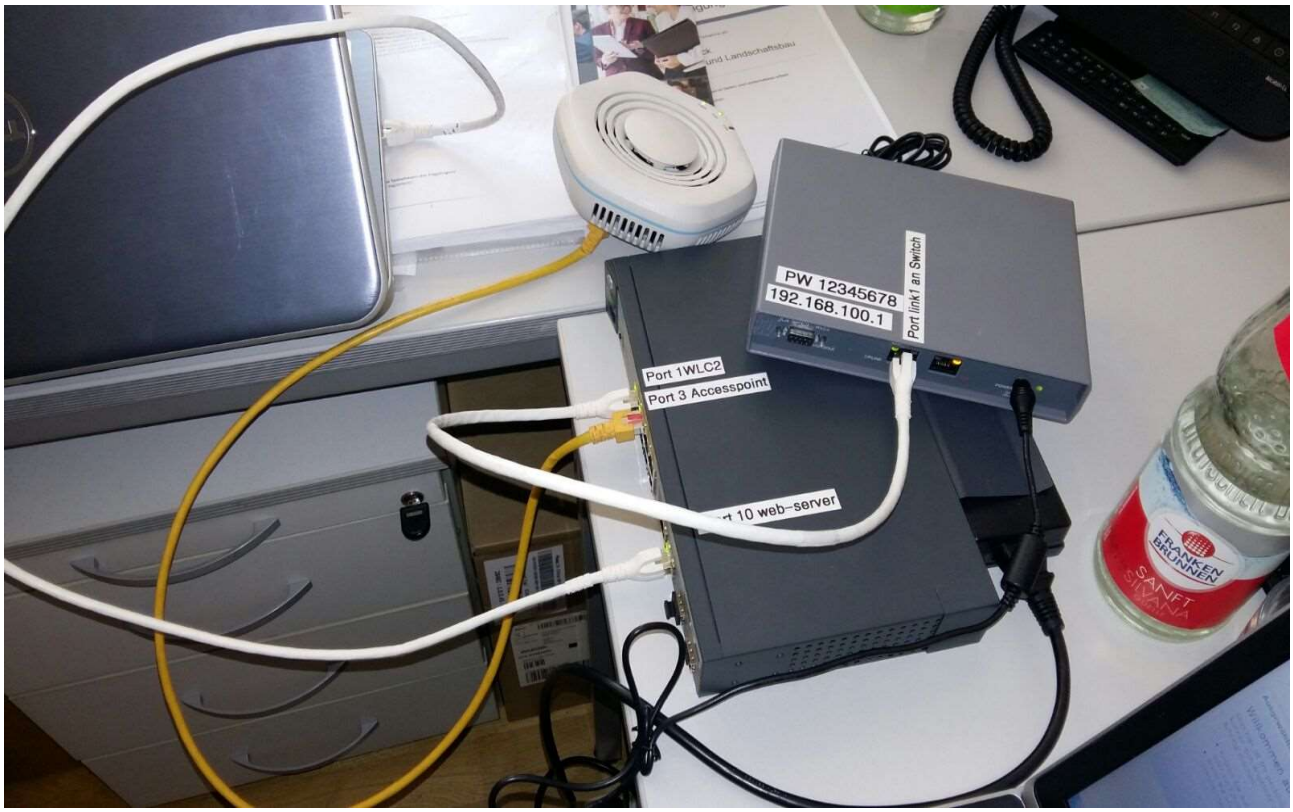
Vorab ein herzliches Dankeschön für die bereits erwiesene Unterstützung und Aufgeschlossenheit bezüglich meines Dissertationsvorhabens.

Ludwigsburg, 03.03.2017

Simone Ries
Wissenschaftliche Mitarbeiterin
Bachelorstudiengang Pflege
Evangelische Hochschule Ludwigsburg
Promoventin an der PTHV Vallendar

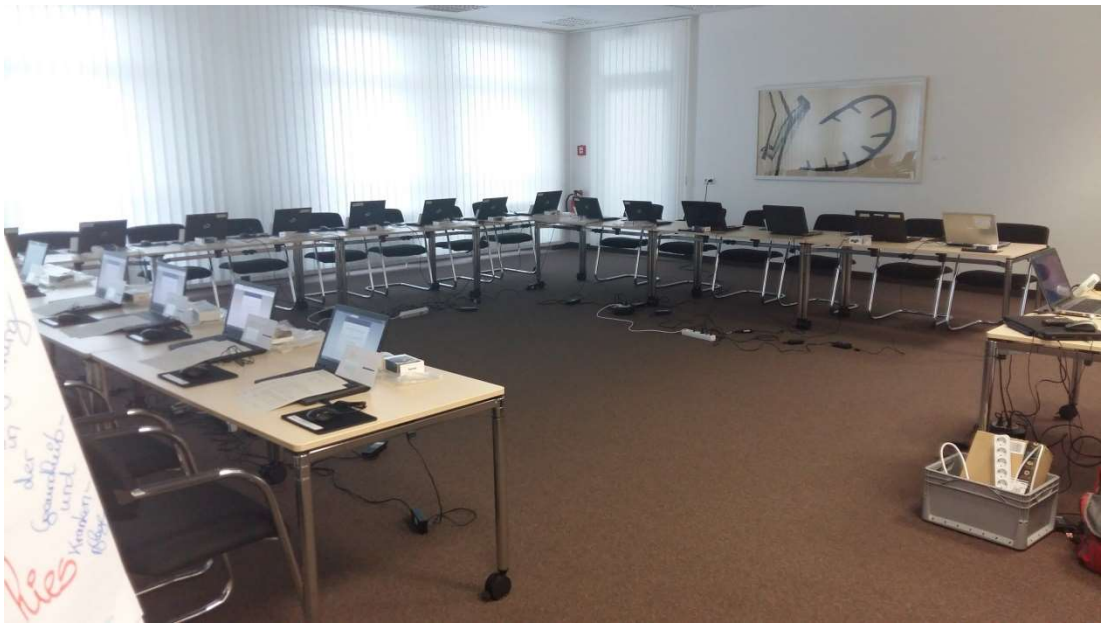
Dissertationsbetreuer: Herr Prof. Dr. Albert Brühl (Lehrstuhl für Statistik und standardisierte Verfahren der Pflegeforschung) an der PTHV Vallendar; abruehl@pthv.de, Tel.: +49-261-6402 257 oder Mobil: +49-170-1653157

Anhang 6: Original „TEMA“ Server (nach Aufbau)



© Ries - eigenes Bildmaterial

Anhang 7: Forschungsequipment für die Feldphase (auf- und abgebaut)



© Ries - eigene Bilder

Anhang 8: Anschreiben „Namensfreigabe“ der Bildungseinrichtung

Ludwigsburg, 20.08.2018

Sehr geehrte teilnehmende Bildungseinrichtung bzw. Hochschuleinrichtung,

es ist schon eine ganze Weile her, dass Sie bzw. Ihre Auszubildenden/Studierenden an der „Kompetenzmessung in der Gesundheits- und Krankenpflege“ teilgenommen haben. Gerne wende ich mich heute an Sie und kann Ihnen mitteilen, dass die Datenauswertung in vollem Gang ist und das Schreiben der Arbeit parallel dazu erfolgt. Sobald alle Auswertungen abgeschlossen sind und die Veröffentlichung der Dissertation stattgefunden hat, werde ich Ihnen gerne die Ergebnisse in Wort oder/und Schrift vorstellen.

Für einen weiteren Schritt im meiner schriftlichen Ausarbeitung ist es wichtig, dass Sie mir **die Zustimmung zur Veröffentlichung des Namens Ihrer Bildungs- bzw. Hochschuleinrichtung** geben.

In der Dissertationsschrift sollen die einzelnen Einrichtungen sowie eine Liste aller Einrichtungen im Anhang der Arbeit aufgelistet werden.

Dies wird voraussichtlich in der Dissertation so aussehen:

Punkt: 11.2
„Am computergestützten Kompetenztest in der Gesundheits- und Krankenpflege haben nachstehende Einrichtungen teilgenommen: “
Der Name der Bildungseinrichtung wird hier benannt!

Im Anhang:
Name der Bildungs- bzw. Hochschuleinrichtung mit Adresse, Schulleitung und wenn abweichend, Ansprechpartner/in im Rahmen des Forschungsprozesses (als Übersichtstabelle).

Deshalb bitte ich Sie, einen der genannten Punkte anzukreuzen oder zu kennzeichnen, damit dies in der Dissertationsschrift exakt berücksichtigt werden kann:

- ○ Die Bildungs- bzw. Hochschuleinrichtung ist mit der oben angeführten Veröffentlichung **einverstanden**.
- ○ Die Bildungs- bzw. Hochschuleinrichtung ist mit der Veröffentlichung **nicht einverstanden**. Diese möchte nur anonym aufgelistet werden.
- ○ Die Bildungs- bzw. Hochschuleinrichtung erlaubt folgende Benennung in der Dissertationsschrift (Bitte führen Sie den gewünschten Text auf!):

Bitte schicken Sie mir den ausgefüllten Bogen bis spätestens 25.09.2018 per Post oder per Mail an die folgende Adresse:

Evangelische Hochschule Ludwigsburg
Simone Ries
Paulusweg 6
71638 Ludwigsburg
oder
per Mail:
s.ries@eh-ludwigsburg.de

Vielen Dank für Ihre Unterstützung und wir bleiben im Austausch!

Ihre Kollegin aus der Pflegepädagogik
Simone Ries

Promotionsbetreuer: Herr Prof. Dr. Albert Brühl - Lehrstuhl für Statistik und standardisierte Verfahren der Pflegeforschung - PTH Vallendar, Pallottistraße 3, 56179 Vallendar

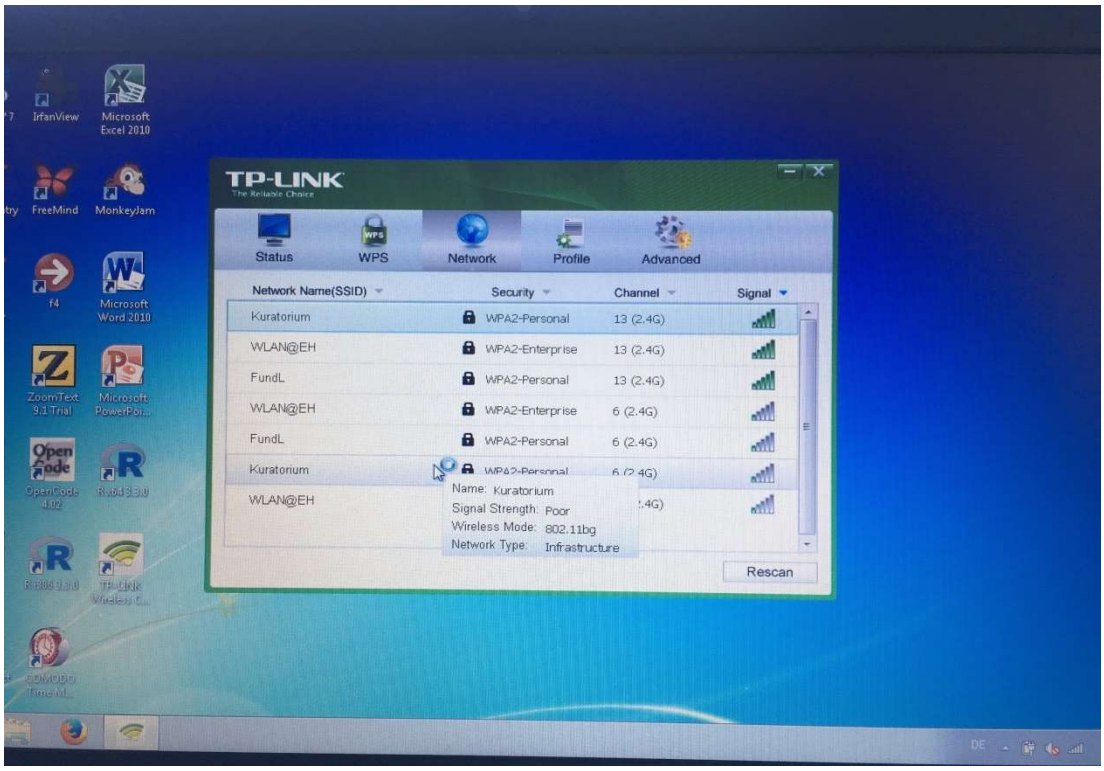
Anhang 9: Übersicht der teilgenommenen Bildungseinrichtungen

Name der Bildungseinrichtung:	Leitung/AnsprechpartnerIn im Projekt:
Schule für Pflegeberufe am Klinikum Crailsheim Gartenstraße 21 74564 Crailsheim	Herr Helmut Hartmann - stellv. Schulleitung, Frau Katharina Weber
Diakonie-Klinikum SHA gGmbH Schule für Pflegeberufe Gelbinger Gasse 25 74523 Schwäbisch Hall	Frau Susanne Fischer - Schulleitung
Gesundheits- und Krankenpflegeschule Künzelsau - Hohenloher Krankenhaus gGmbH Stettenstraße 32 74653 Künzelsau	Herr Martin Boochs - Schulleitung, Herr Marius Rebmann
Evangelische Hochschule Ludwigsburg Bachelorstudiengang „Pflege“ Paulusweg 6 71638 Ludwigsburg	Herr Prof. Dr. Norbert Collmar - Rektor, Frau Prof. Dr. Constanze Eylmann, Br. Prof. Dr. Peter Schiffer OSCam
SLK-Kliniken Gesundheitsakademie Klinikum am Gesundbrunnen Am Gesundbrunnen 20-26 74078 Heilbronn	Frau Gisela Jenkner - Schulleitung, Frau Rosemarie Pauz
Gesundheitsakademie Bodensee-Oberschwaben GmbH Schulen für Pflegeberufe Hähnlehofstraße 5 88250 Weingarten	Herr Andreas Thiel - Schulleitung
Evangelisches Bildungszentrum für Gesundheitsberufe Stuttgart Nordbahnhofstraße 131 70191 Stuttgart	Herr Dr. Johannes Nau - Schulleitung, Herr Claus Cavada
Berufsfachschule für Gesundheits- und Krankenpflege - Donau-Ries Klinik Donauwörth Neudegger Allee 11b 86609 Donauwörth	Frau Karola Rigel - Schulleitung, Herr Werner Kraus
Bildungszentrum für Gesundheits- und Krankenpflege der DRK Krankenhaus GmbH RLP Hachenburg Alte Frankfurter Str.12 57627 Hachenburg	Frau Corinna Kronsteiner-Buschmann - Schulleitung, Frau Ulrike Ehrenfried

Klinikum Esslingen GmbH Gesundheits- und Krankenpflegeschule Hirschlandstraße 97 73730 Esslingen	Frau Lore Schöller - Schulleitung
Katholische Hochschule Freiburg Bachelorstudiengang „Angewandte Pflegewissenschaft“ Karlstraße 63 79104 Freiburg	Herr Prof. Dr. Edgar Kösler - Rektor Frau Prof. Dr. Anne Kellner
Krankenpflegeschule der Rhein-Mosel-Fachklinik Andernach Vulkanstraße 58 56626 Andernach	Herr Udo Hoffmann - Schulleitung Frau Despoina Gkountova
Technische Hochschule Deggendorf Fakultät für „Angewandte Gesundheits-wissenschaften“ Land-Au 27 94469 Deggendorf	Herr Prof. Dr. Christian Rester - Studiendekan
Berufsfachschule für Krankenpflege am Krankenhaus Martha-Maria Nürnberg Stadenstraße 62b 90491 Nürnberg	Herr Friedhelm Harsch - Schulleitung

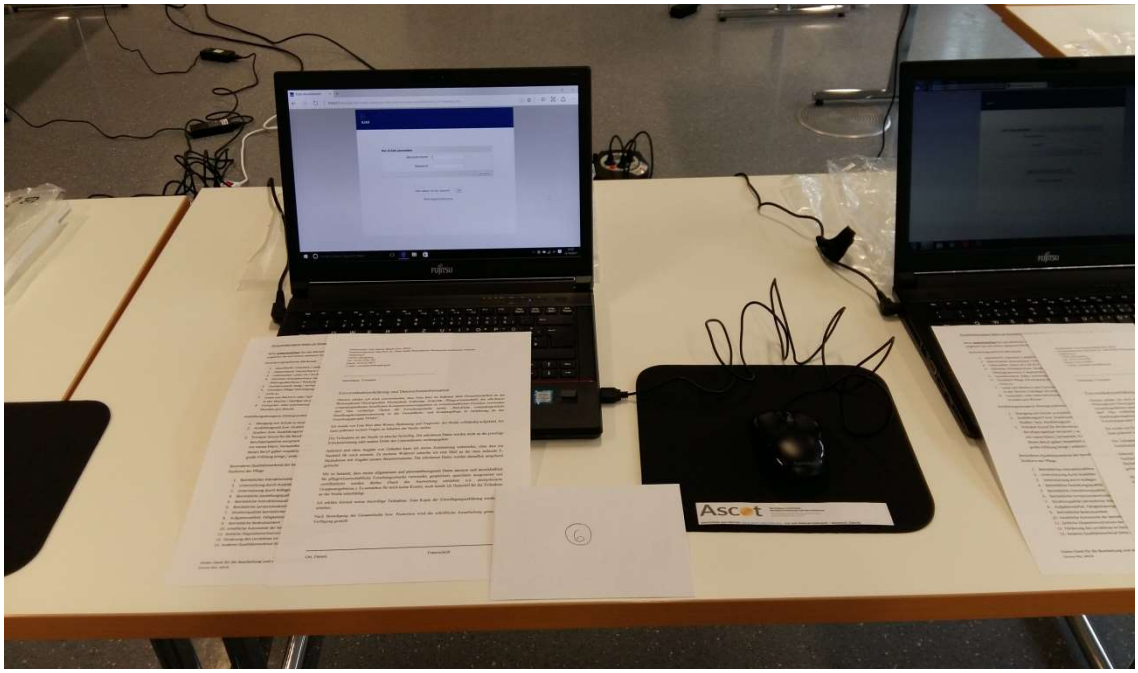
Hinweis: Einverständnis zur namentlichen Benennung der Bildungseinrichtung wurde schriftlich erfragt und von oben aufgeführten erteilt! Alle weiteren Einrichtungen dürfen aus datenschutzrechtlichen Gründen in dieser Auflistung nicht angeführt werden!

