universität wien

Diplomarbeit

Titel der Arbeit

Analyse der Nutzung der Selbsttests im Projekt “Vienna E-Lecturing“ (VEL) und Evaluation des web-basierten attributionalen Feedbacks

Verfasserin

Michaela Stefanie Fasching

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, im November 2008

Studienkennzahl:	298
Studienrichtung:	Psychologie
Betreuer:	Dipl.-Psych. Dr. Ralph Reimann

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die mich durch mein Studium begleitet haben und mich in der Zeit meiner Studiumsabschlussphase unterstützt haben.

Ganz besonders danken möchte ich …

… meinem Diplomarbeitsbetreuer *Ralph Reimann*,
für seine tatkräftige Unterstützung bei meiner Diplomarbeit
und für die aufgewendete Zeit für unsere Besprechungen.

… meiner Betreuerin als Studienassistentin *Barbara Schober*,
für das Motivieren zum Fertigschreiben meiner Diplomarbeit,
für ihr Verständnis und die Möglichkeit der flexiblen
Arbeitseinteilung als Studienassistentin.

… meinen Kolleg/inn/en *Eva-Maria Schiller* und *Takuya Yanagida*,
für die konstruktiven Kaffeepausen und für das fachliche
und inhaltliche Feedback zu meiner Diplomarbeit.

… meiner Freundin *Susi* für das Korrektur lesen meiner Diplomarbeit
und für das gemeinsame Innehalten in den stressigen Zeiten des Lebens.

… meinen Freundinnen *Nina, Alex, Dilber, Kirsti, Katharina, Anne-Marie, iSa und Grete* für die soziale und emotionale Unterstützung während schwieriger Phasen
und die gemeinsame Zeit und dem Spaß in sonnigen Phasen.

… meiner Schwester *Martina*
dafür, dass ich meine Freuden und Leiden mit ihr teilen kann und
für die Wochenenden in Pöchlarn, an denen ich wieder Energie tanken konnte.

… meinem besten Freund auf der ganzen Welt, *Per*,
dafür, dass er mir immer wieder Kraft und Zuversicht gegeben hat
und für seine liebevolle Unterstützung und seine Geduld.

Inhaltsverzeichnis

Einleitung 1

1 Theoretischer Teil 3

1.1 Begriffsdefinitionen von E-Learning bis Blended Learning 3

1.2 Vienna E-Lecturing (VEL) 6

1.2.1 Lehrziele von VEL 6
1.2.2 Technische Umsetzung von VEL 7
1.2.3 Didaktisches Konzept von VEL 8
1.2.4 Aufbau von VEL 9

1.3 Theoretische Grundlagen 15

1.3.1 Selbsttests 15
1.3.2 Feedback 17
1.3.3 Attributionstrainings 22
1.3.4 Selbsttests und attributionales Feedback in VEL 27

1.4 Evaluation von E-Learning 31

1.4.1 Evaluation im VEL 32
1.4.2 Evaluationsergebnisse zu VEL 33

1.5 Ziele und Fragestellungen 35

2 Empirischer Teil 39

2.1 Untersuchungsplan 39

2.1.1 Untersuchungsdesign 39
2.1.2 Stichproben 40
2.1.3 Erhebungsinstrumente 41

2.2 Untersuchungsdurchführung 45

2.2.1 Aufbereitung und Komprimierung der Daten 46
2.2.2 Statistische Auswertungsverfahren 57

2.3 Ergebnisse 58

2.3.1 Stichprobenbeschreibung 58
2.3.2 Deskriptive Ergebnisse zu Ziel 1 60
2.3.3 Prüfung der ST-Durchführungshäufigkeits-Gleichheit 62
2.3.4 Beantwortung der Fragestellungen zu Ziel 1 65
2.3.5 Zusätzliche Ergebnisse zu Ziel 1 82
2.3.6 Deskriptive Ergebnisse zu Ziel 2 88
2.3.7 Beantwortung der Fragestellungen zu Ziel 2 93
2.3.8 Zusätzliche Ergebnisse zu Ziel 2 98

3 Zusammenfassung und Diskussion 101

4 Literaturverzeichnis 109

5 Anhang 117

5.1 Anhang – Terminübersicht für WS 2003/04 und SS 2004 117
5.2 Anhang – Attributionale Rückmeldungen der zehn Selbsttests 118
5.3 Anhang – SRL-Fragebogen des MZP 2 im Studienjahr 2003/04 128
5.4 Anhang – Reliabilität der Skalen des SRL-Fragebogens 134
5.5 Anhang – Kategorieschema der Attributionen 135
5.6 Anhang – Übersicht der Freischaltung 2004/05 137
5.7 Anhang – Korrelationsmatrizen ad Fragestellung 1h 137
5.8 Anhang – Häufigkeiten der Attributionsnennungen 138

Abstract (deutsch) 141
Abstract (englisch) 142

Einleitung

In dieser Diplomarbeit geht es um das Programm “Vienna E-Lecturing“ (VEL), genauer gesagt um ein spezifisches Element aus VEL, nämlich um die Nutzung und den Effekt von Online-Selbsttests mit web-basierten attributionalen Feedbacks.

Was ist VEL? – “Vienna E-Lecturing“ ist ein internetgestütztes Lehr-Lern-Konzept, welches an der Fakultät für Psychologie der Universität Wien – im Arbeitsbereich Bildungspsychologie und Evaluation – entwickelt und im Rahmen der Pflicht-lehrveranstaltung “Forschungsmethoden und Evaluation“ (FM&E) im Diplomstudium Psychologie umgesetzt wurde (vgl. Wagner, Schober, Reimann, Atria & Spiel, 2007). In diesem Konzept leiten zeitlich aufeinanderfolgende und thematisch sukzessiv aufeinander aufbauende Online-Lernmodule (insgesamt 10) die Studierenden systematisch zur inhaltlichen Auseinandersetzung mit dem Lehrstoff an. Diese Auseinandersetzung umfasst jeweils mehrere Arbeitsschritte (siehe Abbildung 1) und endet jeweils mit der termingerechten Abgabe eines in Kleingruppen erarbeiteten "Produkts" (vgl. Schober, Reimann, Wagner, Atria & Spiel, 2006).

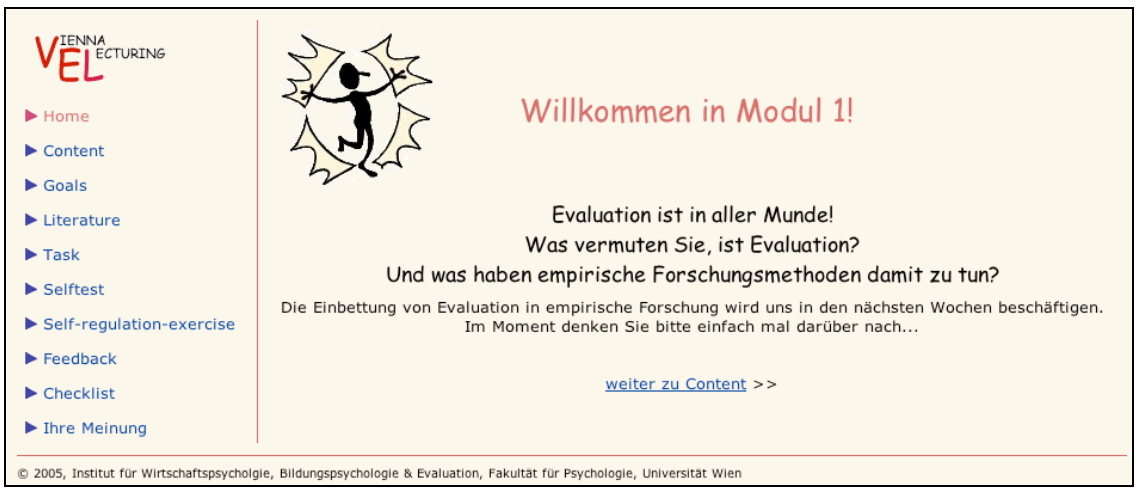


Abbildung 1: Die Einstiegsseite (“Home“) von Modul 1. Die Navigationsleiste auf der linken Seite zeigt die verschiedenen Arbeitsschritte (bzw. Komponenten) des Moduls. Die Struktur dieser Komponenten ist über die Module hinweg gleichbleibend, die Inhalte und konkreten Arbeitsanforderungen modul-spezifisch (vgl. Wagner et al., 2007).

Ein Bestandteil eines jeden Lernmoduls von VEL ist der so genannte “Selftest“, der online durchgeführt wird und ein web-basiertes ergebnisabhängiges Feedback inkludiert. Abbildung 2 zeigt die Einstiegsseite zum Online-Selftest (im Folgenden auch Selbsttest genannt) des ersten Lernmoduls.

Abbildung 2: Screenshot der Einstiegsseite zum Selbsttest 1

Mit diesen Online-Selbsttests, deren Nutzung und Effekten, befasst sich die vorliegende Forschungsarbeit, die im Wesentlichen zwei Ziele verfolgte. Das erste Ziel war die Analyse der Nutzung der Online-Selbsttests im Rahmen der Lehrveranstaltung “Forschungsmethoden und Evaluation“ und die Überprüfung des Zusammenhangs der Selbsttestnutzung mit den erreichten Leistungen in den Lehrveranstaltungsprüfungen. Das zweite Hauptziel war die summative Evaluation des web-basierten attributionalen Feedbacks der Selbsttests.

VEL steht im Kontext von E-Learning, Blended Learning und Selbstreguliertem Lernen (SRL). Diese Aspekte werden im Folgenden zu Beginn des *theoretischen Teils* definiert und von verwandten Begriffen abgegrenzt. Anschließend wird das zu Grunde liegende Konzept von VEL erläutert und detailliert auf die theoretischen Grundlagen der Selbsttests und des attributionalen Feedbacks eingegangen. Im Anschluss daran werden die Selbsttests und das attributionale Feedback des Trainingsprogramms VEL vorgestellt. Ferner wird eine kurze Einführung zum Thema Evaluation bei E-Learning gegeben und es werden auszugsweise bisherige Ergebnisse der Evaluation des Programms VEL berichtet. Den Abschluss des theoretischen Teils bildet die Ableitung der Fragestellungen für beide Ziele der vorliegenden Arbeit. Im *empirischen Teil* werden zu Beginn der Untersuchungsplan und die Durchführung der Untersuchung, sowie die Aufbereitung und Komprimierung der Daten beschrieben. Im Ergebnissteil werden nach der Stichprobenbeschreibung, getrennt nach den beiden Zielen, die deskriptiven Ergebnissen dargestellt, die Fragestellungen beantwortet und zusätzliche Ergebnisse wiedergegeben. Den Abschluss bilden die *Zusammenfassung* der Arbeit und die *Diskussion* der Ergebnisse.

1 Theoretischer Teil

1.1 Begriffsdefinitionen von E-Learning bis Blended Learning

Unter *E-Learning* wird im Allgemeinen das Lernen mit Hilfe elektronischer Medien verstanden. Die Methoden der Umsetzung sind allerdings sehr vielfältig und reichen von Computer-Based-Training (CBT) bzw. Web-Based-Training (WBT) bis hin zu Online-Lernen (Mandl, 2004). Während der Begriff *Computer-Based-Training (CBT)* – *oder auch* Computerunterstütztes Lernen (CUL*)* genannt – Lernprogramme kennzeichnet, die auf Basis von Computern zum Lernen eingesetzt werden (Bruns & Gajewski, 2002), beschreibt der Begriff *Web-Based-Training (WBT)* das Lernen über Netzumgebungen, wie z.B. Intranet oder das Internet (Ihringer, 2000). *Online-Lernen* wiederum zeichnet sich durch die Besonderheit aus, dass zusätzlich zu WBT (oder CBT) verschiedene Kommunikationsformen gewählt werden können (wie z.B. Email, Newsgroup, Chat etc.). Lerninhalte und Kommunikationsbeteiligte können hierbei miteinander vernetzt werden (Meister, 2004).

In der Literatur werden die Begriffe E-Learning, CBT, WBT, Online-Lernen etc. oftmals synonym verwendet (Reinmann & Mandl, 2006), jedoch hat sich der Begriff E-Learning weitgehend als gebräuchlichster durchgesetzt. Vor allem deshalb, weil er anschlussfähig ist an die ökonomischen Begrifflichkeiten, wo sich die Kategorie der E-Begriffe (englisch “e-terms“) zunehmend etabliert (wie z.B. E-Banking, E-Commerce etc.). Das “e“ steht für “electronic“ und bedeutet “auf Elektronenfluss beruhend“ (Meister, 2004). Der Begriff E-Learning wird aber auch im schulischen und universitären Umfeld verwendet (Schüpbach, Guggenbühl, Krehl, Siegenthaler & Kaufmann-Hayoz, 2003). Wobei die Gruppe der E-Begriffe noch durch den Begriff *E-Lecturing* erweitert wurde, unter welchem Spiel et al. (2004) die Anwendung von E-Learning-Methoden auf Vorlesungen verstehen.

Im Rahmen von E-Learning findet man oft auch so genannte *E-Learning-Plattformen* – auch *Lernplattformen* genannt. Darunter versteht man ein System, das es ermöglicht ein virtuelles Bildungszentrum aufzubauen. Der Kern der Plattform ist die Bereitstellung jeglicher Art von Lernmedien. Darüber hinaus bieten viele Lern-plattformen unzählige weitere Möglichkeiten, wie virtuelle Kommunikationsmedien, Suchfunktionen oder separate Arbeitsbereiche für jede/n Lernenden (Mandl, 2004).

Mit der Einführung von E-Learning bzw. E-Lecturing sind meist hohe Erwartungen verknüpft, denn „E-Learning soll ein zeit- und ortsunabhängiges, individuelles, bedarfsgerechtes sowie effektives und effizientes Lernen ermöglichen, das zugleich das Bildungspersonal entlastet und Kosten einspart, aber dennoch eine höhere Qualität erbringt als traditionell organisiertes Lernen in Präsenzveranstaltungen“ (Zimmer, 2002, S. 5). Nach Mandl (2004) ist der Einsatz von E-Learning aber nur dann sinnvoll, wenn dies vor dem Hintergrund einer neuen Lernkultur, nämlich dem *Lernen im Rahmen einer problemorientierten Lernumgebung*, stattfindet. Diese neue Lernkultur basiert auf zwei Säulen, (1) dem Selbstgesteuerten Lernen und (2) dem Kooperativen Lernen.