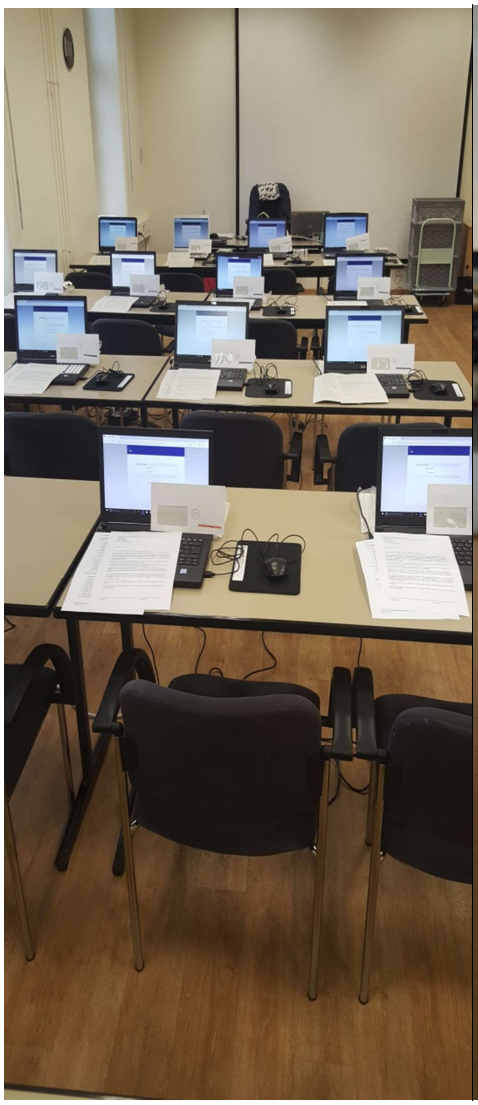
Anhang 10: Zusatzadapter und Einwahl in ein zweites Netzwerk

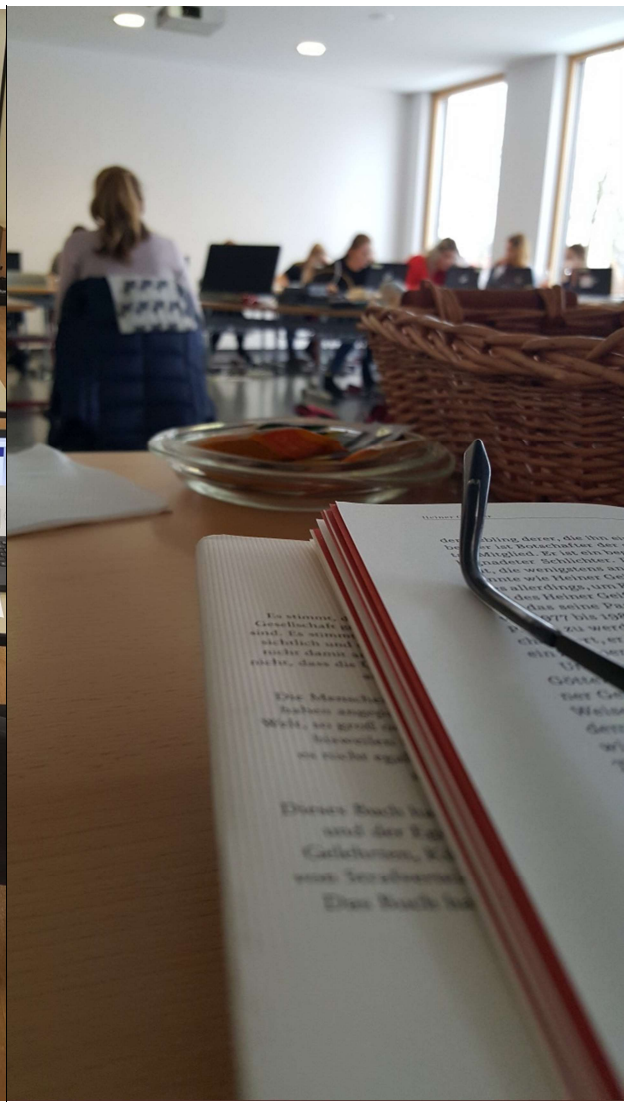


© Ries - eigene Bilder

Anhang 11: Vorbereitete Notebooks für die TN und Beispiele von Forschungsfeldern

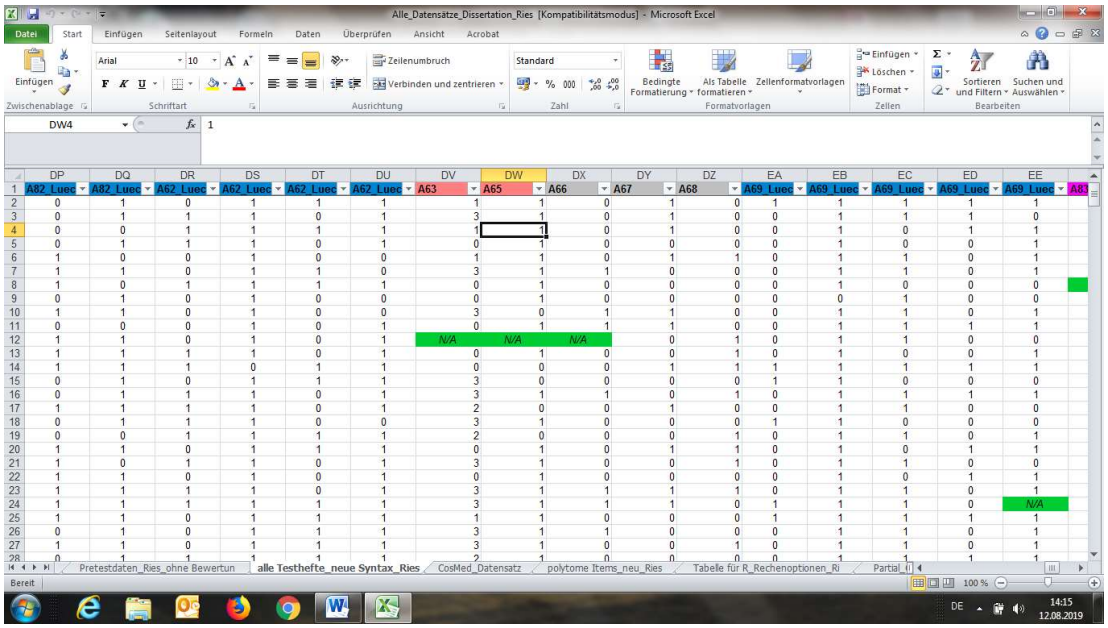




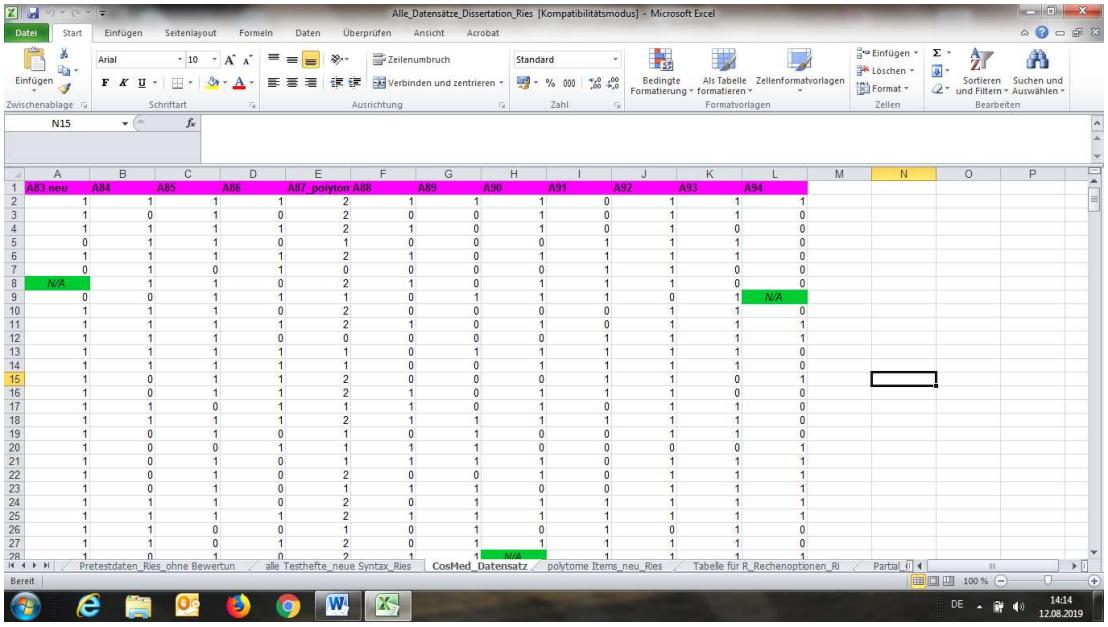


© Ries - eigene Bilder

Anhang 12: Screenshots aus Originaldatensätzen in Microsoft „EXCEL“



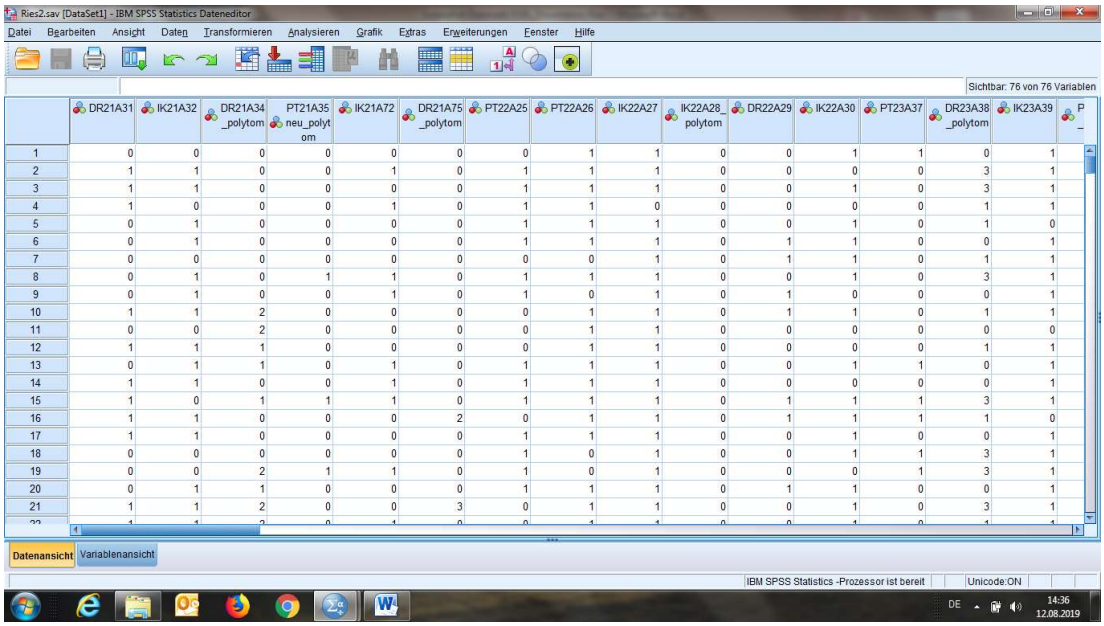
Auszug aus „TEMA“-Datensatz in der Gesundheits- und Krankenpflege_Ries



Auszug aus „CoSMed“-Datensatz in der Gesundheits- und Krankenpflege_Ries

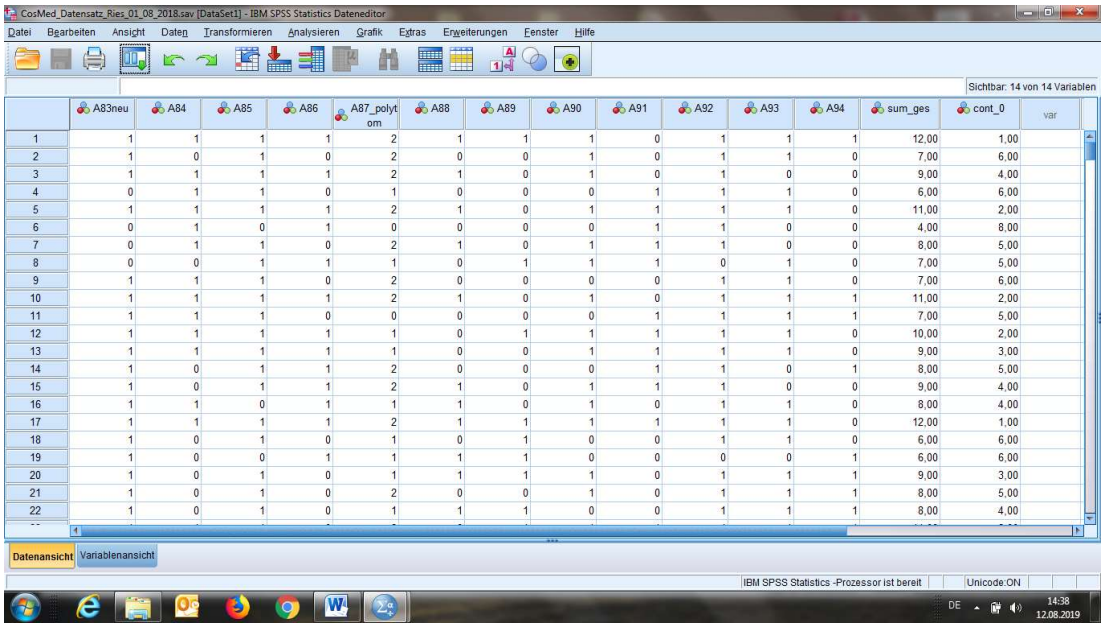
© Ries - eigene Bilder

Anhang 13: Screenshots aus Originaldatensätze nach Recodierung in „SPSS“



	DR21A31	IK21A32	DR21A34_polytom	PT21A35neu_polytom	IK21A72	DR21A75_polytom	PT22A25	PT22A26	IK22A27	IK22A28_polytom	DR22A29	IK22A30	PT23A37	DR23A38_polytom	IK23A39	P_
1	0	0	0	0	0	0	0	1	1	0	0	1	1	0	1	
2	1	1	0	0	1	0	1	1	1	0	0	0	0	3	1	
3	1	1	0	0	0	0	1	1	1	0	0	1	0	3	1	
4	1	0	0	0	1	0	1	1	0	0	0	0	0	1	1	
5	0	1	0	0	0	0	1	1	1	0	0	1	0	1	0	
6	0	1	0	0	0	0	1	1	1	0	1	1	0	0	1	
7	0	0	0	0	0	0	0	0	1	0	1	1	0	1	1	
8	0	1	0	1	1	0	1	1	1	0	0	1	0	3	1	
9	0	1	0	0	1	0	1	0	1	0	1	0	0	0	1	
10	1	1	2	0	0	0	0	1	1	0	1	1	0	1	1	
11	0	0	2	0	0	0	0	1	1	0	0	0	0	0	0	
12	1	1	1	0	0	0	0	1	1	0	0	0	0	1	1	
13	0	1	1	0	1	0	1	1	1	0	0	1	1	0	1	
14	1	1	0	0	1	0	1	1	1	0	0	0	0	0	1	
15	1	0	1	1	1	0	1	1	1	0	1	1	1	3	1	
16	1	1	0	0	0	2	0	1	1	0	1	1	1	1	0	
17	1	1	0	0	0	0	1	1	1	0	0	1	0	0	1	
18	0	0	0	0	0	0	1	0	1	0	0	1	1	3	1	
19	0	0	2	1	1	0	1	0	1	0	0	0	1	3	1	
20	0	1	1	0	0	0	1	1	1	0	1	1	0	0	1	
21	1	1	2	0	0	3	0	1	1	0	0	1	0	3	1	