(1) Der Begriff *Selbstgesteuertes Lernen* stellt kein fest definiertes Konzept dar. Die Grenzen bzw. Übergänge des Konzeptes des Selbstgesteuerten Lernens zu ähnlichen Begriffen, wie z.B. Selbstreguliertes Lernen, Selbstorganisiertes Lernen, Selbstbestimmtes Lernen etc. sind fließend (vgl. Friedrich & Mandl, 1997; Reinmann & Mandl, 2006; Spiel et al., 2004). Im Folgenden werden die beiden am häufigsten verwendeten Begriffe kurz näher beschrieben.
Nach Weinert (1982) ist *Selbstgesteuertes Lernen* eine Lernform, bei der der/die Handelnde die wesentlichen Entscheidungen, ob, was, wann, wie und woraufhin er/sie lernt, gravierend und folgenreich beeinflussen kann. Zimmerman (2001) versteht unter dem Begriff *Selbstreguliertes Lernen*, wenn ein/e Lernende/r selbst Gedanken, Gefühle und Verhaltensweisen erzeugt um sein/ihr Lernziel zu erreichen. Selbstregulation ist hierbei ein zyklischer Prozess, der aus einer Abfolge der drei Phasen “forethought“ (Planung), “performance/volition control“ (Handeln und volitionale Kontrolle) und “self-reflection“ (Selbstreflektion) besteht (Zimmerman, 2000). Die erste Phase beinhaltet die Planung einer Lernhandlung, in der zweiten Phase erfolgt dann die konkrete Umsetzung und in der dritten Phase wird die durchgeführte Lernhandlung bewertet (Wagner et al., 2007). Im Unterschied zum Selbstbestimmten Lernen, wo die Festlegung der Inhalte und Ziele durch den/die Lernende/n selbst erfolgt, werden beim Selbstregulierten Lernen auch fremdgesetzte Ziele verfolgt (Friedrich & Mandl, 1997). In der Literatur werden die Begriffe Selbstgesteuertes Lernen und Selbstreguliertes Lernen häufig synonym verwendet (vgl. Reinmann & Mandl, 2006; Schmitz, 2003). In der vorliegenden Arbeit wird

jedoch nur der Begriff Selbstreguliertes Lernen (SRL) im Sinne der Definition und des Konzeptes von Zimmerman (2000, 2001) benutzt.

(2) *Kooperatives Lernen* – die zweite Säule der von Mandl (2004) propagierten neuen Lernkultur – bedeutet, dass Lernende in einer Gruppe zusammenarbeiten, die so klein ist, dass jede/r an einer kollektiven Aufgabe partizipieren kann (Krapp & Weidenmann, 2006). In der Literatur werden die Bezeichnungen “Kooperatives Lernen“ und “Kollaboratives Lernen“ zum Teil synonym verwendet (z.B. Carell, 2006; Slavin, 1993). Zum Teil werden sie auch definitorisch voneinander abgegrenzt (Perrez, Huber & Geißler, 2006). Das wichtigste Merkmal des Kooperativen Lernens ist die Aufteilung der zu erledigenden Arbeit auf die Gruppenmitglieder – die Teilaufgaben werden von jeder Person für sich gelöst – und eine spätere Zusammenführung der individuellen Ergebnisse zu einem gemeinsamen Arbeitsresultat. Der Arbeitsablauf kann dabei durch bestimmte Vorgaben strukturiert und gesteuert werden. Beim *Kollaborativen Lernen* können die Mitglieder einer Arbeitsgruppe ihre Ziele und Vorgehensweisen selbstständig bestimmen, denn hier besteht zumeist nur ein geringes Ausmaß an Strukturierung von außen (vgl. Perrez et al., 2006; Reinmann-Rothmeier & Mandl, 2002). Im deutschsprachigen Raum ist die Unterscheidung zwischen kooperativem und kollaborativem Lernen wenig bekannt und demzufolge leider auch wenig verbreitet (Reinmann-Rothmeier & Mandl, 2002). In der vorliegenden Arbeit jedoch wird dieser Divergenz Rechnung getragen, indem nur die Definition des Kollaborativen Lernens verwendet wird (siehe Kapitel 1.2.1 – Lehrziele von VEL).

Gerade mittels der heutigen Informations- und Kommunikationstechnologien (IKT) besteht ein großes Realisierungspotenzial für selbstregulierte und kollaborative computerbasierte Lernsettings (Mandl, 2004). Leider wurde und wird die Diskussion zu E-Learning durch die Polarisierung zwischen Präsenzlernen auf der einen Seite und E-Learning auf der anderen Seite belastet. Dies ist jedoch den Erfahrungen zum computerunterstützten Lernen zufolge nur selten zielführend. Es sollte deshalb nicht um die Frage der Überlegenheit der einen Lehrform gegenüber der anderen gehen, sondern um deren effiziente Kombination, im Sinne von Blended Learning (Kerres & Jechle, 2002).

Blended Learning heißt direkt übersetzt “vermischtes Lernen“ (Reinmann-Rothmeier, 2003) und bedeutet eben die Kombination von rein virtuellen Phasen – in denen verschiedenen Formen des Online-Lernens zur Anwendung kommen können – und Präsenzphasen – in denen die Teilnehmer/innen die Möglichkeit haben, sich “*face-to-face*“ (d.h. “von Angesicht zu Angesicht“) zu treffen und sich auszutauschen. Die Variationsmöglichkeiten für die Integration von E-Learning und Präsenzphasen sind sehr vielfältig (vgl. Kerres, 2001; Mandl, 2004; Reinmann-Rothmeier, 2003). Im nächsten Kapitel wird nun ein Beispiel eines solchen Blended-Learning-Konzepts, das Programm Vienna E-Lecturing (VEL), vorgestellt.

1.2 Vienna E-Lecturing (VEL)

Das internetgestützte Lehr-Lern-Konzept VEL wurde im Arbeitsbereich Bildungspsychologie und Evaluation an der Fakultät für Psychologie der Universität Wien in mehreren Phasen entwickelt, evaluiert und optimiert (Spiel et al., 2004).