Auszug aus „TEMA“-Datensatz in der Gesundheits- und Krankenpflege_Ries



	A83neu	A84	A85	A86	A87_polytom	A88	A89	A90	A91	A92	A93	A94	sum_ges	cont_0	var
1	1	1	1	1	2	1	1	1	0	1	1	1	12,00	1,00	
2	1	0	1	0	2	0	0	1	0	1	1	0	7,00	6,00	
3	1	1	1	1	2	1	0	1	0	1	0	0	9,00	4,00	
4	0	1	1	0	1	0	0	0	1	1	1	0	6,00	6,00	
5	1	1	1	1	2	1	0	1	1	1	1	0	11,00	2,00	
6	0	1	0	1	0	0	0	0	1	1	0	0	4,00	8,00	
7	0	1	1	0	2	1	0	1	1	1	0	0	8,00	5,00	
8	0	0	1	1	1	0	1	1	1	0	1	0	7,00	5,00	
9	1	1	1	0	2	0	0	0	0	1	1	0	7,00	6,00	
10	1	1	1	1	2	1	0	1	0	1	1	1	11,00	2,00	
11	1	1	1	0	0	0	0	0	1	1	1	1	7,00	5,00	
12	1	1	1	1	1	0	1	1	1	1	1	0	10,00	2,00	
13	1	1	1	1	1	0	0	1	1	1	1	0	9,00	3,00	
14	1	0	1	1	2	0	0	0	1	1	0	1	8,00	5,00	
15	1	0	1	1	2	1	0	1	1	1	0	0	9,00	4,00	
16	1	1	0	1	1	1	0	1	0	1	1	0	8,00	4,00	
17	1	1	1	1	2	1	1	1	1	1	1	0	12,00	1,00	
18	1	0	1	0	1	0	1	0	0	1	1	0	6,00	6,00	
19	1	0	0	1	1	1	1	0	0	0	0	1	6,00	6,00	
20	1	0	1	0	1	1	1	1	0	1	1	1	9,00	3,00	
21	1	0	1	0	2	0	0	1	0	1	1	1	8,00	5,00	
22	1	0	1	0	1	1	1	0	0	1	1	1	8,00	4,00	

Auszug aus „CoSMed“-Datensatz in der Gesundheits- und Krankenpflege_Ries

© Ries - eigene Bilder

Anhang 14: Erstellte Punkteliste - Scoring der einzelnen Items in Anlehnung an „TEMA“

Punktevergabe der einzelnen Testaufgaben - Kompetenzmessung Pflege

f=falsch
r=richtig

Testheft_2

Aufgabe	korrekte Antwort Nr.	korr. num. Antwort	dichotom/polytom	Punkte								Gesamtpunkte	Fragetypus	Fragenkategorie		
I. Setting 2												maximal				
Ambulant			maxim. Scoringwert	Itemanzahl												
Video																
A31	3		1	1									MC	Emotion		
A32	1		1	1									MC			
A34	f/f/r/r		2	1	1	1	1						T/F Aufgabe	Subfragestellungen	4 Unterfragen	
A35 neu	fa/ri/ri/fa		2	1	1	1	1						T/F Aufgabe	Subfragestellungen	4 Unterfragen	nicht in Antwo
A72	5		1	1									MC	Emotion		
A75	r/f/f/r/r/f		3	1	1	1	1	1	1				T/F Aufgabe	Subfragestellungen	6 Unterfragen	
Video																
A25	4		1	1									MC			
A26	4		1	1									MC			
A27	3		1	1									MC			
A28	f/r/f/f/f		3	1	1	1	1	1					T/C Aufgabe	Subfragestellungen	5 Unterfragen	
A29	1		1	1									MC	Emotion		
A30	1		1	1									MC	Emotion		
Video																
A37	4		1	1									MC			
A38	f/f/f/f/r		3	1	1	1	1	1					T/F Aufgabe	Emotion	5 Unterfragen	
A39	3		1	1									MC			
A40	f/f/f/f/f/r		2	1	1	1	1	1	1				T/F Aufgabe		6 Unterfragen	
A41	1		1	1									MC			
A42	1		1	1									MC			
A76	1		1	1									MC			
Video																
A43	2		1	1									MC	Emotion		
A44	f/r/f/f		2	1	1	1	1						T/F Aufgabe		4 Unterfragen	
A45	f/f/r/r/f		2	1	1	1	1	1					T/F Aufgabe		5 Unterfragen	
A47	4		1	1									MC			
A46	3		1	1									MC			
A48	7 Injektionsstellen		1	1									Image-Map	technische Frage	1 = ganz richtig	
														nicht in gelähmte Seite	0 = falsch	
														injizieren	2 = teilweise richtig?	
A77	4		1	1									MC			
A78	2		1	1									MC			

II. Setting 1																
Wohngruppe																
Video																
A1	1		1	1									MC			4 Antwortmöglichkeiten
A2	3		1	1									MC			4 Antwortmöglichkeiten
A3	4		1	1									MC			4 Antwortmöglichkeiten
A4	r/f/r/f/f	A und C richtig	2	1	1	1	1	1				2	T/F Aufgabe	polytom	5 Unterfragen	
A5	3		1	1									MC			5 Antwortmöglichkeiten
A73	2		1	1									MC			4 Antwortmöglichkeiten
A74	3		1	1									MC			5 Antwortmöglichkeiten
Video																
A6	4		1	1									MC	Emotion		4 Antwortmöglichkeiten
A7	2		1	1									MC			5 Antwortmöglichkeiten
A8	1		1	1									MC			4 Antwortmöglichkeiten
A9	3		1	1									MC			5 Antwortmöglichkeiten
A10	3		1	1									MC			4 Antwortmöglichkeiten
A11	4		1	1									MC			4 Antwortmöglichkeiten
A12	1		1	1									MC			4 Anwortmöglichkeiten
A71	3		1	1									MC			4 Antwortmöglichkeiten
Video																
A14	1		1	1									MC			4 Antwortmöglichkeiten
A15	3		1	1									MC			5 Antwortmöglichkeiten
A16	2		1	1									MC			5 Antwortmöglichkeiten
A17	r/f/r/f	A und C richtig	2	1	0	1	0					2	T/F Aufgabe	Emotion/Subfragestellungen/po	4 Unterfragen	
Video																
A18	3		1	1									MC			4 Antwortmöglichkeiten
A19	1		1	1									MC			3 Antwortmöglichkeiten
A20	3		1	1									MC			6 Antwortmöglichkeiten
Video																
A21	4		1	1									MC			4 Antwortmöglichkeiten
A22	2		1	1									MC			4 Antwortmöglichkeiten
A23	3		1	1									MC			5 Antwortmöglichkeiten
A24	1		1	1									MC	Emotion		4 Antwortmöglichkeiten

III. Setting 3																		
Stationär																		
Video																		
A53	4		1	1									MC			4 Antwortmöglichkeiten		
A54	1		1	1									MC	Emotion		4 Antwortmöglichkeiten		
A55	3		1	1									MC			5 Antwortmöglichkeiten		
A52	3		1	1									MC			4 Antwortmöglichkeiten		
A49	4		1	1									MC			7 Antwortmöglichkeiten		
A50	3		1	1									MC			4 Antwortmöglichkeiten		
A51	1		1	1									MC			4 Antwortmöglichkeiten		
A56	1		1	1									MC			5 Antwortmöglichkeiten		
A70	f/f/r/r		2	1	1	1	1						T/F Aufgabe	Subfragestellungen	4 Unterfragen			
A57	F,B,D,A,G,C,E		1	1									Drag and Drop	Reihenfolge festlegen	1 = richtig	7 Antwortmöglichkeiten		
															0 = falsch			
															2 = teilweise richtig?			
A79	1		1	1									Image-Map	Punkt benennen/Sigma	1 = richtig; 0 = falsch			
A80	3		1	1									MC			4 Antwortmöglichkeiten		
A81	3		1	1									MC			3 Antwortmöglichkeiten		
Video																		
A59	2		1	1									MC	Emotion		6 Antwortmöglichkeiten		
A60	3		1	1									MC	Emotion		4 Antwortmöglichkeiten		
A61	2		1	1									MC			4 Antwortmöglichkeiten		
A82	ri/ri/ri/ri/ri/fa/fa/ri		2	1	1	1	1	1	1	1	1		T/F Aufgabe		8 Unterfragen			
Video																		
A62	fa/ri/fa/ri		2	1	1	1	1						T/F Aufgabe		4 Unterfragen			
A63	1,2,3		2	1	1	1							MC (ordinal gescort)	Emotion		4 Antwortmöglichkeiten (drei richtig)		
A65	3		1	1									MC	Emotion		4 Antwortmöglichkeiten		
A66	4		1	1									MC			4 Antwortmöglichkeiten		
A67	3		1	1									MC			4 Antwortmöglichkeiten		
A68	1		1	1									MC			4 Antwortmöglichkeiten		
A69	fa/ri/ri/fa/fa		2	1	1	1	1	1					T/F Aufgabe		5 Unterfragen			
Abschluss																		
Gesamtpunkte:			95	77	14	15	13	8	3	1	1	4	55	gesamte Subantworten		76+54	130 Punkte gesamt	
			Scoring_gesamt	Items	Subantworten													

Cosmed_Antworten																			
A83	d,c,b,a,e		1	1									MC	Reihenfolge					
A84	4		1	1									MC			4 Antwortmöglichkeiten			
A85	3		1	1									MC			4 Antwortmöglichkeiten			
A86	1		1	1									MC			4 Antwortmöglichkeiten			
A87	1 und 2		2	1									MC			4 Antwortmöglichkeiten			
A88	2		1	1									MC			4 Antwortmöglichkeiten			
A89	2		1	1									MC			4 Antwortmöglichkeiten			
A90	3		1	1									MC			4 Antwortmöglichkeiten			
A91	3		1	1									MC			4 Antwortmöglichkeiten			
A92	2		1	1									MC			4 Antwortmöglichkeiten			
A93	2		1	1									MC			4 Antwortmöglichkeiten			
A94	1		1	1									MC			4 Antwortmöglichkeiten			
Abschluss																			
Gesamtpunkte:			13	12	Punkte									14 Emotionaspekte					
			Scoring_gesamt	Items															
														62 without Emotionaspekte					
Maximale Gesamtpunktzahl im Test (Pflege): 80?																			
MW: 50 Punkte														keine Hinweise auf Emotion (48)					
SD: 10 Punkte														(14) ?					
WLE-Werte: 45,5 vs. 56,0 > 1 SD auseinander																			
														gesamt: 76 Punkte im Text zu					
Varianzbestimmung														entnehmen					
Lebensalter und Testleistung der TN -> Pearsons r = 0.14, p = .015 / kleinster statistischer Effekt																			
62 stufiger computergestützter Test																			
	emotionale Items																		

Anhang 15: Ergebnisse zu Itemhäufigkeitsverteilungen in SPSS (ergänzend zu Kapitel 14.2)

A32

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	170	43,7	44,6	44,6
	1	211	54,2	55,4	100,0
	Gesamt	381	97,9	100,0	
Fehlend	System	8	2,1		
Gesamt		389	100,0		

A34_polytom

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	192	49,4	50,5	50,5
	1	53	13,6	13,9	64,5
	2	135	34,7	35,5	100,0
	Gesamt	380	97,7	100,0	
Fehlend	System	9	2,3		
Gesamt		389	100,0		

A35neu_polytom

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	315	81,0	82,0	82,0
	1	67	17,2	17,4	99,5
	2	2	,5	,5	100,0
	Gesamt	384	98,7	100,0	
Fehlend	System	5	1,3		
Gesamt		389	100,0		

A72

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	190	48,8	49,6	49,6
	1	193	49,6	50,4	100,0
	Gesamt	383	98,5	100,0	
Fehlend	System	6	1,5		
Gesamt		389	100,0		

A25

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	114	29,3	29,8	29,8
	1	268	68,9	70,2	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A26

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	54	13,9	14,2	14,2
	1	326	83,8	85,8	100,0
	Gesamt	380	97,7	100,0	
Fehlend	System	9	2,3		
Gesamt		389	100,0		

A27

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	10	2,6	2,6	2,6
	1	373	95,9	97,4	100,0
	Gesamt	383	98,5	100,0	
Fehlend	System	6	1,5		
Gesamt		389	100,0		

A28_polytom

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	371	95,4	96,9	96,9
	1	2	,5	,5	97,4
	2	1	,3	,3	97,7
	3	9	2,3	2,3	100,0
	Gesamt	383	98,5	100,0	
Fehlend	System	6	1,5		
Gesamt		389	100,0		

A29

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	223	57,3	58,4	58,4
	1	159	40,9	41,6	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A30

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	192	49,4	50,7	50,7
	1	187	48,1	49,3	100,0
	Gesamt	379	97,4	100,0	
Fehlend	System	10	2,6		
Gesamt		389	100,0		

A37

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	255	65,6	67,1	67,1
	1	125	32,1	32,9	100,0
	Gesamt	380	97,7	100,0	
Fehlend	System	9	2,3		
Gesamt		389	100,0		

A38_polytom

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	117	30,1	30,7	30,7
	1	154	39,6	40,4	71,1
	2	20	5,1	5,2	76,4
	3	90	23,1	23,6	100,0
	Gesamt	381	97,9	100,0	
Fehlend	System	8	2,1		
Gesamt		389	100,0		

A39_polytom

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	34	8,7	8,9	8,9
	1	347	89,2	90,8	99,7
	3	1	,3	,3	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A40_polytom

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	240	61,7	62,8	62,8
	1	40	10,3	10,5	73,3
	2	102	26,2	26,7	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A41

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	276	71,0	72,1	72,1
	1	107	27,5	27,9	100,0
	Gesamt	383	98,5	100,0	
Fehlend	System	6	1,5		
Gesamt		389	100,0		

A42

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	360	92,5	93,8	93,8
	1	24	6,2	6,3	100,0
	Gesamt	384	98,7	100,0	
Fehlend	System	5	1,3		
Gesamt		389	100,0		

A76

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	149	38,3	39,0	39,0
	1	233	59,9	61,0	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A43

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	171	44,0	44,9	44,9
	1	210	54,0	55,1	100,0
	Gesamt	381	97,9	100,0	
Fehlend	System	8	2,1		
Gesamt		389	100,0		

A44_polytom

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	39	10,0	10,3	10,3
	1	104	26,7	27,4	37,6
	2	237	60,9	62,4	100,0
	Gesamt	380	97,7	100,0	
Fehlend	System	9	2,3		
Gesamt		389	100,0		

A45_polytom

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	297	76,3	77,7	77,7
	1	7	1,8	1,8	79,6
	2	78	20,1	20,4	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A47

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	54	13,9	14,1	14,1
	1	329	84,6	85,9	100,0
	Gesamt	383	98,5	100,0	
Fehlend	System	6	1,5		
Gesamt		389	100,0		

A46

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	141	36,2	37,0	37,0
	1	240	61,7	63,0	100,0
	Gesamt	381	97,9	100,0	
Fehlend	System	8	2,1		
Gesamt		389	100,0		

A77

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	144	37,0	37,7	37,7
	1	238	61,2	62,3	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A78

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	93	23,9	24,9	24,9
	1	281	72,2	75,1	100,0
	Gesamt	374	96,1	100,0	
Fehlend	System	15	3,9		
Gesamt		389	100,0		

A01

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	102	26,2	26,7	26,7
	1	280	72,0	73,3	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A02

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	128	32,9	34,0	34,0
	1	249	64,0	66,0	100,0
	Gesamt	377	96,9	100,0	
Fehlend	System	12	3,1		
Gesamt		389	100,0		

A03

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	102	26,2	26,7	26,7
	1	280	72,0	73,3	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A04_polytom

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	337	86,6	88,2	88,2
	1	5	1,3	1,3	89,5
	2	40	10,3	10,5	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A05

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	246	63,2	65,1	65,1
	1	132	33,9	34,9	100,0
	Gesamt	378	97,2	100,0	
Fehlend	System	11	2,8		
Gesamt		389	100,0		

A73

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	237	60,9	63,5	63,5
	1	136	35,0	36,5	100,0
	Gesamt	373	95,9	100,0	
Fehlend	System	16	4,1		
Gesamt		389	100,0		

A74

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	196	50,4	51,4	51,4
	1	185	47,6	48,6	100,0
	Gesamt	381	97,9	100,0	
Fehlend	System	8	2,1		
Gesamt		389	100,0		

A06

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	206	53,0	53,9	53,9
	1	176	45,2	46,1	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A07

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	142	36,5	37,4	37,4
	1	238	61,2	62,6	100,0
	Gesamt	380	97,7	100,0	
Fehlend	System	9	2,3		
Gesamt		389	100,0		

A08

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	125	32,1	32,8	32,8
	1	256	65,8	67,2	100,0
	Gesamt	381	97,9	100,0	
Fehlend	System	8	2,1		
Gesamt		389	100,0		

A09

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	92	23,7	24,3	24,3
	1	286	73,5	75,7	100,0
	Gesamt	378	97,2	100,0	
Fehlend	System	11	2,8		
Gesamt		389	100,0		

A10

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	122	31,4	31,9	31,9
	1	261	67,1	68,1	100,0
	Gesamt	383	98,5	100,0	
Fehlend	System	6	1,5		
Gesamt		389	100,0		

A11

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	127	32,6	33,5	33,5
	1	252	64,8	66,5	100,0
	Gesamt	379	97,4	100,0	
Fehlend	System	10	2,6		
Gesamt		389	100,0		

A12

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	237	60,9	62,4	62,4
	1	143	36,8	37,6	100,0
	Gesamt	380	97,7	100,0	
Fehlend	System	9	2,3		
Gesamt		389	100,0		

A71

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	159	40,9	41,6	41,6
	1	223	57,3	58,4	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A14

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	157	40,4	41,3	41,3
	1	223	57,3	58,7	100,0
	Gesamt	380	97,7	100,0	
Fehlend	System	9	2,3		
Gesamt		389	100,0		

A15

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	116	29,8	30,8	30,8
	1	261	67,1	69,2	100,0
	Gesamt	377	96,9	100,0	
Fehlend	System	12	3,1		
Gesamt		389	100,0		

A16

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	193	49,6	51,1	51,1
	1	185	47,6	48,9	100,0
	Gesamt	378	97,2	100,0	
Fehlend	System	11	2,8		
Gesamt		389	100,0		

A17_polytom

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	106	27,2	28,2	28,2
	1	91	23,4	24,2	52,4
	2	179	46,0	47,6	100,0
	Gesamt	376	96,7	100,0	
Fehlend	System	13	3,3		
Gesamt		389	100,0		

A18

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	159	40,9	42,3	42,3
	1	217	55,8	57,7	100,0
	Gesamt	376	96,7	100,0	
Fehlend	System	13	3,3		
Gesamt		389	100,0		

A19

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	71	18,3	18,6	18,6
	1	310	79,7	81,4	100,0
	Gesamt	381	97,9	100,0	
Fehlend	System	8	2,1		
Gesamt		389	100,0		

A20

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	340	87,4	89,9	89,9
	1	38	9,8	10,1	100,0
	Gesamt	378	97,2	100,0	
Fehlend	System	11	2,8		
Gesamt		389	100,0		

A21

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	200	51,4	53,2	53,2
	1	176	45,2	46,8	100,0
	Gesamt	376	96,7	100,0	
Fehlend	System	13	3,3		
Gesamt		389	100,0		

A22

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	265	68,1	69,9	69,9
	1	114	29,3	30,1	100,0
	Gesamt	379	97,4	100,0	
Fehlend	System	10	2,6		
Gesamt		389	100,0		

A23

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	239	61,4	63,4	63,4
	1	138	35,5	36,6	100,0
	Gesamt	377	96,9	100,0	
Fehlend	System	12	3,1		
Gesamt		389	100,0		

A24

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	166	42,7	43,9	43,9
	1	212	54,5	56,1	100,0
	Gesamt	378	97,2	100,0	
Fehlend	System	11	2,8		
Gesamt		389	100,0		

A53

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	37	9,5	9,7	9,7
	1	344	88,4	90,3	100,0
	Gesamt	381	97,9	100,0	
Fehlend	System	8	2,1		
Gesamt		389	100,0		

A54

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	240	61,7	63,5	63,5
	1	138	35,5	36,5	100,0
	Gesamt	378	97,2	100,0	
Fehlend	System	11	2,8		
Gesamt		389	100,0		

A55

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	173	44,5	45,8	45,8
	1	205	52,7	54,2	100,0
	Gesamt	378	97,2	100,0	
Fehlend	System	11	2,8		
Gesamt		389	100,0		

A52

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	87	22,4	22,9	22,9
	1	293	75,3	77,1	100,0
	Gesamt	380	97,7	100,0	
Fehlend	System	9	2,3		
Gesamt		389	100,0		

A49

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	123	31,6	32,5	32,5
	1	255	65,6	67,5	100,0
	Gesamt	378	97,2	100,0	
Fehlend	System	11	2,8		
Gesamt		389	100,0		

A50

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	69	17,7	18,1	18,1
	1	313	80,5	81,9	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A51

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	180	46,3	47,2	47,2
	1	201	51,7	52,8	100,0
	Gesamt	381	97,9	100,0	
Fehlend	System	8	2,1		
Gesamt		389	100,0		

A56

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	113	29,0	29,7	29,7
	1	268	68,9	70,3	100,0
	Gesamt	381	97,9	100,0	
Fehlend	System	8	2,1		
Gesamt		389	100,0		

A70_polytom

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	239	61,4	63,1	63,1
	1	90	23,1	23,7	86,8
	2	50	12,9	13,2	100,0
	Gesamt	379	97,4	100,0	
Fehlend	System	10	2,6		
Gesamt		389	100,0		

A57

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	123	31,6	32,5	32,5
	1	255	65,6	67,5	100,0
	Gesamt	378	97,2	100,0	
Fehlend	System	11	2,8		
Gesamt		389	100,0		

A79

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	158	40,6	42,5	42,5
	1	214	55,0	57,5	100,0
	Gesamt	372	95,6	100,0	
Fehlend	System	17	4,4		
Gesamt		389	100,0		

A80

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	176	45,2	46,3	46,3
	1	204	52,4	53,7	100,0
	Gesamt	380	97,7	100,0	
Fehlend	System	9	2,3		
Gesamt		389	100,0		

A81

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	54	13,9	14,1	14,1
	1	328	84,3	85,9	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A59

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	197	50,6	52,0	52,0
	1	182	46,8	48,0	100,0
	Gesamt	379	97,4	100,0	
Fehlend	System	10	2,6		
Gesamt		389	100,0		

A60

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	213	54,8	55,8	55,8
	1	169	43,4	44,2	100,0
	Gesamt	382	98,2	100,0	
Fehlend	System	7	1,8		
Gesamt		389	100,0		

A61

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	295	75,8	78,7	78,7
	1	80	20,6	21,3	100,0
	Gesamt	375	96,4	100,0	
Fehlend	System	14	3,6		
Gesamt		389	100,0		

A82_polytom

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	292	75,1	76,8	76,8
	1	55	14,1	14,5	91,3
	2	24	6,2	6,3	97,6
	3	9	2,3	2,4	100,0
	Gesamt	380	97,7	100,0	
Fehlend	System	9	2,3		
Gesamt		389	100,0		

A62_polytom

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	270	69,4	71,1	71,1
	1	10	2,6	2,6	73,7
	2	100	25,7	26,3	100,0
	Gesamt	380	97,7	100,0	
Fehlend	System	9	2,3		
Gesamt		389	100,0		

A63_polytom_ordinal

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	148	38,0	39,2	39,2
	1	112	28,8	29,6	68,8
	2	29	7,5	7,7	76,5
	3	89	22,9	23,5	100,0
	Gesamt	378	97,2	100,0	
Fehlend	System	11	2,8		
Gesamt		389	100,0		

A65

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	81	20,8	21,3	21,3
	1	299	76,9	78,7	100,0
	Gesamt	380	97,7	100,0	
Fehlend	System	9	2,3		
Gesamt		389	100,0		

A66

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	253	65,0	67,3	67,3
	1	123	31,6	32,7	100,0
	Gesamt	376	96,7	100,0	
Fehlend	System	13	3,3		
Gesamt		389	100,0		

A67

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	249	64,0	65,7	65,7
	1	130	33,4	34,3	100,0
	Gesamt	379	97,4	100,0	
Fehlend	System	10	2,6		
Gesamt		389	100,0		

A68

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	161	41,4	42,6	42,6
	1	217	55,8	57,4	100,0
	Gesamt	378	97,2	100,0	
Fehlend	System	11	2,8		
Gesamt		389	100,0		

A69_polytom

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	0	173	44,5	46,0	46,0
	1	182	46,8	48,4	94,4
	2	21	5,4	5,6	100,0
	Gesamt	376	96,7	100,0	
Fehlend	System	13	3,3		
Gesamt		389	100,0		

Anhang 16: Veränderungen des Cronbachs Alpha und Item-Trennschärfe

Item-Skala-Statistiken in SPSS

	Skalenmittelwert, wenn Item wegge-lassen	Skalenvarianz, wenn Item wegge-lassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item wegge-lassen
DR21A31	43,09	98,332	,226	,783
IK21A32	43,21	100,338	,009	,788
DR21A34_polytom	42,93	95,126	,264	,782
PT21A35neu_polytom	43,57	99,359	,145	,785
IK21A72	43,26	97,971	,248	,782
DR21A75_polytom	43,44	97,201	,156	,786
PT22A25	43,07	99,073	,150	,785
PT22A26	42,92	99,143	,190	,784
IK22A27	42,80	99,126	,381	,783
IK22A28_polytom	43,68	99,766	,074	,786
DR22A29	43,35	98,119	,237	,783
IK22A30	43,28	99,406	,102	,786
PT23A37	43,43	99,911	,058	,787
DR23A38_polytom	42,56	94,345	,232	,784
IK23A39	42,86	98,211	,366	,782
PT23A40_polytom	43,13	95,215	,276	,781
DR23A41	43,48	98,828	,185	,784
DR23A42	43,69	100,630	-,003	,787
DR23A76	43,16	98,540	,194	,784
IK24A43	43,22	96,855	,363	,780
IK24A44_polytom	42,27	94,914	,383	,778
DR24A45_polytom	43,34	97,492	,160	,785
PT24A47	42,91	99,654	,123	,785
DR24A46	43,14	98,532	,197	,784
DR24A77	43,14	98,392	,211	,783
DR24A78	43,03	97,677	,316	,781
DR11A01	43,04	97,782	,303	,782
DR11A02	43,12	97,819	,276	,782
DR11A03	43,04	97,746	,307	,781
PT11A04_polytom	43,54	97,765	,208	,783
PT11A05	43,42	98,445	,213	,783
DR11A73	43,41	98,061	,252	,782
DR11A74	43,28	98,491	,195	,784
IK12A06	43,30	97,717	,275	,782
DR12A07	43,14	99,221	,125	,785

PT12A08	43,10	99,258	,126	,785
IK12A09	43,02	98,814	,190	,784
DR12A10	43,08	98,908	,165	,784
PT12A11	43,11	99,643	,084	,786
PT12A12	43,39	98,805	,171	,784
PT12A71	43,18	97,521	,297	,781
DR13A14	43,18	98,144	,233	,783
PT13A15	43,08	97,820	,283	,782
IK13A16	43,28	98,109	,234	,783
IK13A17_polytom	42,60	96,426	,207	,784
IK14A18	43,20	99,329	,111	,786
PT14A19	42,96	98,333	,273	,782
IK14A20	43,66	99,885	,118	,785
IK15A21	43,30	99,191	,124	,785
PT15A22	43,46	98,167	,255	,782
IK15A23	43,40	98,957	,156	,785
DR15A24	43,21	98,388	,206	,783
IK31A53	42,87	98,932	,257	,783
IK31A54	43,40	98,050	,252	,782
DR31A55	43,23	96,966	,351	,780
DR31A52	43,00	98,724	,206	,784
PT31A49	43,10	97,771	,285	,782
PT31A50	42,95	98,510	,255	,783
DR31A51	43,24	97,151	,332	,781
DR31A56	43,07	98,114	,255	,782
DR31A70_polytom	43,27	98,397	,125	,786
PT31A57	43,10	99,003	,153	,785
DR31A79	43,21	97,963	,250	,782
DR31A80	43,23	98,003	,245	,783
DR31A81	42,91	98,235	,320	,782
IK32A59	43,29	98,092	,236	,783
IK32A60	43,32	97,384	,311	,781
IK32A61	43,55	99,831	,084	,786
DR32A82_polytom	43,42	98,636	,112	,786
PT33A62_polytom	43,22	95,969	,230	,783
IK33A63_polytom_ordinal	42,63	100,599	-,056	,800
DR33A65	42,99	97,843	,318	,781
DR33A66	43,44	100,201	,028	,787
DR33A67	43,42	100,054	,042	,787
IK33A68	43,20	98,329	,213	,783
PT33A69_polytom	43,18	98,854	,124	,786

Anhang 17: Item-Mittelwerte in der Gesundheits- und Krankenpflege aus SPSS

Itemstatistiken

	N	Mittelwert	Std.-Abweichung
DR21A31	389	,66	,473
IK21A32	389	,54	,499
DR21A34_polytom	389	,83	,915
PT21A35neu_polytom	389	,18	,400
IK21A72	389	,50	,501
DR21A75_polytom	389	,31	,878
PT22A25	389	,69	,464
PT22A26	389	,84	,369
IK22A27	389	,96	,199
IK22A28_polytom	389	,08	,467
DR22A29	389	,41	,492
IK22A30	389	,48	,500
PT23A37	389	,32	,468
DR23A38_polytom	389	1,19	1,124
IK23A39	389	,90	,325
PT23A40_polytom	389	,63	,872
DR23A41	389	,28	,447
DR23A42	389	,06	,241
DR23A76	389	,60	,491
IK24A43	389	,54	,499
IK24A44_polytom	389	1,49	,706
DR24A45_polytom	389	,42	,804
PT24A47	389	,85	,362
DR24A46	389	,62	,487
DR24A77	389	,61	,488
DR24A78	389	,72	,448
DR11A01	389	,72	,450
DR11A02	389	,64	,481
DR11A03	389	,72	,450
PT11A04_polytom	389	,22	,614
PT11A05	389	,34	,474
DR11A73	389	,35	,477
DR11A74	389	,48	,500
IK12A06	389	,45	,498
DR12A07	389	,61	,488
PT12A08	389	,66	,475

IK12A09	389	,74	,442
DR12A10	389	,67	,470
PT12A11	389	,65	,478
PT12A12	389	,37	,483
PT12A71	389	,57	,495
DR13A14	389	,57	,495
PT13A15	389	,67	,470
IK13A16	389	,48	,500
IK13A17_polytom	389	1,15	,863
IK14A18	389	,56	,497
PT14A19	389	,80	,403
IK14A20	389	,10	,297
IK15A21	389	,45	,498
PT15A22	389	,29	,456
IK15A23	389	,35	,479
DR15A24	389	,54	,499
IK31A53	389	,88	,320
IK31A54	389	,35	,479
DR31A55	389	,53	,500
DR31A52	389	,75	,432
PT31A49	389	,66	,476
PT31A50	389	,80	,397
DR31A51	389	,52	,500
DR31A56	389	,69	,464
DR31A70_polytom	389	,49	,713
PT31A57	389	,66	,476
DR31A79	389	,55	,498
DR31A80	389	,52	,500
DR31A81	389	,84	,364
IK32A59	389	,47	,500
IK32A60	389	,43	,496
IK32A61	389	,21	,405
DR32A82_polytom	389	,33	,697
PT33A62_polytom	389	,54	,874
IK33A63_polytom_ordinal	389	1,12	1,178
DR33A65	389	,77	,422
DR33A66	389	,32	,466
DR33A67	389	,33	,472
IK33A68	389	,56	,497
PT33A69_polytom	389	,58	,594
Gültige Werte (Listenweise)	389		

Anhang 18: Entwickeltes R-Skript für die IRT-Berechnungen

```
### ----------------------------------------------------------------------------------------------------
```