1.2.1 Lehrziele von VEL

Dieses Programm VEL verfolgt *vier Lehrziele*. Diese Lehrziele sind die Vermittlung von Fachwissen, Lernkompetenz, Kompetenz zum kollaborativen Lernen und E-Kompetenz (Schober, Wagner, Reimann, Atria & Spiel, 2006; Wagner et al., 2007). Das *erste Lehrziel*, die Vermittlung von *Fachwissen*, bezieht sich auf die fachspezifischen Inhalte der Pflichtvorlesung “Forschungsmethoden und Evaluation“. Den Studierenden soll mittels dieser Lehrveranstaltung jenes Wissen und Können vermittelt werden, das für das Verstehen, Planen, Durchführen und Bewerten empirischer Studien notwendig ist (Wagner et al., 2007). Das *zweite Lehrziel* ist die systematische Förderung von *Lernkompetenz*. Das VEL-Programm zielt nach Wagner et al. (2007) darauf ab, „die Studierenden sukzessive in die systematische Steuerung ihrer eigenen Lernprozesse einzuführen“ (S. 293) und orientiert sich dabei am Phasenmodell zum Selbstregulierten Lernen von Zimmermann (2000; vgl. dazu Kapitel 1.2.3 – Didaktisches Konzept von VEL). Das *dritte Lehrziel* ist das *Kollaborative Lernen und Arbeiten*. Hierbei orientierte man sich bei der Konzeption des VEL-Programms an der Selbstbestimmungstheorie von Deci und Ryan (1993), wonach ein Individuum vor allem dann intrinsisch motiviert arbeitet und lernt, wenn

es sich als *autonom* handelnd erlebt und sich *kompetent* sowie *sozial eingebunden* fühlt (Schober, Wagner et al., 2006; Wagner et al., 2007). Die Vermittlung von *E-Kompetenz* ist das *vierte Lehrziel* von VEL. Darunter verstehen Wagner et al. (2007) jene Fertigkeiten, die notwendig sind, um mit einer E-Learning-Plattform umgehen und sich an einem elektronisch-gestützten Informationsaustausch beteiligen zu können. E-Kompetenz kann nur vermittelt werden, wenn online bzw. auf einer E-Learning-Plattform gearbeitet wird. Deshalb wird im Folgenden die technische Umsetzung des VEL-Programms auf einer Lernplattform vorgestellt.

1.2.2 Technische Umsetzung von VEL

Die Onlinekomponenten des Programms VEL wurden auf einer passwortgeschützten E-Learning-Plattform präsentiert, wodurch nur angemeldete Studierende Zugang zum VEL-Programm hatten (vgl. Schober, Reimann et al., 2006; Schober & Spiel, 2005). Im Studienjahr 2003/04 wurde das VEL-Konzept technisch auf der Lernplattform *ILIAS* (Integrietes Lern-, Informations- und ArbeitskooperationsSystem) realisiert. Im Wintersemester 2004/05 erfolgte der Umstieg auf die E-Learning-Plattform *Blackboard Vista*[1], welche von der Universität Wien betrieben und technisch unterstützt wurde (vgl. Schober, Reimann et al., 2006; Schober & Spiel, 2005).

Die umfangreichen Möglichkeiten von Lernplattformen (wie z.B. selektive Freigabe von Inhalten, Zuweisung von Aufgaben, Einteilung der Studierenden in Lerngruppen, sowie Diskussionsforen und Chats; vgl. ZID, 2008a) wurden im Rahmen von VEL einerseits für die administrativen Arbeiten innerhalb der Lehrveranstaltung genutzt (Schober, Reimann et al., 2006) und andererseits für die virtuelle Kommunikation und Zusammenarbeit. Die unmittelbare Darbietung der Lerninhalte, inklusive der Selbsttests, erfolgte über HTML[2]- bzw. PHP[3]-Dateien. Durch die zusätzliche

[1] Die Lernplattform heißt aufgrund der Fusion der Herstellerfirma mit einem anderen Anbieter nicht mehr *WebCT Vista*, sondern Blackboard Learning System Vista Enterprise License, kurz *Blackboard Vista* (Zentraler Informatikdienst der Universität Wien [ZID], 2008b).

[2] HTML (*Hyper Text Markup Language*) ist eine Beschreibungssprache mit der Webseiten für das Internet erstellt werden (vgl. Immler, 2006; Winkler, 2006)

[3] PHP (*Personal Home Page*) ist eine in HTML eingebettete Skriptsprache für die Erstellung dynamischer Webseiten (vgl. Vogt, 2004; Winkler, 2006; ATLAS Gesellschaft für Internationale Kommunikation, 2003).

Verwendung von Stylesheets[4] und Javascript[5] ergaben sich umfangreiche Möglichkeiten der Darstellung. Aufgrund der Nutzung von HTML- bzw. PHP-Dateien ist das VEL-Programm *plattformunabhängig*, d.h. nicht auf Blackboard Vista angewiesen (vgl. Schober, Reimann et al., 2006; Spiel et al., 2007). Trotz unzähliger Möglichkeiten in der Gestaltung der Onlinekomponenten wurde darauf geachtet, dass die Realisierung des didaktischen Konzeptes nicht vernachlässigt wurde (Schober, Reimann et al., 2006). Im Folgenden wird dieses didaktische Konzept vorgestellt.

1.2.3 Didaktisches Konzept von VEL

Das didaktische Konzept des Trainingsprogramms VEL basiert auf zwei zentralen Prinzipien – dem Prinzip der Vernetzung und dem Prinzip einer möglichst optimalen Instruktion (vgl. Wagner et al., 2007): (1) Das *Prinzip der Vernetzung* betrifft *erstens* den Einsatz von Blended Learning, wobei die Lehrinhalte, welche in den Online-Modulen vermittelten werden mit jenen der Präsenzeinheiten verknüpft sind. *Zweitens* werden im VEL-Programm die vier Lehrziele (Fachwissen, Lernkompetenz, Kompetenz zum kollaborativen Lernen und E-Kompetenz) nicht hintereinander und unabhängig, sondern gleichzeitig und ineinander greifend vermittelt. *Drittens* werden in den Lernmodulen explizite und implizite Vermittlungsformen angewendet und systematisch kombiniert. So werden z.B. die Studierenden in den Online-Modulen explizit angeleitet, sich mit einer bestimmten Literatur auseinanderzusetzen, während implizites Lernen durch Gruppenarbeiten gefördert werden soll. Und v*iertens* vermittelt VEL simultan drei verschiedene Wissensarten: deklaratives, prozedurales und konditionales Wissen (bzgl. Details siehe Wagner et al., 2007). (2) Das zweite Prinzip, das *Prinzip einer möglichst optimalen Instruktion,* basiert auf den Lehrfunktionen nach Klauer (1985) bzw. den Lehr-Lern-Schritten nach Gagné (1985). Die Kombination dieser beiden Ansätze führte zu sieben Lehrfunktionen (siehe Tabelle 1), die sowohl bei der Instruktion in den Online-Modulen als auch bei der Umsetzung der Präsenzeinheiten angewendet wurden (Wagner et al., 2007). Diese sieben Lehrfunktionen können nach Wagner et al. (2007) auf einer Metaebene auch

[4] Stylesheets sind Textdateien, die eine Erweiterung der Möglichkeiten von HTML darstellen (vgl. ATLAS Gesellschaft für Internationale Kommunikation, 2003; Winkler, 2006)

[5] JavaScript ist eine Skriptsprache deren Befehle in HTML-Dokumente eingebettet werden können und die es ermöglicht interaktive Webseiten zu erstellen (vgl. Vogt, 2004; Winkler, 2006).

den drei Phasen des Prozessmodells des Selbstregulierten Lernens von Zimmerman (2000) zugeordnet werden (siehe Tabelle 1).