R-Skript für IRT-Rasch-Modellierung_Ries Kompetenzmessung Dissertation

```
### ----------------------------------------------------------------------------------------------------

# Pakete laden

library(ltm)

library(TAM)

library(WrightMap)

library(foreign)

###

# Rasch-Modell

path <- "C:/Users/s.ries/Desktop/Ries2/"   # mein Pfad

data <- Ries2

# Deskriptive Itemstatistiken

descript(data)

mod1 <- tam.mml(data)

summary(mod1)

###

# Itemschwierigkeiten

mod1$xsi

mod1$xsi$xsi

mod1$xsi$se.xsi # SE der Itemschwierigkeiten

dfr <- mod1$xsi # Tabelle der Schwierigkeiten und deren SE

dfr$t <- dfr$xsi / dfr$se.xsi

write.csv2( dfr, "my_itemparameters.csv")

###
```

```
# Gesamtübersicht aller Items erzeugen

IRT.WrightMap(mod1, show.thr.lab = TRUE, label.items.rows = 3)

###

### Mehrparametrige Modelle: 2-PL Modell

mod2 <- tam.mml.2pl(data)

summary(mod2)

mod2$xsi # Aufgabenschwierigkeit

mod2$B   # Diskriminanz

###

# Modellvergleich

anova(mod1, mod2)

###

### Schätzung der Personenfertigkeit

# Expected a posteriori

# Prior-basierter Personenschätzer

head(mod1$person)

hist(mod1$person$EAP)

mean(mod1$person$EAP)

sd(mod1$person$EAP)

# Weighted Likelihood Estimate

# Likelihood-basierter Personenschätzer (geringerer Bias als EAPs)

ability <- tam.wle(mod1)

summary(ability)

WLE.ability <- ability$theta

head(WLE.ability)
```

```
hist(WLE.ability,las=1)

mean(WLE.ability)

sd(WLE.ability)


cor(mod1$person$EAP, WLE.ability)

plot(mod1$person$EAP, WLE.ability)

-----------------------------------------------------------------------------------------------------------------------

### Modell- und Itemfit

-----------------------------------------------------------------------------------------------------------------------

# Pakete laden

library(irtoys)

library(TAM)

library(haven)

# Pfad eingeben

Ries2 <- read_sav("C:/Users/s.ries/Desktop/Ries2.sav")

View(Ries2)


mod1 <- tam.mml(Ries2)

###

### Modellvergleiche

mod2 <- tam.mml.2pl(Ries2)

anova(mod1, mod2)

###

### Globaler Modelfit

summary(IRT.modelfit(mod1))

# maxX2: Pruefstatistik X2 Test auf globalen Modelfit
```

```
# SRMSR: Standardized mean square root of squared residuals

# MADaQ3: Mean of absolute values of centered Q_3 statistic

###

### Itemfit

# itemcharakteristische Funktionen und empirische Response-Kurven I

item <- "IK21A32"

tam.ItemDiff <- mod1$xsi$xsi

ability <- tam.wle(mod1)$theta


curve(exp(x-tam.ItemDiff[item]) / (1+ exp(x-tam.ItemDiff[item])), xlim=c(-3,3), ylim=c(0,1),
las=1, col="red",

    xlab="Latent Trait (logits)", ylab="Probability")

npp(Ries2, ability, items=item, bands=T, add=T)

ordner <- paste(getwd(), "/Plot_", sep="") pdf_name <- paste(paste(ordner, i,sep=""),".pdf",
sep="")

pdf(file=pdf_name)

plot(mod1, ask=FALSE)

dev.off()

plot(mod1, ask=FALSE) # expected Response-Kurven mit TAM


# itemcharakteristische Funktionen und empirische Response-Kurven II

item <- 3

tam.ItemDiff <- mod1$xsi$xsi

ability <- tam.wle(mod1)$theta


curve(exp(x-tam.ItemDiff[item]) / (1+ exp(x-tam.ItemDiff[item])), xlim=c(-3,3), ylim=c(0,1),
las=1, col="red",
```

```
xlab="Latent Trait (logits)", ylab="Probability")

npp(data, ability, items=item, bands=T, add=T)


plot(mod1, ask=FALSE) # expected Response-Kurven mit TAM

###

# Residuenbasierten (Infit/Outfit)

tam.fit <- tam.fit(mod1)


range(tam.fit$itemfit$Outfit)

range(tam.fit$itemfit$Infit)


crit.infit.min <- 0.8

crit.infit.max <- 1.2

crit.outfit.min <- 0.7

crit.outfit.max <- 1.3


par(mfrow=c(1,2))

plot(tam.ItemDiff, tam.fit$itemfit$Infit, xlab="Difficulty", ylab="Infit", main="Infit",
ylim=c(0.5,1.5))

abline(h=crit.infit.min,col="red")

abline(h=crit.infit.max,col="red")

write.csv2(tam.fit$itemfit$Infit,"infit_Rohwerte.csv")

plot(tam.ItemDiff, tam.fit$itemfit$Outfit, xlab="Difficulty", ylab="Outfit", main="Outfit",
ylim=c(0.5,1.5))

abline(h=crit.outfit.min,col="red")

abline(h=crit.outfit.max,col="red")

write.csv2(tam.fit$itemfit$Outfit,"Outfit_Rohwerte.csv")
```

```
# Ausgabe aller Infit- und Outfitwerte mit p-Werten

infit_wert <- tam.fit$itemfit$Infit

infit_p <- round(tam.fit$itemfit$Infit_p,3)

infit_code <- tam.fit$itemfit$parameter

write.csv2(data.frame(infit_code, infit_wert, infit_p), "infit_all.csv")

outfit_wert <- tam.fit$itemfit$Outfit

outfit_p <- round(tam.fit$itemfit$Outfit_p,3)

outfit_code <- tam.fit$itemfit$parameter

write.csv2(data.frame(outfit_code, outfit_wert, outfit_p), "outfit_all.csv")
```

Anhang 19: Itemparameter unter der Bedingung des PCM nach Ausgabe in R

	xsi	se.xsi	t
DR21A31_Cat1	-0,68215395	0,11015409	-6,19272496
IK21A32_Cat1	-0,14706561	0,10443681	-1,40817801
DR21A34_polytom_Cat1	1,314475	0,1076382	12,2119748
DR21A34_polytom_Cat2	-0,77868559	0,11265591	-6,91207076
PT21A35neu_polytom_Cat1	1,65715396	0,13461692	12,3101461
PT21A35neu_polytom_Cat2	3,77843757	0,71004131	5,32143342
IK21A72_Cat1	0,04815866	0,10399491	0,46308672
DR21A75_polytom_Cat1	3,96388516	0,16454225	24,0903793
DR21A75_polytom_Cat2	0,39252391	0,1771838	2,21534881
DR21A75_polytom_Cat3	-1,36011745	0,18989504	-7,16246972
PT22A25_Cat1	-0,80594682	0,11244462	-7,16749978
PT22A26_Cat1	-1,69834354	0,1409033	-12,0532559
IK22A27_Cat1	-3,2611744	0,25871507	-12,605274
IK22A28_polytom_Cat1	5,35222456	0,29926344	17,8846592
IK22A28_polytom_Cat2	0,97462721	0,32749656	2,97599218
IK22A28_polytom_Cat3	-1,77463881	0,34458555	-5,15006739
DR22A29_Cat1	0,41953181	0,10556518	3,97414949
IK22A30_Cat1	0,11306686	0,10404068	1,08675624
PT23A37_Cat1	0,81493892	0,11081126	7,35429682
DR23A38_polytom_Cat1	-0,2899602	0,11462176	-2,52971333
DR23A38_polytom_Cat2	2,14065844	0,12116061	17,6679404
DR23A38_polytom_Cat3	-1,25978908	0,12884582	-9,77749282
IK23A39_Cat1	-2,213588	0,1685887	-13,1301091
IK23A39_Cat2	12,2165965	1,00240974	12,1872284
IK23A39_Cat3	-5,90560989	1,00340659	-5,88556023
PT23A40_polytom_Cat1	1,83818793	0,11116443	16,5357565
PT23A40_polytom_Cat2	-0,74755526	0,12143566	-6,15597785
DR23A41_Cat1	1,0451723	0,11566176	9,03645525
DR23A42_Cat1	2,8377836	0,21186238	13,3944666
DR23A76_Cat1	-0,39057039	0,10621103	-3,67730531
IK24A43_Cat1	-0,13616359	0,1043892	-1,30438391
IK24A44_polytom_Cat1	-0,9613812	0,16232272	-5,9226532

IK24A44_polytom_Cat2	-0,77704377	0,10833471	-7,17262031
DR24A45_polytom_Cat1	3,82540318	0,12902501	29,6485399
DR24A45_polytom_Cat2	-2,18873533	0,13315385	-16,4376427
PT24A47_Cat1	-1,75907067	0,14368132	-12,2428631
DR24A46_Cat1	-0,47016984	0,10708849	-4,39047992
DR24A77_Cat1	-0,44729165	0,10682114	-4,1872952
DR24A78_Cat1	-0,97562952	0,11620228	-8,39595885
DR11A01_Cat1	-0,96216438	0,11587747	-8,30329108
DR11A02_Cat1	-0,57466746	0,10846614	-5,29812782
DR11A03_Cat1	-0,96216438	0,11587747	-8,30329108
PT11A04_polytom_Cat1	4,31968537	0,16353873	26,4138379
PT11A04_polytom_Cat2	-1,82243269	0,17221038	-10,5825949
PT11A05_Cat1	0,73012991	0,10936491	6,67608949
DR11A73_Cat1	0,68261164	0,10863246	6,2836804
DR11A74_Cat1	0,13472305	0,10407738	1,29445081
IK12A06_Cat1	0,23245812	0,10437678	2,22710571
DR12A07_Cat1	-0,44729165	0,10682114	-4,1872952
PT12A08_Cat1	-0,6579756	0,10975022	-5,99520993
IK12A09_Cat1	-1,04413465	0,11792707	-8,85407096
DR12A10_Cat1	-0,71876491	0,11079255	-6,48748446
PT12A11_Cat1	-0,61013181	0,10899249	-5,59792544
PT12A12_Cat1	0,60088483	0,10750177	5,58953426
PT12A71_Cat1	-0,27885783	0,10522767	-2,6500429
DR13A14_Cat1	-0,27885783	0,10522767	-2,6500429
PT13A15_Cat1	-0,71876491	0,11079255	-6,48748446
IK13A16_Cat1	0,13472305	0,10407738	1,29445081
IK13A17_polytom_Cat1	0,18179274	0,11642907	1,56140333
IK13A17_polytom_Cat2	-0,57137044	0,10696813	-5,34150189
IK14A18_Cat1	-0,21271324	0,10478097	-2,03007503
PT14A19_Cat1	-1,40799761	0,12919284	-10,898418
IK14A20_Cat1	2,33243796	0,17217176	13,5471576
IK15A21_Cat1	0,23245812	0,10437678	2,22710571
PT15A22_Cat1	0,95325982	0,11356049	8,39429131
IK15A23_Cat1	0,65908444	0,1082903	6,0862738
DR15A24_Cat1	-0,15797787	0,10448718	-1,51193544

IK31A53_Cat1	-2,10652165	0,1619468	-13,0074915
IK31A54_Cat1	0,65908444	0,1082903	6,0862738
DR31A55_Cat1	-0,08178672	0,10419226	-0,78495968
DR31A52_Cat1	-1,1436956	0,12065249	-9,47925435
PT31A49_Cat1	-0,64595202	0,10955463	-5,89616385
PT31A50_Cat1	-1,45878871	0,13105982	-11,130709
DR31A51_Cat1	-0,0384109	0,10408355	-0,36903909
DR31A56_Cat1	-0,80594682	0,11244462	-7,16749978
DR31A70_polytom_Cat1	1,05695957	0,10948728	9,65372028
DR31A70_polytom_Cat2	0,80081665	0,15568877	5,14370224
PT31A57_Cat1	-0,64595202	0,10955463	-5,89616385
DR31A79_Cat1	-0,17983543	0,10459628	-1,7193292
DR31A80_Cat1	-0,070934	0,10416104	-0,68100318
DR31A81_Cat1	-1,73856336	0,14273	-12,1807845
IK32A59_Cat1	0,1672416	0,10415266	1,60573522
IK32A60_Cat1	0,30898666	0,10476474	2,94933823
IK32A61_Cat1	1,44044455	0,12734467	11,3113844
DR32A82_polytom_Cat1	1,7676943	0,12505869	14,1349181
DR32A82_polytom_Cat2	1,06685932	0,18733727	5,69485896
DR32A82_polytom_Cat3	1,360311	0,3423538	3,97340702
PT33A62_polytom_Cat1	3,36244819	0,11929455	28,1861004
PT33A62_polytom_Cat2	-2,10058266	0,12290602	-17,0909669
IK33A63_polytom_ordinal_Cat1	0,28743419	0,10995151	2,61419047
IK33A63_polytom_ordinal_Cat2	1,46112548	0,11949444	12,22756
IK33A63_polytom_ordinal_Cat3	-0,86950983	0,12971665	-6,70314746
DR33A65_Cat1	-1,23294298	0,12332066	-9,99786268
DR33A66_Cat1	0,83959732	0,11126563	7,54588217
DR33A67_Cat1	0,75413633	0,10975614	6,87101743
IK33A68_Cat1	-0,21271324	0,10478097	-2,03007503
PT33A69_polytom_Cat1	0,04047404	0,10454216	0,38715519
PT33A69_polytom_Cat2	2,36059387	0,22653449	10,4204612

Item Parameters Xsi-> Mod1 - PCM (xsi=Itemparameter, se.xsi=Standardfehler)

	xsi	se.xsi
DR21A31_Cat1	-0.682	0.110
IK21A32_Cat1	-0.147	0.104
DR21A34_polytom_Cat1	1.314	0.108
DR21A34_polytom_Cat2	-0.779	0.113
PT21A35neu_polytom_Cat1	1.657	0.135
PT21A35neu_polytom_Cat2	3.778	0.710
IK21A72_Cat1	0.048	0.104
DR21A75_polytom_Cat1	3.964	0.165
DR21A75_polytom_Cat2	0.393	0.177
DR21A75_polytom_Cat3	-1.360	0.190
PT22A25_Cat1	-0.806	0.112
PT22A26_Cat1	-1.698	0.141
IK22A27_Cat1	-3.261	0.259
IK22A28_polytom_Cat1	5.352	0.299
IK22A28_polytom_Cat2	0.975	0.327
IK22A28_polytom_Cat3	-1.775	0.345
DR22A29_Cat1	0.420	0.106
IK22A30_Cat1	0.113	0.104
PT23A37_Cat1	0.815	0.111
DR23A38_polytom_Cat1	-0.290	0.115
DR23A38_polytom_Cat2	2.141	0.121
DR23A38_polytom_Cat3	-1.260	0.129
IK23A39_Cat1	-2.214	0.169
IK23A39_Cat2	12.217	1.002
IK23A39_Cat3	-5.906	1.003
PT23A40_polytom_Cat1	1.838	0.111
PT23A40_polytom_Cat2	-0.748	0.121
DR23A41_Cat1	1.045	0.116
DR23A42_Cat1	2.838	0.212
DR23A76_Cat1	-0.391	0.106
IK24A43_Cat1	-0.136	0.104
IK24A44_polytom_Cat1	-0.961	0.162
IK24A44_polytom_Cat2	-0.777	0.108
DR24A45_polytom_Cat1	3.825	0.129
DR24A45_polytom_Cat2	-2.189	0.133
PT24A47_Cat1	-1.759	0.144
DR24A46_Cat1	-0.470	0.107
DR24A77_Cat1	-0.447	0.107
DR24A78_Cat1	-0.976	0.116
DR11A01_Cat1	-0.962	0.116
DR11A02_Cat1	-0.575	0.108
DR11A03_Cat1	-0.962	0.116
PT11A04_polytom_Cat1	4.320	0.164
PT11A04_polytom_Cat2	-1.822	0.172
PT11A05_Cat1	0.730	0.109
DR11A73_Cat1	0.683	0.109
DR11A74_Cat1	0.135	0.104
IK12A06_Cat1	0.232	0.104
DR12A07_Cat1	-0.447	0.107
PT12A08_Cat1	-0.658	0.110
IK12A09_Cat1	-1.044	0.118
DR12A10_Cat1	-0.719	0.111
PT12A11_Cat1	-0.610	0.109
PT12A12_Cat1	0.601	0.108
PT12A71_Cat1	-0.279	0.105
DR13A14_Cat1	-0.279	0.105

PT13A15_Cat1	-0.719	0.111
IK13A16_Cat1	0.135	0.104
IK13A17_polytom_Cat1	0.182	0.116
IK13A17_polytom_Cat2	-0.571	0.107
IK14A18_Cat1	-0.213	0.105
PT14A19_Cat1	-1.408	0.129
IK14A20_Cat1	2.332	0.172
IK15A21_Cat1	0.232	0.104
PT15A22_Cat1	0.953	0.114
IK15A23_Cat1	0.659	0.108
DR15A24_Cat1	-0.158	0.104
IK31A53_Cat1	-2.107	0.162
IK31A54_Cat1	0.659	0.108
DR31A55_Cat1	-0.082	0.104
DR31A52_Cat1	-1.144	0.121
PT31A49_Cat1	-0.646	0.110
PT31A50_Cat1	-1.459	0.131
DR31A51_Cat1	-0.038	0.104
DR31A56_Cat1	-0.806	0.112
DR31A70_polytom_Cat1	1.057	0.109
DR31A70_polytom_Cat2	0.801	0.156
PT31A57_Cat1	-0.646	0.110
DR31A79_Cat1	-0.180	0.105
DR31A80_Cat1	-0.071	0.104
DR31A81_Cat1	-1.739	0.143
IK32A59_Cat1	0.167	0.104
IK32A60_Cat1	0.309	0.105
IK32A61_Cat1	1.440	0.127
DR32A82_polytom_Cat1	1.768	0.125
DR32A82_polytom_Cat2	1.067	0.187
DR32A82_polytom_Cat3	1.360	0.342
PT33A62_polytom_Cat1	3.362	0.119
PT33A62_polytom_Cat2	-2.101	0.123
IK33A63_polytom_ordinal_Cat1	0.287	0.110
IK33A63_polytom_ordinal_Cat2	1.461	0.119
IK33A63_polytom_ordinal_Cat3	-0.870	0.130
DR33A65_Cat1	-1.233	0.123
DR33A66_Cat1	0.840	0.111
DR33A67_Cat1	0.754	0.110
IK33A68_Cat1	-0.213	0.105
PT33A69_polytom_Cat1	0.040	0.105
PT33A69_polytom_Cat2	2.361	0.227

Anhang 20: Itemparameter unter der Bedingung des GPCM in R

```
> mod2$xsi -> GPCM -> Aufgabenschwierigkeit und Standardfehler

                                    xsi    se.xsi
DR21A31_Cat1                 -0.71604277 0.1107437
IK21A32_Cat1                 -0.17008728 0.1017716
DR21A34_polytom_Cat1          1.34015905 0.1070505
DR21A34_polytom_Cat2         -0.91458407 0.1100287
PT21A35neu_polytom_Cat1       1.62000830 0.1343719
PT21A35neu_polytom_Cat2       3.66473780 0.7097235
IK21A72_Cat1                  0.02545876 0.1053830
DR21A75_polytom_Cat1          3.94657687 0.1577555
DR21A75_polytom_Cat2          0.19811915 0.1691149
DR21A75_polytom_Cat3         -1.63814294 0.1829285
PT22A25_Cat1                 -0.81487116 0.1109294
PT22A26_Cat1                 -1.73350736 0.1409666
IK22A27_Cat1                 -4.23674949 0.3020042
IK22A28_polytom_Cat1          5.24234427 0.2943067
IK22A28_polytom_Cat2          0.99810667 0.3218367
IK22A28_polytom_Cat3         -2.27579939 0.3392299
DR22A29_Cat1                  0.40533396 0.1068171
IK22A30_Cat1                  0.07782223 0.1018484
PT23A37_Cat1                  0.75006301 0.1087337
DR23A38_polytom_Cat1         -0.28596340 0.1147509
DR23A38_polytom_Cat2          2.06171777 0.1138497
DR23A38_polytom_Cat3         -1.51989269 0.1211871
IK23A39_Cat1                 -2.68088643 0.1847426
IK23A39_Cat2                 12.18936067 1.0005184
IK23A39_Cat3                 -6.31591259 1.0013879
PT23A40_polytom_Cat1          1.79485016 0.1114223
PT23A40_polytom_Cat2         -0.73532642 0.1227177
DR23A41_Cat1                  1.04618571 0.1169153
DR23A42_Cat1                  2.72309070 0.2107462
DR23A76_Cat1                 -0.41985873 0.1066018
IK24A43_Cat1                 -0.16611866 0.1138959
IK24A44_polytom_Cat1         -1.28268280 0.1716523
IK24A44_polytom_Cat2         -0.78283913 0.1102653
DR24A45_polytom_Cat1          3.80309746 0.1241231
DR24A45_polytom_Cat2         -2.38337628 0.1284519
PT24A47_Cat1                 -1.74393997 0.1417565
DR24A46_Cat1                 -0.50346954 0.1079833
DR24A77_Cat1                 -0.48027008 0.1077290
DR24A78_Cat1                 -1.09056517 0.1218226
DR11A01_Cat1                 -1.04292498 0.1193512
DR11A02_Cat1                 -0.62208973 0.1108070
DR11A03_Cat1                 -1.05457278 0.1201159
PT11A04_polytom_Cat1          5.13363378 0.1654589
PT11A04_polytom_Cat2         -2.71389097 0.1719133
PT11A05_Cat1                  0.72408769 0.1106556
DR11A73_Cat1                  0.72866858 0.1132124
DR11A74_Cat1                  0.10987194 0.1045628
IK12A06_Cat1                  0.22436953 0.1071332
DR12A07_Cat1                 -0.46084333 0.1047588
PT12A08_Cat1                 -0.67299245 0.1085012
IK12A09_Cat1                 -1.07247447 0.1178801
DR12A10_Cat1                 -0.73539609 0.1097850
PT12A11_Cat1                 -0.61802412 0.1069247
```

PT12A12_Cat1	0.58666136	0.1084472
PT12A71_Cat1	-0.31706245	0.1088600
DR13A14_Cat1	-0.31198761	0.1071696
PT13A15_Cat1	-0.78340253	0.1141274
IK13A16_Cat1	0.11436965	0.1054705
IK13A17_polytom_Cat1	0.22415107	0.1133565
IK13A17_polytom_Cat2	-0.67837566	0.1028066
IK14A18_Cat1	-0.23487338	0.1027212
PT14A19_Cat1	-1.50539241	0.1323575
IK14A20_Cat1	2.32098288	0.1725182
IK15A21_Cat1	0.19393401	0.1025498
PT15A22_Cat1	0.99398395	0.1167583
IK15A23_Cat1	0.62479153	0.1078590
DR15A24_Cat1	-0.18446986	0.1038927
IK31A53_Cat1	-2.27790926	0.1671495
IK31A54_Cat1	0.69474914	0.1123489
DR31A55_Cat1	-0.10272774	0.1118662
DR31A52_Cat1	-1.19176031	0.1217037
PT31A49_Cat1	-0.70207509	0.1124688
PT31A50_Cat1	-1.57440489	0.1349697
DR31A51_Cat1	-0.05634009	0.1096180
DR31A56_Cat1	-0.87340774	0.1155961
DR31A70_polytom_Cat1	1.04294143	0.1064369
DR31A70_polytom_Cat2	0.55917994	0.1515148
PT31A57_Cat1	-0.66066545	0.1082534
DR31A79_Cat1	-0.20923185	0.1065094
DR31A80_Cat1	-0.09586091	0.1064981
DR31A81_Cat1	-1.95918208	0.1504459
IK32A59_Cat1	0.14889224	0.1057439
IK32A60_Cat1	0.33391569	0.1108124
IK32A61_Cat1	1.39195177	0.1267380
DR32A82_polytom_Cat1	1.70580063	0.1217048
DR32A82_polytom_Cat2	0.82715041	0.1825797
DR32A82_polytom_Cat3	1.23827906	0.3395531
PT33A62_polytom_Cat1	3.35728438	0.1188472
PT33A62_polytom_Cat2	-2.13661427	0.1237276
IK33A63_polytom_ordinal_Cat1	0.34579232	0.1034748
IK33A63_polytom_ordinal_Cat2	1.39341140	0.1108896
IK33A63_polytom_ordinal_Cat3	-1.15842475	0.1209009
DR33A65_Cat1	-1.38094789	0.1295152
DR33A66_Cat1	0.77279949	0.1091377
DR33A67_Cat1	0.68930948	0.1074889
IK33A68_Cat1	-0.23847598	0.1041264
PT33A69_polytom_Cat1	0.01848948	0.1021045
PT33A69_polytom_Cat2	2.22066777	0.2251262

Anhang 21: Diskriminationsparameter im GPCM = Mod2 in “TAM”

```
> mod2$B -> Diskriminationsparameter („lambda“) des GPCM = Mod2

, , Dim01

                        Cat0         Cat1       Cat2       Cat3
DR21A31                    0  0.550495518  0.0000000  0.0000000
IK21A32                    0  0.007571891  0.0000000  0.0000000
DR21A34_polytom            0  0.705574699  0.7708023  0.0000000
PT21A35neu_polytom         0  0.436177486  0.7775528  0.0000000
IK21A72                    0  0.615088563  0.0000000  0.0000000
DR21A75_polytom            0  0.425117942 -0.3795615  0.7468584
PT22A25                    0  0.340112443  0.0000000  0.0000000
PT22A26                    0  0.529175315  0.0000000  0.0000000
IK22A27                    0  1.556528129  0.0000000  0.0000000
IK22A28_polytom            0 -0.068695848 -0.6739307  0.7409547
DR22A29                    0  0.602811086  0.0000000  0.0000000
IK22A30                    0  0.172926017  0.0000000  0.0000000
PT23A37                    0  0.119212695  0.0000000  0.0000000
DR23A38_polytom            0  0.751720267  0.8369310  0.7733229
IK23A39                    0  1.300987725  1.1211077  1.4322598
PT23A40_polytom            0  0.381867684  1.0138749  0.0000000
DR23A41                    0  0.607206966  0.0000000  0.0000000
DR23A42                    0  0.055225807  0.0000000  0.0000000
DR23A76                    0  0.527816837  0.0000000  0.0000000
IK24A43                    0  1.200286031  0.0000000  0.0000000
IK24A44_polytom            0  0.968223313  1.4901882  0.0000000
DR24A45_polytom            0 -0.181771433  0.4560904  0.0000000
PT24A47                    0  0.349483345  0.0000000  0.0000000
DR24A46                    0  0.574641429  0.0000000  0.0000000
DR24A77                    0  0.575549949  0.0000000  0.0000000
DR24A78                    0  0.922511435  0.0000000  0.0000000
DR11A01                    0  0.776666395  0.0000000  0.0000000
DR11A02                    0  0.695067865  0.0000000  0.0000000
DR11A03                    0  0.830077665  0.0000000  0.0000000
PT11A04_polytom            0  1.654135661  0.9287104  0.0000000
PT11A05                    0  0.607838668  0.0000000  0.0000000
DR11A73                    0  0.880914767  0.0000000  0.0000000
DR11A74                    0  0.530647715  0.0000000  0.0000000
IK12A06                    0  0.731731349  0.0000000  0.0000000
DR12A07                    0  0.240976346  0.0000000  0.0000000
PT12A08                    0  0.365344031  0.0000000  0.0000000
IK12A09                    0  0.501596423  0.0000000  0.0000000
DR12A10                    0  0.395463965  0.0000000  0.0000000
PT12A11                    0  0.250831602  0.0000000  0.0000000
PT12A12                    0  0.574037166  0.0000000  0.0000000
PT12A71                    0  0.795883846  0.0000000  0.0000000
DR13A14                    0  0.662933249  0.0000000  0.0000000
PT13A15                    0  0.769700032  0.0000000  0.0000000
IK13A16                    0  0.615439874  0.0000000  0.0000000
IK13A17_polytom            0  0.543693421  0.5320454  0.0000000
IK14A18                    0  0.225820937  0.0000000  0.0000000
PT14A19                    0  0.746727568  0.0000000  0.0000000
IK14A20                    0  0.518893244  0.0000000  0.0000000
IK15A21                    0  0.239203687  0.0000000  0.0000000
PT15A22                    0  0.779114567  0.0000000  0.0000000
IK15A23                    0  0.424635958  0.0000000  0.0000000
```

DR15A24	0	0.423220721	0.0000000	0.0000000
IK31A53	0	0.829367161	0.0000000	0.0000000
IK31A54	0	0.840825791	0.0000000	0.0000000
DR31A55	0	1.082623377	0.0000000	0.0000000
DR31A52	0	0.595720675	0.0000000	0.0000000
PT31A49	0	0.738948763	0.0000000	0.0000000
PT31A50	0	0.793652611	0.0000000	0.0000000
DR31A51	0	0.937370001	0.0000000	0.0000000
DR31A56	0	0.755365914	0.0000000	0.0000000
DR31A70_polytom	0	0.346318444	0.1357696	0.0000000
PT31A57	0	0.358597805	0.0000000	0.0000000
DR31A79	0	0.660483712	0.0000000	0.0000000
DR31A80	0	0.695725165	0.0000000	0.0000000
DR31A81	0	0.995938005	0.0000000	0.0000000
IK32A59	0	0.633027402	0.0000000	0.0000000
IK32A60	0	0.979567995	0.0000000	0.0000000
IK32A61	0	0.382311429	0.0000000	0.0000000
DR32A82_polytom	0	0.174243727	0.1542865	0.8328820
PT33A62_polytom	0	-0.116457604	0.9682240	0.0000000
IK33A63_polytom_ordinal	0	-0.075762888	-0.3854703	-0.2066932
DR33A65	0	0.948015359	0.0000000	0.0000000
DR33A66	0	0.092976956	0.0000000	0.0000000
DR33A67	0	0.014664781	0.0000000	0.0000000
IK33A68	0	0.417967752	0.0000000	0.0000000
PT33A69_polytom	0	0.193517620	0.4906405	0.0000000

Anhang 22: Entwickeltes R-Skript in „eRm“ u.a. zur globalen Modelgeltung

###----------

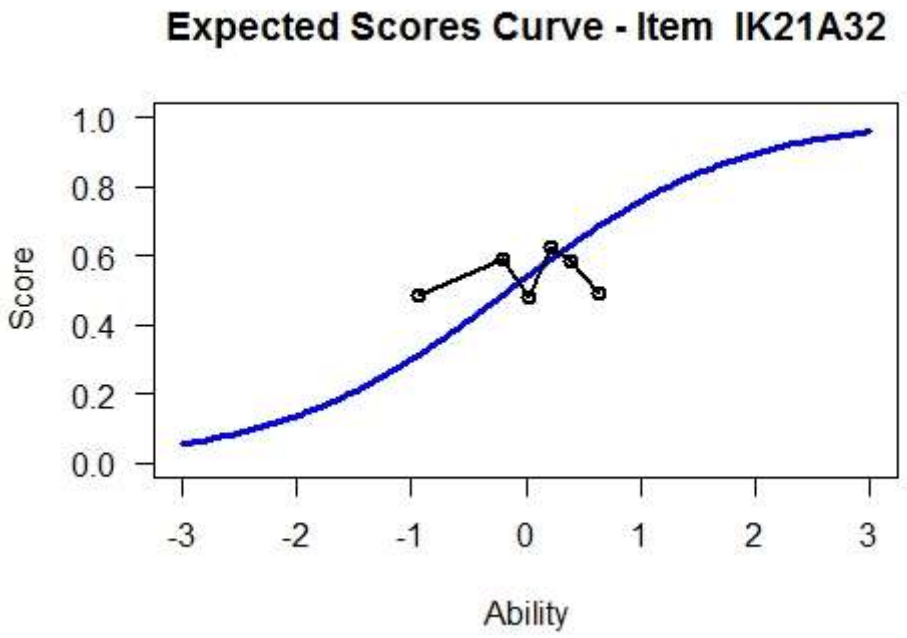
R-Skript in „eRm“ u.a. für die Berechnung des Likelihood-Quotienten_Ries