Tabelle 1: Lehrfunktionen als Prinzip einer möglichst optimalen Instruktion nach Wagner et al. (2007)

Phasen des Prozessmodells von Zimmerman (2000)	**Lehrfunktionen angelehnt an Klauer (1985) und Gagné (1985)**
Forethought	a) Aufmerksamkeit der Lernenden gewinnen
	b) Lernende über die Lernziele explizit informieren
	c) Vorwissen bei den Lernenden aktivieren
	d) Bedeutsamkeit des Lernstoffes herausstellen
Performance	e) Lerninhalte üben und aktiv umsetzen lassen
Self-reflection	f) Feedback zu Vorgehen und Leistung der Lernenden
	g) Überprüfung und Steuerung des Lernprozesses ermöglichen

Nachdem die didaktischen Prinzipien des Trainingsprogramms VEL vorgestellt wurden, wird im Folgenden die Programmstruktur beschrieben.

1.2.4 Aufbau von VEL

Da es für das bessere Verständnis des folgenden Kapitels wichtig erscheint, wird nun schon vorweg erwähnt, dass im Projekt VEL zum Zwecke der summativen Evaluation zwei Varianten des VEL-Programms implementiert wurden. Einerseits die "*Intensiv-Variante"* (Track A), welche das vollständige VEL-Trainingsprogramm für die Versuchsgruppe beinhaltete, und andererseits die "*Standard-Variante"* (Track B), eine reduzierte Version des VEL-Programms für die Kontrollgruppe. Die Entscheidung der Studierenden, an welcher der beiden Varianten von VEL sie teilnehmen wollten, erfolgte auf freiwilliger Basis (Schober, Reimann et al., 2006; Spiel et al., 2007). Genaue Details zum Evaluationsdesign befinden sich in Kapitel 2.1.1 – Untersuchungsdesign.

Das Programm VEL verfolgte, wie zuvor schon erwähnt, einen Blended-Learning-Ansatz, d.h. insgesamt wurden zehn Online-Module mit 18 (Track A) bzw. sieben (Track B) Präsenzeinheiten, den so genannten "Face-to-face"-Einheiten, systematisch kombiniert (siehe Tabelle 2).

Die *Online-Module* wurden von den Studierenden auf einer passwortgeschützten Lernplattform bearbeitet und behandelten einerseits in sich abgeschlossene, aufeinander aufbauende Inhalte aus dem Vorlesungscurriculum und andererseits Inhalte zum Selbstregulierten Lernen (Wagner et al., 2007). Die Lernmodule wurden sukzessive zu feststehenden Terminen (siehe Tabelle 2) freigegeben, d.h. sie waren ab diesem Zeitpunkt für die Studierenden sichtbar und verfügbar bis mindestens zum letzten Prüfungstermin der zugehörigen Vorlesung, i.e. bis ca. ein Jahr nach Vorlesungsende (vgl. Wagner et al., 2007).

Die *Präsenzeinheiten* (auch *"Face-to-face"-Einheiten* genannt) umfassten für alle Studierenden (Track A und Track B) eine *"Kick-off"-Veranstaltung*, in welcher die Studierenden grundlegende Informationen zur Organisation und dem Ablauf von VEL erhielten, und sechs "Meet-the-Expert"-Einheiten, die in ihrer praktischen Durchführung einer "klassischen" Vorlesung ähnelten. Jedoch wurden in den "Meet-the-Expert"-Einheiten aufbauend auf den Kenntnissen der Studierenden, die sie sich durch die Online-Module und das Literaturstudium erworben hatten, bestimmte Fachinhalte vertiefend diskutiert (Schober & Spiel, 2005). Die Studierenden der "Intensiv-Variante" (Track A) absolvierten zusätzlich zu den soeben genannten Einheiten elf weitere Präsenzeinheiten im Zuge eines Proseminars zum Selbstregulierten Lernen.[6] Tabelle 2 zeigt eine Übersicht der genauen Daten der Online-Module und Präsenzeinheiten exemplarisch für das Studienjahr 2004/05. Die Übersicht für das Studienjahr 2003/04 befindet sich im Anhang 5.1.

[6] Da die Präsenzeinheiten für die vorliegende Arbeit nicht im Detail von Bedeutung sind, wird auf diese hier nicht näher eingegangen (genaue Details dazu siehe Schober, Wagner et al., 2006; Wagner et al., 2007).

Tabelle 2: Übersicht der Termine für die Online-Module und “Face-to-face“-Einheiten des Wintersemester 2004/05 und des Sommersemester 2005

Online-Module Track A + B	**“Face-to-face”-Einheiten**
	14.10.04 Proseminar Track A
	18.10.04 Kick-off Track A + B
	21.10.04 Proseminar Track A
26.10. - 08.11.04 Modul 1	
	28.10.04 Proseminar Track A
	08.11.04 Meet-the-Expert I Track A + B
09.11. - 22.11.04 Modul 2	
	18.11.04 Proseminar Track A
23.11. - 06.12.04 Modul 3	
07.12. - 20.12.04 Modul 4	
	09.12.04 Proseminar Track A
	13.12.04 Meet-the-Expert II Track A + B
11.01. - 24.01.05 Modul 5	
	13.01.05 Proseminar Track A
	17.01.05 Meet-the-Expert III Track A + B
	31.01.05 1. Prüfungstermin WS 2004/05
	07.03.05 2. Prüfungstermin WS 2004/05
	20.06.05 3. Prüfungstermin WS 2004/05
08.03. - 04.04.05 Modul 6	
	10.03.05 Proseminar Track A
	07.04.05 Proseminar Track A
12.04. - 25.04.05 Modul 7	
	18.04.05 Meet-the-Expert IV Track A + B
	21.04.05 Proseminar Track A
26.04. - 09.05.05 Modul 8	
	09.05.05 Meet-the-Expert V Track A + B
10.05. - 30.05.05 Modul 9	
	19.05.05 Proseminar Track A
31.05. - 13.06.05 Modul 10	
	13.06.05 Meet-the-Expert VI Track A + B
	23.06.05 Proseminar Track A
	27.06.05 1. Prüfungstermin SS 2005
	03.10.05 2. Prüfungstermin SS 2005
	23.01.06 3. Prüfungstermin SS 2005

Die zehn *Online-Module* von VEL waren alle einheitlich aufgebaut und beinhalteten für die Studierenden beider Varianten (Track A und Track B) folgende Seiten (vgl. Spiel et al., 2007; Wagner et al., 2007):

- *Home* war die Startseite, auf welche man automatisch beim Einstieg in das Online-Modul gelangte. Hier wurde das Thema des Moduls genannt und eine Anknüpfung an bisheriges Wissen geschaffen (siehe Abbildung 1 in der Einleitung).

- Auf der Seite *Content* wurde das Thema des Moduls genau benannt und das Modul in der Gesamtstruktur der Vorlesung verortet. Ferner wurde auf die praktische Relevanz der Inhalte hingewiesen.

- Die *Goals*-Seite spezifizierte die Lernziele des jeweiligen Moduls.

- Auf der Seite *Literatur* wurden Angaben zur Fachliteratur gegeben, welche für das jeweilige Modul durchzuarbeiten waren.

- Unter *Task* wurden die Aufgaben erläutert, die mittels der angegebenen Literatur, hauptsächlich in Kleingruppen, zu bearbeiten waren. Eine Aufgabe bestand immer aus mehreren Schritten und erforderte die Verschriftlichung der Resultate. Diese waren bis zu einem bestimmten, vorher bekannt gegebenen, Abgabetermin auf die Lernplattform zu stellen.

- Auf der Seite *Selftest* (i.e. Selbsttest) konnten die Studierenden individuell und auf freiwilliger Basis ihr deklaratives Wissen zum Modulinhalt überprüfen. Da der Schwerpunkt dieser Diplomarbeit auf den Selbsttests und dem, den Studierenden des Track A dargebotenen, attributionalen Feedback liegt, werden diese in einem eigenen Kapitel genau erläutert (siehe Kapitel 1.3.4 – Selbsttests und attributionales Feedback in VEL).

Die Online-Module der Intensiv-Variante (Track A) beinhalteten zusätzlich zu den oben genannten HTML- bzw. PHP-Seiten noch folgende (vgl. Spiel et al., 2007; Wagner et al., 2007):

- Auf der Seite *Self-regulation-exercise* wurde eine Gruppenaufgabe gestellt, die dem Lehrziel 2, der Förderung der Kompetenz zum Selbstregulierten Lernen, diente. Auch hier waren die Resultate in schriftlicher Form bis zu einem bestimmten, vorher bekannt gegebenen, Abgabetermin auf der E-Learning-Plattform abzuspeichern.

- Auf der *Feedback*-Seite erhielten die Studierenden Rückmeldung zu ihren Aufgabenbearbeitungen des Task und der Self-regulation-exercise. Hierbei wurde unterschieden zwischen (1) allgemeinem Feedback – schriftlich durch Lehrende – und (2) gruppenspezifischem Feedback – (a) schriftlich durch Lehrende, (b) mündlich durch Lehrende und (c) schriftlich durch "Peers", d.h. durch Studienkolleg/inn/en. Der Einsatz der verschiedenen Feedback-Arten erfolgte systematisch, nämlich über das Jahr hinweg in Richtung zunehmender Selbstregulation (d.h. sukzessive weniger Feedback durch die Lehrenden, dafür immer mehr Peer-Feedback).

- Die Seite *Checklist* beinhaltete eine zusammenfassende Auflistung aller im Modul zu erledigender Arbeiten, einschließlich der zu beachtenden Abgabetermine.

- Die Seite *Ihre Meinung* beinhaltete einen Online-Fragebogen und diente einerseits dem Reflexionsprozess der Studierenden und andererseits der Begleitforschung.

Um den Zusammenhang zwischen dem Aufbau der Online-Module und den didaktischen Prinzipien nochmals zu verdeutlichen, wird in Tabelle 3 die Zuordnung der verwendeten Lehrfunktionen (Gagné, 1985; Klauer, 1985) inkl. der drei Phasen des Prozessmodels des Selbstregulierten Lernens von Zimmerman (2000) zu den einzelnen Modulkomponenten abgebildet (vgl. Wagner et al., 2007).

Tabelle 3: Zuordnung der Lehrfunktionen zu den Modulkomponenten nach Wagner et al. (2007)

Modul-komponenten	**Lehrfunktionen (Gagné, 1985; Klauer, 1985) in den verschiedenen Phasen nach Zimmerman (2000)**		
	1.) Forethought	2.) Performance	3.) Self-reflection
Home	a) Aufmerksamkeit gewinnen		
	c) Vorwissen aktivieren		
Content	d) Bedeutsamkeit herausstellen		
Goals	b) Ziele explizieren		
Literature		e) Inhalte üben und aktiv umsetzen	
Task		e) Inhalte üben und aktiv umsetzen	
Selftest		e) Inhalte üben und aktiv umsetzen	f) Feedback zu Lernfortschritt und Leistung
SRL-Exercise		e) Inhalte üben und aktiv umsetzen	
Feedback			f) Feedback zu Lernfortschritt und Leistung
Checklist			g) Überprüfung und Steuerung des Lernens
Ihre Meinung			g) Überprüfung und Steuerung des Lernens

Nach dieser allgemeinen Darstellung des Trainingsprogramms VEL wird im Folgenden speziell auf die Modulkomponente Selbsttest inkl. des attributionalen Feedbacks eingegangen, wobei vorweg die theoretischen Grundlagen erläutert werden.

1.3 Theoretische Grundlagen

In diesem Kapitel wird zu Beginn ein allgemeiner Überblick zum Thema Selbsttests gegeben. Anschließend werden Theorien und Forschungsergebnisse zum Thema Feedback vorgestellt. Weiters wird auf Attributionstrainings und Befunde zur Wirksamkeit dieser Maßnahmen eingegangen. Zum Abschluss dieses Kapitels wird eine detaillierte Darstellung der im VEL-Programm verwendeten Selbsttests inkl. attributionalen Feedbacks gegeben.

1.3.1 Selbsttests

Selbsttests sind eine spezielle Art der Lernkontrolle (siehe unten). Lernkontrollen werden in der Literatur auch Lernerfolgskontrollen, Tests oder Checks genannt (vgl. Baeßler, Lücke, Koring, Kinnebrock & Rössler, 2003; Kerres, 2001; Schüpbach et al., 2003).

Bezüglich der Erstellung bzw. Bewertung von E-Learning-Angeboten wird in der Literatur häufig darauf hingewiesen, dass am Ende von Lerneinheiten oder Online-Modulen eine Lernkontrolle verankert sein sollte, wodurch eine konzentrierte Auseinandersetzung mit dem Lernstoff von den Lernenden verlangt wird (vgl. Baeßler et al., 2003; Bruns & Gajewski, 2002). Es sollte jedoch nicht der Prüfungscharakter, sondern die Übungsmöglichkeit betont werden (Schüpbach et al., 2003). Dies bedeutet, dass die Lernenden über ihre Lernfortschritte informiert bzw. auf eventuelle Defizite aufmerksam gemacht werden sollten (Schüpbach et al., 2003; Zimmer, 2002). Lernkontrollen sollen außerdem verhindern, dass Lernende der Meinung sind, z.B. einen Lernstoff verstanden zu haben, wobei sie ihn aber tatsächlich nur flüchtig gelesen haben (Kerres, 2001). Kerres (2001) fordert, dass Lernkontrollen, um nicht als Fremdsteuerung sondern als Unterstützung aufgefasst zu werden, als Selbsttests konzipiert sein sollten, als Angebot zur Selbstkontrolle, d.h. nicht automatisch an einer bestimmten Stelle und nicht verpflichtend.

Baeßler et al. (2003) berichten, dass in ihrem Projekt CLIC („Computer-based Learning: Introduction to Communications“) Lernerfolgskontrollen mit direktem,

sofortigem Feedback (so genannten Checks) von der Mehrheit der Studierenden (über 90%) als generell gute Einrichtung beurteilt wurden.