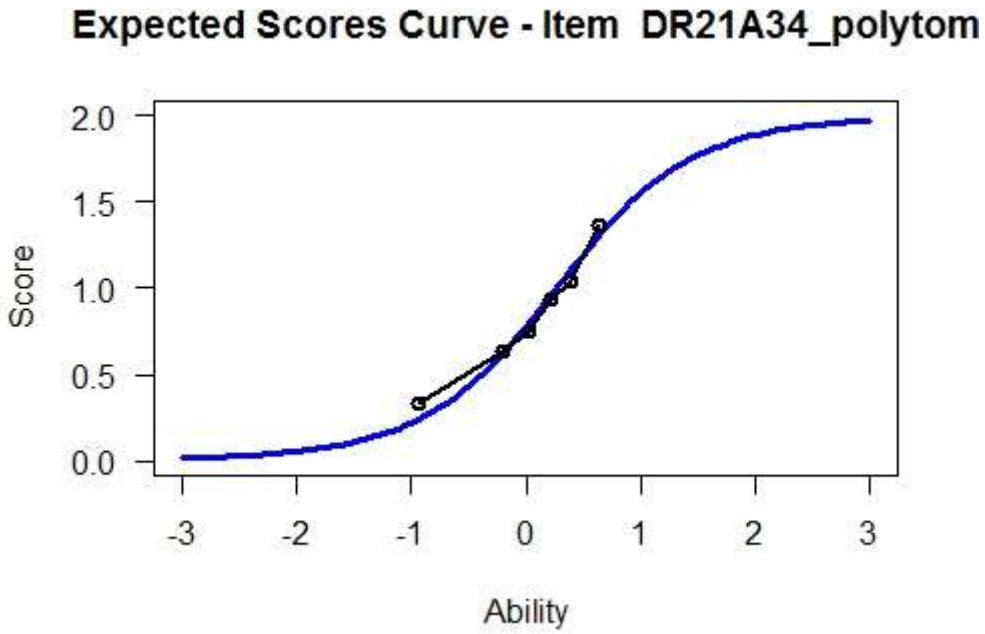
###----------

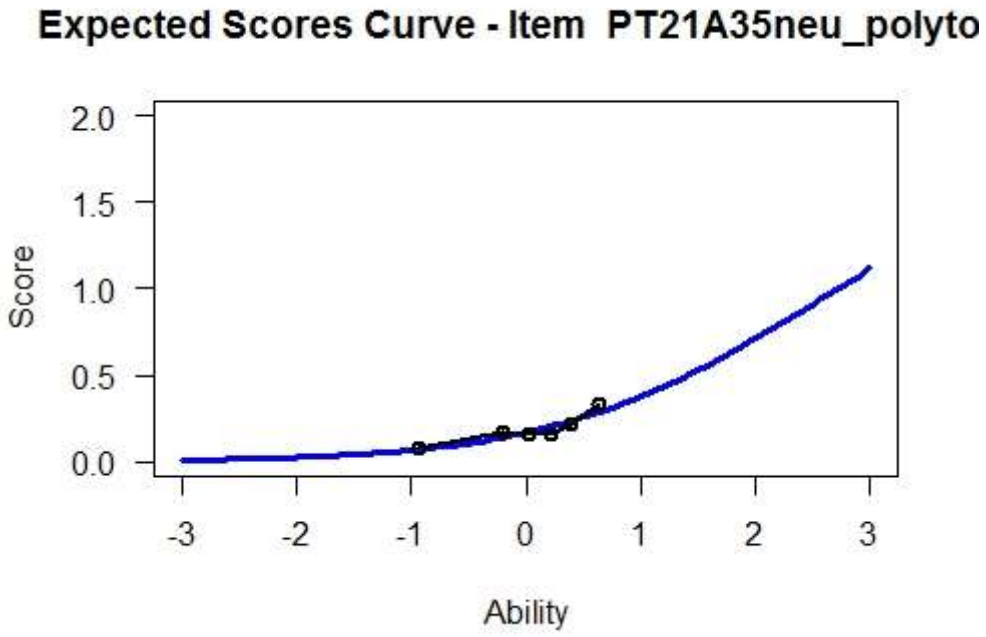
```
# Einlesen der Daten
library(haven)
Ries2 <- read_sav("C:/Users/Administrator/Downloads/Ries2.sav")
#
# installiere Paket
install.packages("eRm")
# lade Paket
library(eRm)

# Partial Credit Modell
# Umkodierung Person 313
Ries2$IK23A39[313] <- 1
## Berechnung des Modells (Partial Credit Modell)
out2 <- PCM(Ries2)
summary(out2)#
# Alles in eine Tabelle
write.csv2(data.frame(
  out2$etapar,
  out2$se.eta), "Ausgabe_Difficutly_eRm.csv")

## Likelihoodratio-Test bzw. Likelihood-Quotienten-Test
out2_lrt <- LRtest(out2)
out2_lrt

# Grafik mit Diagonale
# Nummer
plotGOF(out2_lrt,tlab = "number")
# nur Punkte
plotGOF(out2_lrt,tlab = "none")
# Waldtest
wald_out2 <- Waldtest(out2)
wald_out2
write.csv2(wald_out2$coef.table, "wald_test_pro_Item.csv" )
####################
# Item und Personenspezifika
pp <- person.parameter(out2)
##### Itemfit Statistics:
t1 <- itemfit(pp)
p_werte_i <- 1-pchisq(t1$i.fit, t1$i.df-1)
item_stat <- round(data.frame(t1$i.fit,t1$i.df, p_werte_i, t1$i.outfitMSQ,t1$i.infitMSQ,
                  t1$i.outfitZ, t1$i.infitZ ),5)
colnames(item_stat) <- c("Chi2", "df", "p-Wert", "Outfit_MS", "Infit_MS",
                         "Outfit_Z", "Infit_Z")
```

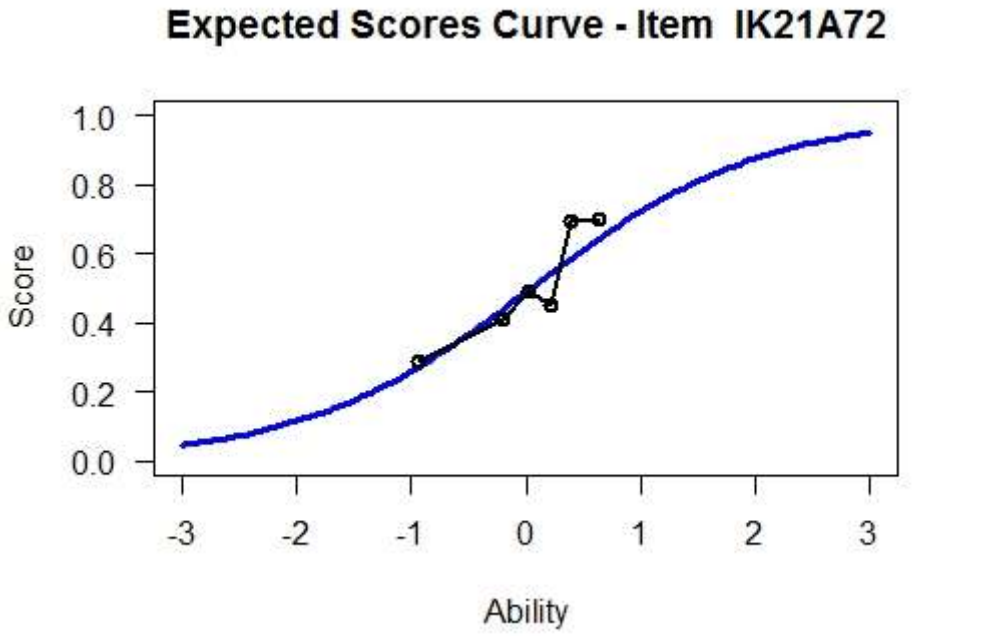
```
write.csv2(item_stat, "item_statistics_eRm.csv")
##### Person Statistics:
p1 <- personfit(pp)
plot(pp)
p_werte_p <- 1-pchisq(p1$p.fit, p1$p.df-1)
person_stat <- round(data.frame(p1$p.fit,p1$p.df-1, p_werte_p, p1$p.outfitMSQ,p1$p.infit
                    MSQ, p1$p.outfitZ, p1$p.infitZ ),5)
colnames(person_stat) <- c("Chi2", "df", "p-Wert", "Outfit_MS", "Infit_MS",
                                  "Outfit_Z", "Infit_Z")
write.csv2(person_stat, "person_statistics_eRm.csv")
########## AIC BIC
IC(pp)
plot(pp)
```

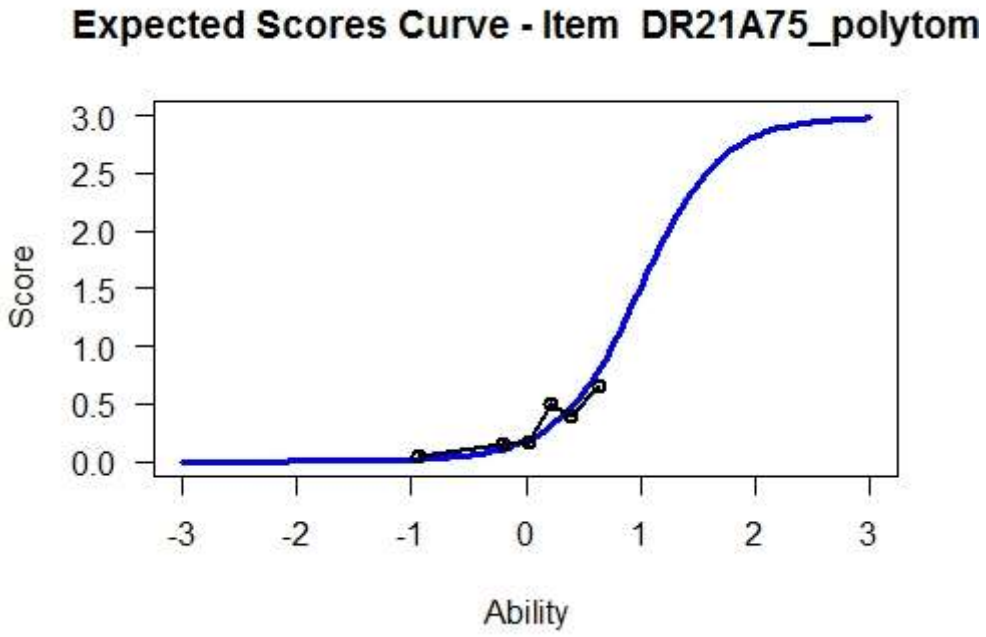
Anhang 23: Itemcharakteristische Funktionen im 1 PL-Modell in „TAM“ und „irtoys“ mit beobachteten und erwarteten Lösungshäufigkeiten (weitere Beispiele)

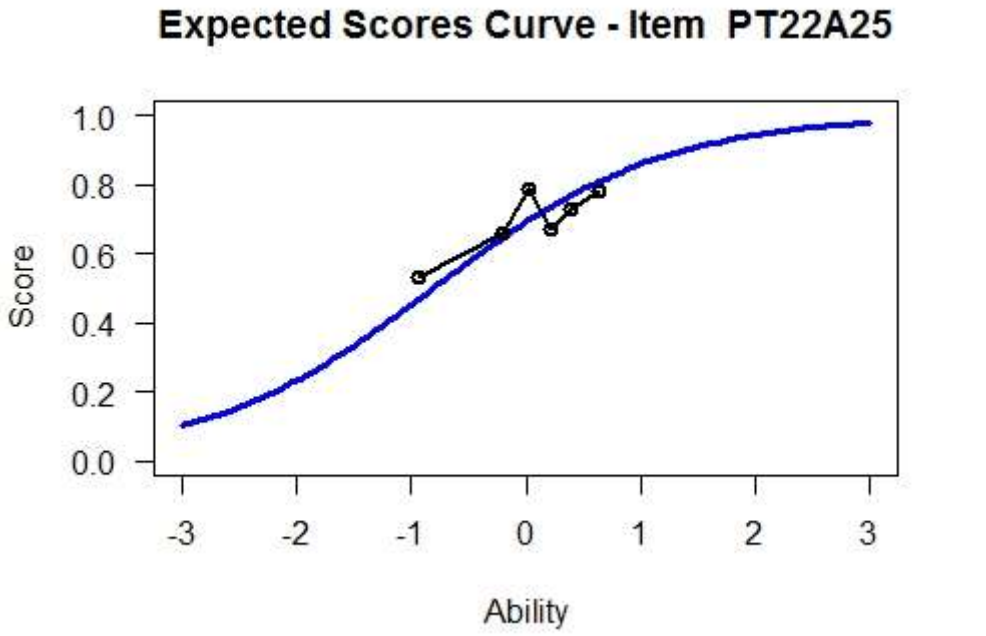


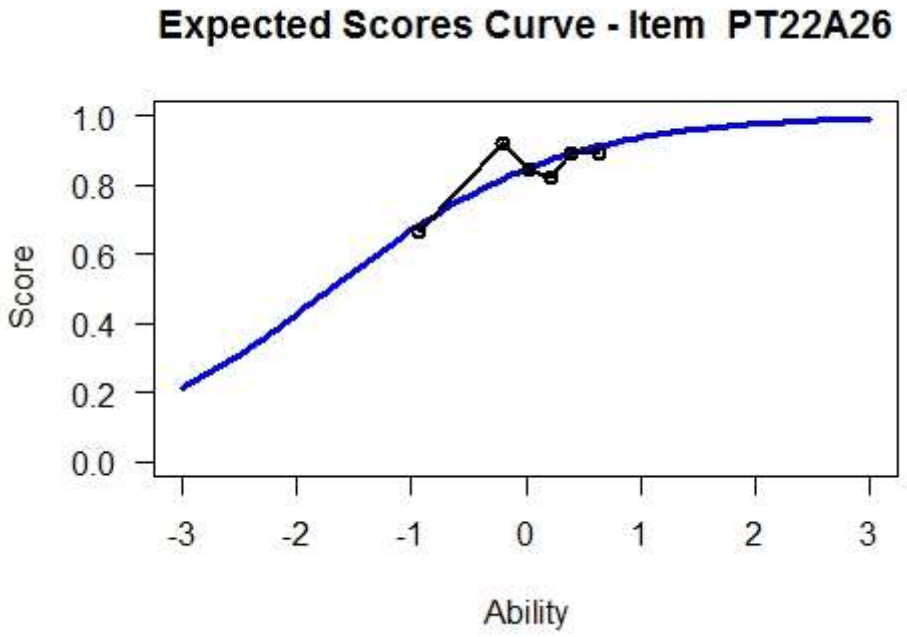


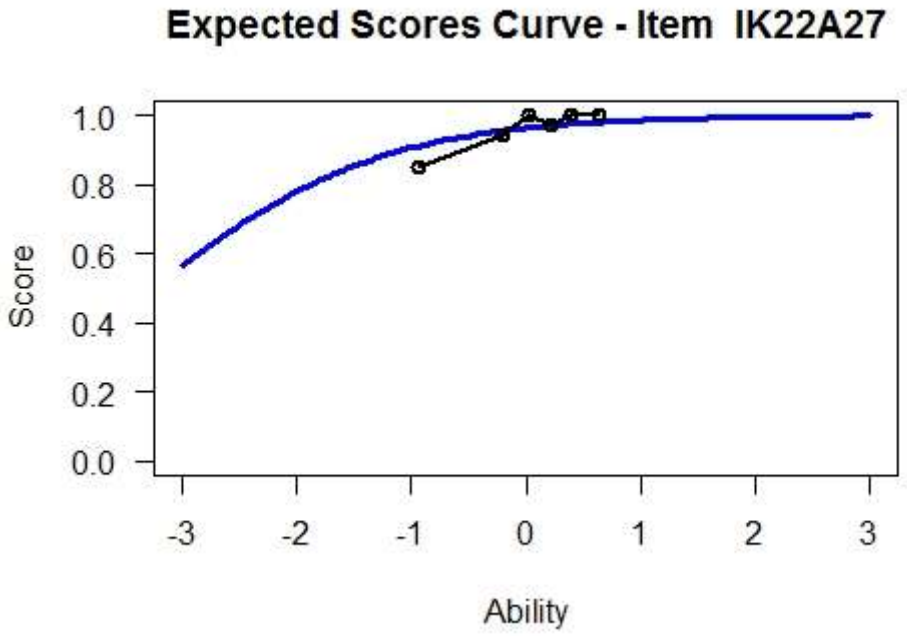

Expected Scores Curve - Item PT21A35neu_polyto
Score
2.0
1.5
1.0
0.5
0.0
-3
-2
-1
0
1
2
3
Ability


Expected Scores Curve - Item IK21A72
Score
1.0
0.8
0.6
0.4
0.2
0.0
-3
-2
-1
0
1
2
3
Ability


Expected Scores Curve - Item DR21A75_polytom
Score
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-3
-2
-1
0
1
2
3
Ability


Expected Scores Curve - Item PT22A25
Score
1.0
0.8
0.6
0.4
0.2
0.0
-3
-2
-1
0
1
2
3
Ability





Erläuterung: Die schwarze Kurve zeigt die beobachteten Häufigkeiten und die blaue Kurve die erwarteten Häufigkeiten in den aufgeführten graphischen „Item-Fit"-Beispielen.

© Ries - Abbildungen aus R Skript „IRT"

Anhang 24: Infit- und Outfitwerte einschließlich p-Werten

infit_code	infit_wert	infit_p
DR21A31_Cat1	0,98622442	0,741
IK21A32_Cat1	1,08423876	0
DR21A34_polytom_Cat1	1,02400817	0,449
DR21A34_polytom_Cat2	1,04270998	0,31
PT21A35neu_polytom_Cat1	0,99509947	0,975
PT21A35neu_polytom_Cat2	1,00536675	0,81
IK21A72_Cat1	0,97497525	0,23
DR21A75_polytom_Cat1	1,1054188	0,335
DR21A75_polytom_Cat2	1,11697156	0,332
DR21A75_polytom_Cat3	1,08244439	0,524
PT22A25_Cat1	1,01711272	0,701
PT22A26_Cat1	0,97949287	0,848
IK22A27_Cat1	0,87943755	0,637
IK22A28_polytom_Cat1	1,0548842	0,772
IK22A28_polytom_Cat2	1,05994208	0,767
IK22A28_polytom_Cat3	1,05444962	0,785
DR22A29_Cat1	0,98779057	0,649
IK22A30_Cat1	1,0428588	0,044
PT23A37_Cat1	1,04551071	0,27
DR23A38_polytom_Cat1	0,99431107	0,911
DR23A38_polytom_Cat2	1,13474556	0,02
DR23A38_polytom_Cat3	1,13448431	0,048
IK23A39_Cat1	0,89715511	0,432
IK23A39_Cat2	1,00695896	0,736
IK23A39_Cat3	1,00896236	0,734
PT23A40_polytom_Cat1	1,0149421	0,704
PT23A40_polytom_Cat2	1,00807055	0,878
DR23A41_Cat1	0,99659487	0,96
DR23A42_Cat1	1,02196394	0,857
DR23A76_Cat1	1,00481997	0,865
IK24A43_Cat1	0,93390589	0,003
IK24A44_polytom_Cat1	0,91588334	0,474
IK24A44_polytom_Cat2	0,97966188	0,559
DR24A45_polytom_Cat1	1,0846685	0,222
DR24A45_polytom_Cat2	1,07525214	0,31
PT24A47_Cat1	1,00252817	0,954
DR24A46_Cat1	1,0049965	0,871
DR24A77_Cat1	0,99649699	0,918
DR24A78_Cat1	0,96007125	0,46
DR11A01_Cat1	0,97061586	0,586
DR11A02_Cat1	0,97267469	0,452
DR11A03_Cat1	0,9576812	0,427

PT11A04_polytom_Cat1	0,96843848	0,809
PT11A04_polytom_Cat2	0,98643485	0,943
PT11A05_Cat1	0,99185245	0,836
DR11A73_Cat1	0,97563319	0,497
DR11A74_Cat1	0,99269989	0,733
IK12A06_Cat1	0,97542645	0,268
DR12A07_Cat1	1,03323128	0,293
PT12A08_Cat1	1,02744393	0,486
IK12A09_Cat1	0,99572527	0,954
DR12A10_Cat1	1,01241157	0,76
PT12A11_Cat1	1,05111726	0,178
PT12A12_Cat1	1,00608677	0,845
PT12A71_Cat1	0,96881185	0,231
DR13A14_Cat1	0,98509446	0,571
PT13A15_Cat1	0,96858478	0,458
IK13A16_Cat1	0,98690212	0,537
IK13A17_polytom_Cat1	1,03619706	0,472
IK13A17_polytom_Cat2	1,08335269	0,006
IK14A18_Cat1	1,04100011	0,098
PT14A19_Cat1	0,96462662	0,655
IK14A20_Cat1	0,99602704	0,988
IK15A21_Cat1	1,03239102	0,15
PT15A22_Cat1	0,96899311	0,516
IK15A23_Cat1	1,0151271	0,658
DR15A24_Cat1	1,01018564	0,66
IK31A53_Cat1	0,95167836	0,713
IK31A54_Cat1	0,98132982	0,595
DR31A55_Cat1	0,94177408	0,007
DR31A52_Cat1	0,99264496	0,92
PT31A49_Cat1	0,96608316	0,386
PT31A50_Cat1	0,96476392	0,669
DR31A51_Cat1	0,94912362	0,016
DR31A56_Cat1	0,98523956	0,757
DR31A70_polytom_Cat1	1,03777109	0,309
DR31A70_polytom_Cat2	1,05984191	0,561
PT31A57_Cat1	1,02317894	0,549
DR31A79_Cat1	0,98080452	0,42
DR31A80_Cat1	0,9846106	0,482
DR31A81_Cat1	0,940462	0,546
IK32A59_Cat1	0,98591796	0,512
IK32A60_Cat1	0,9570855	0,067
IK32A61_Cat1	1,01972245	0,77
DR32A82_polytom_Cat1	1,04133008	0,526
DR32A82_polytom_Cat2	1,03571583	0,77
DR32A82_polytom_Cat3	1,02029401	0,869

PT33A62_polytom_Cat1	1,04867127	0,367
PT33A62_polytom_Cat2	1,03279688	0,573
IK33A63_polytom_ordinal_Cat1	1,24056313	0
IK33A63_polytom_ordinal_Cat2	1,32328425	0
IK33A63_polytom_ordinal_Cat3	1,2647116	0
DR33A65_Cat1	0,93989148	0,367
DR33A66_Cat1	1,05486126	0,197
DR33A67_Cat1	1,05521307	0,156
IK33A68_Cat1	1,00087841	0,967
PT33A69_polytom_Cat1	1,04747603	0,044
PT33A69_polytom_Cat2	1,01292669	0,897

outfit_code	outfit_wert	outfit_p
DR21A31_Cat1	0,99530183	0,916
IK21A32_Cat1	1,08515042	0
DR21A34_polytom_Cat1	1,01696471	0,591
DR21A34_polytom_Cat2	1,06923363	0,105
PT21A35neu_polytom_Cat1	0,97981073	0,828
PT21A35neu_polytom_Cat2	1,00302127	0,824
IK21A72_Cat1	0,96741152	0,117
DR21A75_polytom_Cat1	1,22040548	0,057
DR21A75_polytom_Cat2	1,39676167	0,003
DR21A75_polytom_Cat3	1,36565815	0,014
PT22A25_Cat1	1,03011662	0,508
PT22A26_Cat1	1,00516687	0,933
IK22A27_Cat1	0,63681028	0,079
IK22A28_polytom_Cat1	1,4811114	0,097
IK22A28_polytom_Cat2	1,2135252	0,464
IK22A28_polytom_Cat3	1,11130299	0,684
DR22A29_Cat1	0,97990767	0,449
IK22A30_Cat1	1,03766909	0,077
PT23A37_Cat1	1,06287174	0,131
DR23A38_polytom_Cat1	1,04726627	0,314
DR23A38_polytom_Cat2	1,18275726	0,002
DR23A38_polytom_Cat3	1,26734785	0
IK23A39_Cat1	0,80780817	0,119
IK23A39_Cat2	1,46517225	0,477
IK23A39_Cat3	1,4691083	0,475
PT23A40_polytom_Cat1	1,00022699	0,991
PT23A40_polytom_Cat2	0,98984225	0,864
DR23A41_Cat1	0,99766852	0,975
DR23A42_Cat1	1,1861032	0,307
DR23A76_Cat1	1,00421474	0,881
IK24A43_Cat1	0,92119864	0
IK24A44_polytom_Cat1	0,88542116	0,314

IK24A44_polytom_Cat2	0,96698913	0,338
DR24A45_polytom_Cat1	1,14007656	0,049
DR24A45_polytom_Cat2	1,15167416	0,049
PT24A47_Cat1	1,06905968	0,468
DR24A46_Cat1	1,00191263	0,947
DR24A77_Cat1	0,98942928	0,743
DR24A78_Cat1	0,94391706	0,294
DR11A01_Cat1	0,95580527	0,406
DR11A02_Cat1	0,96126657	0,282
DR11A03_Cat1	0,94935944	0,339
PT11A04_polytom_Cat1	0,98482287	0,91
PT11A04_polytom_Cat2	1,08771034	0,484
PT11A05_Cat1	0,98069401	0,611
DR11A73_Cat1	0,96385239	0,309
DR11A74_Cat1	0,98670124	0,53
IK12A06_Cat1	0,96367125	0,1
DR12A07_Cat1	1,03304825	0,296
PT12A08_Cat1	1,03508859	0,376
IK12A09_Cat1	0,99833423	0,99
DR12A10_Cat1	1,0157645	0,701
PT12A11_Cat1	1,06328471	0,097
PT12A12_Cat1	0,9983761	0,967
PT12A71_Cat1	0,96474285	0,175
DR13A14_Cat1	0,97360369	0,312
PT13A15_Cat1	0,96126999	0,357
IK13A16_Cat1	0,97945325	0,33
IK13A17_polytom_Cat1	1,07674866	0,137
IK13A17_polytom_Cat2	1,07690495	0,01
IK14A18_Cat1	1,03633666	0,142
PT14A19_Cat1	0,96971639	0,704
IK14A20_Cat1	0,98267311	0,931
IK15A21_Cat1	1,02477475	0,27
PT15A22_Cat1	0,94023598	0,202
IK15A23_Cat1	1,01053992	0,756
DR15A24_Cat1	1,00379686	0,867
IK31A53_Cat1	0,93748301	0,622
IK31A54_Cat1	0,95940508	0,24
DR31A55_Cat1	0,93019345	0,001
DR31A52_Cat1	1,00345601	0,941
PT31A49_Cat1	0,9517451	0,213
PT31A50_Cat1	0,9553406	0,582
DR31A51_Cat1	0,94520178	0,01
DR31A56_Cat1	0,97723409	0,626
DR31A70_polytom_Cat1	1,04365491	0,242
DR31A70_polytom_Cat2	1,43606677	0

PT31A57_Cat1	1,02948015	0,449
DR31A79_Cat1	0,97691769	0,33
DR31A80_Cat1	0,98336502	0,446
DR31A81_Cat1	0,89400557	0,263
IK32A59_Cat1	0,97843446	0,313
IK32A60_Cat1	0,94515879	0,019
IK32A61_Cat1	1,02454185	0,721
DR32A82_polytom_Cat1	1,17713972	0,011
DR32A82_polytom_Cat2	5,29154116	0
DR32A82_polytom_Cat3	1,09238142	0,789
PT33A62_polytom_Cat1	1,12889619	0,021
PT33A62_polytom_Cat2	1,05763962	0,333
IK33A63_polytom_ordinal_Cat1	1,32058486	0
IK33A63_polytom_ordinal_Cat2	1,67224685	0
IK33A63_polytom_ordinal_Cat3	1,85169361	0
DR33A65_Cat1	0,93663103	0,34
DR33A66_Cat1	1,07863282	0,067
DR33A67_Cat1	1,05975369	0,126
IK33A68_Cat1	0,9965782	0,893
PT33A69_polytom_Cat1	1,04530584	0,054
PT33A69_polytom_Cat2	1,06482817	0,707

Erläuterung: Die Ausgabe der Infit- und Outfitwerte einschließlich deren jeweiligen p-Werten in R / „TAM“ wurde vom 29.04.2020 um 11:07h verwendet. Die rot markierten Antwortkategorien zeigen die p-Werte unter 0,05 auf. Dies bedeutet, dass die roten Ergebnisse unter Annahme des Rasch-Modells unwahrscheinlich sind. Davon sind im Infit 12 Antwortkategorien und im Outfit 18 Antwortkategorien betroffen.

Hinweis: Da die Berechnung der „Itemstatistik“ in dem Paket „TAM“ auf Monte Carlo Simulationen beruhen, sind bei jeder neuen Berechnung ganz leicht andere Fit-Werte zu beobachten. Diese Veränderungen beziehen sich allerdings erst im Bereich der dritten Nachkommastelle sichtbar. Für die Interpretation der Ergebnisse sind diese Abweichungen praktisch bedeutungslos und wie bei jedem simulationsbasierten Berechnungsverfahren zu erwarten.

Anhang 25: Entwickeltes R-Skript für die CFA-Berechnungen

R-Skript für CFA-Berechnungen_Kompetenzmessung_Dissertation_Ries mit Anmerkungen

---

```
# Einlesen der Daten

library(readr)

# Eigenen Pfad einfügen

data <- read_csv("C:/Users/s.ries/Desktop/R_CFA_Ries/Ries_data.csv")

getwd()

colnames(data)

# Datenaufbereitung

dr1 = rowMeans(data[,c("DR11A01", "DR11A02","DR11A03","DR11A73","DR11A74",

                  "DR13A14","DR15A24")])


dr2 = rowMeans(data[, c("DR21A31","DR21A34_polytom","DR21A75_polytom",

                  "DR22A29","DR23A38_polytom","DR23A42",

                  "DR23A76","DR24A45_polytom","DR24A46",

                  "DR24A77","DR24A78")])


dr3 = rowMeans(data[, c("DR31A55","DR31A52","DR31A51","DR31A56",

                  "DR31A70_polytom","DR31A79","DR31A80",

                  "DR31A81", "DR32A82_polytom","DR33A65")])


pt1 = rowMeans(data[, c("PT11A04_polytom","PT11A05","PT12A08","PT12A12",

                  "PT12A71","PT13A15","PT14A19","PT15A22")])
```

```
pt2 = rowMeans(data[, c("PT22A26","PT23A37","PT23A40_polytom")])

pt3 = rowMeans(data[, c("PT31A49","PT31A50","PT31A57",
                "PT33A62_polytom", "PT33A69_polytom")])

ik1 = rowMeans(data[, c("IK12A06","IK12A09","IK13A16","IK13A17_polytom",
                "IK15A23")])

ik2 = rowMeans(data[, c("IK21A72","IK22A27","IK22A28_polytom","IK22A30",
                "IK23A39","IK24A43","IK24A44_polytom")])

ik3 = rowMeans(data[, c("IK31A53","IK31A54","IK32A59","IK32A60",
                "IK33A63_polytom_ordinal","IK33A68")])
#
SR <- data.frame(dr1,pt1,ik1,dr2,pt2,ik2,dr3,pt3,ik3)
#######################################################################
# Model 1: Eindimensional Pflege = Pflege_1D
library(lavaan)
library(semPlot)
model_pfl_1d <- '
Pflege =~ dr1+pt1+ik1+dr2+pt2+ik2+dr3+pt3+ik3
'
fit_mod1 <- cfa(model_pfl_1d, data=SR)
summary(fit_mod1, standardized=TRUE, fit.measures=T)
semPaths(fit_mod1, what="std", posCol = "black")
#####################################################################
```

```
# Model 2: “Subdomain Model”
model_sd_3d <- '
DR =~ dr1+dr2+dr3
PT =~ pt1+pt2+pt3
IK =~ ik1+ik2+ik3
'
fit_mod2 <- cfa(model_sd_3d, data=SR)
summary(fit_mod2, standardized=TRUE, fit.measures=T)
semPaths(fit_mod2, what="std", posCol = "black")

#######################################################################
# Model 3: “Testlet Model”
model_tl_3d <- '
S1 =~ dr1+pt1+ik1
S2 =~ dr2+pt2+ik2
S3 =~ dr3+pt3+ik3
'
fit_mod3 <- cfa(model_tl_3d, data=SR)
summary(fit_mod3, standardized=TRUE, fit.measures=T)
semPaths(fit_mod3, what="std", posCol = "black")

#####################################################################
# Vergleich der Modelle
# Modell 1 vs. Modell 2
anova(fit_mod1, fit_mod2)
# Modell 1 vs. Modell 3
```

```
anova(fit_mod1, fit_mod3)

# Modell 2 vs. Modell 3 kann nicht verglichen werden, da df gleich

anova(fit_mod2, fit_mod3)

round(cbind(m1=inspect(fit_mod1, 'fit.measures'),

    m2=inspect(fit_mod2, 'fit.measures'),

    m3=inspect(fit_mod3, 'fit.measures')),3)

lavTestLRT(fit_mod1, fit_mod2, fit_mod3)

##########################################################################
```

Eidesstattliche Erklärung zur Einhaltung guter wissenschaftlicher Praxis

(Formular zur Verwendung im Anhang schriftlicher Qualifikationsarbeiten an der PTHV - Stand Nov. 2019)

1. Hiermit versichere ich,

   .....

   (VORNAME, NACHNAME, MATRIKELNUMMER)

2. dass ich mich bei der Erstellung der schriftlichen Arbeit an die aktuell geltenden „Richtlinie zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten“ der PTHV genannten Regeln und Verpflichtungen gehalten habe.

3. Mir ist bekannt, dass die Übernahme von Texten, Ideen oder Daten anderer ohne eine eindeutige Kenntlichmachung des Urhebers ein Plagiat ist und gegen die Regeln der guten wissenschaftlichen Praxis verstößt.
   Ich erkläre hiermit, dass ich die vorliegende schriftliche Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel einschließlich der Internetquellen und anderer elektronischer Quellen benutzt habe. Alle Zitate und Halbzitate aus gedruckten, ungedruckten und elektronischen und sonstigen Quellen, aus jeder Art von Publikationen oder sonstiger Verbreitung von Arbeitsergebnissen, aus bildlichen Darstellungen und dergleichen habe ich als solche eindeutig und im Einzelnen kenntlich gemacht.

4. Die Arbeit wurde vorher nicht in einem anderen Prüfungsverfahren eingereicht.

5. Die eingereichte elektronische Fassung (z.B. auf CD-ROM) ist im Wortlaut identisch mit der eingereichten schriftlichen Fassung.

6. Ich bin mir bewusst, dass die Aufdeckung von nicht gekennzeichneten Passagen, die nicht von mir stammen, als Plagiatsversuch interpretiert wird.
   Dies kann eine Bewertung der Arbeit mit „nicht ausreichend“ und gegebenenfalls rechtliche Konsequenzen zur Folge haben. Sollte im Nachhinein ein Fehlverhalten gemäß der PTHV-Richtlinie „Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten“ nachgewiesen werden, kann der mit der Arbeit erworbene akademische Grad oder Titel entzogen werden.

_____

Ort, Datum, Unterschrift