Lernkontrollen werden wegen der Einfachheit der Programmierung oftmals im Multiple-Choice-Format erstellt. Hierbei muss der/die Lernende die richtige(n) Antwort(en) unter einer bestimmten Anzahl von vorgegebenen Alternativen bestimmen. Die Schwierigkeit einer Multiple-Choice-Aufgabe lässt sich erhöhen, wenn aus einer bestimmten Anzahl von Alternativen nicht nur eine, sondern mehrere richtige Antworten zu identifizieren sind (Kerres, 2001). Auch wenn der Einsatz von Multiple-Choice-Fragen vielfach als didaktisch minderwertig eingeschätzt wird, betont Kerres (2001), dass gerade anspruchvollere Antwortformate des Multiple-Choice-Typs eine intensive Beschäftigung mit den gestellten Fragen nach sich ziehen können. Jedoch kann mit Hilfe von Lernkontrollen im Multiple-Choice-Format hauptsächlich nur die Aneignung von Faktenwissen, eventuell noch die Anwendung von Fertigkeiten, nicht aber Zusammenhangswissen überprüft werden (vgl. Baeßler et al., 2003; Kerres, 2001). Deshalb schlagen Baeßler et al. (2003) vor, zusätzlich zu den computerbasierten Lernkontrollen im Multiple-Choice-Format (mit automatisch generierten Rückmeldungen), am Ende jeder Lerneinheit eine oder mehrere komplexe Aufgaben zu positionieren, wobei das neu erworbene Wissen selbstständig zur Problemlösung oder auf Fallbeispiele angewendet werden muss. Die Bewertung dieser Aufgaben erfolgt durch Lehrpersonen oder so genannte Teletutor/inn/en (Arnold & Thillosen, 2002; Baeßler et al., 2003).

Im Trainingsprogramm VEL wurden, wie in Kapitel 1.2 bereits dargestellt, einerseits freiwillige Lernkontrollen, die so genannten Selbsttests, sowie auch komplexere (Gruppen-)Aufgaben realisiert. In dieser Arbeit wird jedoch der Forschungsschwerpunkt auf die im Multiple-Choice-Format erstellten Selbsttests gelegt und auf die komplexeren – durch Personen bewerteten – Aufgaben nicht weiter eingegangen.

Obwohl in der Literatur zu E-Learning diverse Empfehlungen zu Lernkontrollen gegeben werden (z.B. Arnold & Thillosen, 2002; Kerres, 2001; Schüpbach et al., 2003), gibt es keine expliziten Untersuchungen oder Ergebnisse bzgl. des Effekts von Lernkontrollen in E-Learning-Angeboten auf die Lernleistung (z.B. gemessen anhand von entsprechenden Lehrveranstaltungsprüfungen oder spezifischen

Wissenstests), wobei doch die Lernleistung bzw. der Lerneffekt einer der zentralen Aspekte in Lehr-Lern-Settings ist. Aber Forschungsarbeiten zu Lernerfolgskontrollen beinhalten stets auch die Analyse von unterschiedlichen Feedback-Methoden, weil – wie z.B. Krause (2007) schreibt – Lernkontrollen ohne anschließendem Feedback aus pädagogischer Perspektive als wenig sinnvoll anzusehen sind (Krause, 2007). Und auch Schüpbach et al. (2003) betonen, dass Lernkontrollen unbedingt ein Feedback beinhalten sollten. Bei automatisierten Multiple-Choice-Tests können quantitative Feedbacks (z.B. 4 von 10 Aufgaben richtig gelöst) auf sehr einfache Weise programmiert werden, weil die dafür notwendigen Daten sofort nach Bearbeitung der Multiple-Choice-Aufgabe/n im Computer bzw. in einer Datenbank vorhanden sind. Hingegen ist die Möglichkeit, computerbasierte, qualitative Feedbacks zu geben, begrenzt und verlangt je nach Differenziertheit einen entsprechenden Entwicklungsaufwand. Gerade aber im Hinblick auf den Einsatz von Lernkontrollen als Selbsttest- und Übungsinstrument sind qualitative Feedbacks von entscheidender Bedeutung (Schüpbach et al., 2003). Die Beliebtheit von automatisch auswertbaren Lernkontrollen liegt unter anderem an der schnellen individuellen Rückmeldung über den eigenen Lernstand (Arnold & Thillosen, 2002). Jedoch fordert Kerres (2001), dass Lernkontrollen einen diagnostischen Wert haben sollen, welcher dem/der Lernenden ersichtlich ist bzw. sichtbar gemacht wird. Dies hat ein reines Mitteilen eines Punktewertes (i.e. Knowledge of Performance – vgl. Kapitel 1.3.2 – Feedback) allerdings nicht. Und auch Arnold und Thillosen (2002) weisen besonders darauf hin, dass als Rückmeldung nicht nur das richtige Ergebnis einer Aufgabe angezeigt werden sollte (i.e. Knowledge of Correct Response – vgl. Kapitel 1.3.2 – Feedback). Dies führt zu den verschiedenen Arten der Rückmeldung und somit zu dem Kapitel Feedback.

1.3.2 Feedback

In diesem Kapitel wird zunächst der Begriff *Feedback* kurz definiert. Anschließend wird ein kurzer Überblick über die Annahmen der verschiedenen lernpsychologischen Sichtweisen in Bezug auf Feedback gegeben. Zum Abschluss dieses Kapitels werden einige Ergebnisse aus der Feedbackforschung wiedergegeben, welche für die vorliegende Arbeit wichtig erscheinen.

Der Begriff *Feedback* kommt ursprünglich aus der Kybernetik (i.e. die Lehre von Reglungsprozessen) und bezeichnet die Rückmeldung bzw. Rückkopplung (oder wörtlich übersetzt “Rückfütterung“) von Informationen (Fengler, 2004). Die Basis für eine Rückmeldung ist ein Soll-Ist-Vergleich, wodurch eine eventuell bestehende Diskrepanz zwischen Soll- und Ist-Zustand festgestellt werden kann (vgl. Carver & Scheier, 2000). Innerhalb von Lehr-Lernkontexten ist der Ist-Zustand beispielsweise eine erbrachte Leistung und der Soll-Zustand ein zu erreichendes Lernziel (Krause, 2007). In der vorliegenden Arbeit werden die Begriffe Feedback und Rückmeldung synonym verwendet.

In der Feedbackforschung dominierte über mehrere Jahrzehnte die *behavioristische Auffassung* vom Lehren und Lernen. Hierbei wurde *Feedback als Verstärker* verstanden. Ausgangspunkt der behavioristischen Herangehensweise war das „Law of effect“ von Thorndike (1913, zitiert nach Bangert-Drowns, C. C. Kulik, J. A. Kulik & Morgan, 1991, S. 213) und die Annahme, dass positives Feedback, welches auf richtiges Verhalten (z.B. richtige Antworten) folgt, automatisch Stimulus-Response-Assoziationen stärkt und somit die Auftrittswahrscheinlichkeit dieser Verhaltensweise (oder Antworten) erhöht, während negatives Feedback oder auch das Ausbleiben eines positiven Feedbacks diese Wahrscheinlichkeit verringert (vgl. Krause, 2007; Musch, 1999).

Im Zuge der kognitiven Wende veränderte sich das Verständnis von Feedback (Narciss, 2006). Zwar blieb die Auffassung bestehen, dass sich die Konsequenzen eines Verhaltens auf künftiges Verhalten auswirken, jedoch wird *Feedback* innerhalb des *kognitivistischen Ansatzes* nicht als Verstärker, sondern *als Information* verstanden. *Informatives Feedback* zeigt, ob eine Diskrepanz beispielsweise zwischen erbrachter Leistung und einem angestrebten Lernziel besteht (vgl. Krause, 2007; Narciss, 2006). Somit hat das informative Feedback aus kognitivistischer Sicht auch eine *korrigierende Funktion*, vor allem beim Auftreten falscher Antworten (vgl. Kulhavy & Stock, 1989; Narciss, 2006).

Es gibt viele verschiedene Arten des *informativen Feedbacks*. Im Folgenden wird eine Übersicht und Ordnung nach Narciss (2006) von weit verbreiteten und in zahlreichen Lernexperimenten verwendeten Feedback-Arten wieder gegeben:

- *Knowledge of performance (KP):* Hierbei erhalten die Lernenden nach der Bearbeitung von Aufgaben ein summatives Feedback über die erbrachte Leistung (z.B. 80% der Aufgaben wurden korrekt gelöst).
- *Knowledge of result/response (KR):* Hierbei wird angegeben, ob die gegenwärtige Antwort richtig oder falsch ist.
- *Knowledge of the correct response (KCR):* Hierbei wird nach Bearbeitung der Aufgaben die korrekte Antwort oder Lösung präsentiert.
- *Answer until correct* oder *multiple try feedback (AUC/MTF):* Hierbei erhalten die Lernenden als Feedback KR und können anschließend die Aufgabe nochmals bearbeiten. Bei AUC geschieht dies, bis die richtige Antwort gefunden wurde. Bei MTF kann die Aufgabe bis zu einer festgelegten Anzahl an Wiederholungen bearbeitet werden, danach wird dann die richtige Antwort oder Lösung präsentiert.
- *Elaborated feedback (EF):* Hierbei wird als Feedback KR oder KCR präsentiert, jedoch mit zusätzlichen Informationen, welche zur Fehlerkorrektur bzw. für die Lösung von künftigen Aufgaben dienen sollen (Narciss, 2006).

Nach Narciss (2006) haben mehrere Autor/inn/en versucht, die zahlreichen Feedback-Arten zu klassifizieren. Diese Klassifikationen decken sich aber nur bei der Einordnung der so genannten “einfachen Feedback-Arten“. Zwar werden die Feedback-Arten *Knowledge of result/response (KR)* und *Knowledge of the correct response (KCR)* von verschiedenen Autor/inn/en unterschiedlich benannt, aber übereinstimmend als “einfache Feedback-Art“ bewertet (vgl. Narciss, 2006). Die Feedback-Arten *Knowledge of performance (KP)*, *Knowledge of result/response (KR)* und *Knowledge of the correct response (KCR)* werden weiters auch noch “ergebnisorientierte Feedback-Arten“ genannt (Narciss, 2006). Bei der Einordnung und dem Verständnis von Feedback-Arten, die über *KP*, *KR* und *KCR* hinausgehen, stimmen die Klassifikationsvorschläge jedoch nicht mehr überein. Diese werden nach Narciss (2006) unter dem Begriff “elaborierte Feedback-Arten“ zusammengefasst. Da beim Programm VEL nach der obigen Klassifikation nur “einfache Feedback-Arten“ zum Einsatz kamen, wird auf die “elaborierten Feedback-Arten“ nicht näher

eingegangen. Details zu den im Trainingsprogramm VEL verwendeten Feedback-Arten befinden sich im Kapitel 1.3.4 – Selbsttests und attributionales Feedback in VEL.

Neuere, *konstruktivistische Ansätze* verstehen *Feedback als Angebot*, welches die Lernenden für die individuelle oder auch kollektive Wissenskonstruktion nutzen können (vgl. Krause, 2007; Musch, 1999). Für die Lernwirksamkeit von Feedback ist in diesem Rahmen besonders wichtig, dass die Lernenden das Feedback aufmerksam rezipieren (vgl. Bangert-Drowns et al., 1991).

Ferner sind Feedback-Informationen abzugrenzen, die nicht extern dargeboten, sondern von den Lernenden im Lernprozess direkt wahrgenommen werden können, also *internes Feedback* oder *Selbstbeurteilung* (vlg. Krause, 2007; Narciss, 2006). In der vorliegenden Arbeit wird jedoch nur Feedback betrachtet, welches Lernenden von außen und systematisch dargeboten wird. Neben den hier genannten Feedback-Arten, gibt es auch noch zahlreiche Varianten von motivierenden Feedbacks (vgl. Narciss, 2006), unter anderem *Reattributions-Feedbacks*. Auf diese wird im nächsten Kapitel (1.3.3 – Attributionstrainings) genauer eingegangen, da das attributionale Feedback Gegenstand der vorliegenden Untersuchung ist.

Zu Feedback in Lehr-Lernsituationen werden nun schon seit fast einem Jahrhundert Untersuchungen durchgeführt, entsprechend vielfältig sind die Untersuchungsansätze. Die Ergebnisse sind sehr inkonsistent und zum Teil sogar widersprüchlich (Narciss, 2006). Dies liegt zum einen an den unterschiedlichen theoretischen Sichtweisen (von behavioristischen Ansätzen bis hin zu Modellen des Selbstregulierten Lernens) vor deren Hintergrund die Untersuchungen durchgeführt wurden, zum anderen, daran dass die verschiedenen theoretischen Ansätze Feedback unterschiedliche Funktionen zugeschrieben haben (z.B. Verstärkerfunktion, Informationsfunktion oder eine motivierende Funktion). Somit wurden auch unterschiedliche Forschungsfragen und –strategien entwickelt, welche wiederum zu unterschiedlichen methodischen Vorgehensweisen führten (vgl. Bangert-Drowns et al., 1991; Kluger & DeNisi, 1996; Narciss, 2006).

Im Folgenden werden exemplarisch einige Ergebnisse zum Thema Feedback wiedergegeben, welche für die vorliegende Arbeit als wichtig erscheinen.

Nach Hattie und Timperley (2007) sind Feedbackmaßnahmen im Durchschnitt beinahe doppelt so wirksam (d = 0.79) wie klassische Maßnahmen der Beschulung, wie z.B. Lehrqualität und –quantität, Klassenumfeld, Hausaufgaben (d = 0.40), jedoch sind die verschiedenen Feedback-Arten unterschiedlich effektiv (siehe Metaanalyse von Hattie, 1999). Bzgl. des Lernerfolgs sind elaborierte Feedback-Arten meist wirksamer als einfache Feedback-Arten, vor allem bei komplexeren Aufgaben (vgl. Moreno, 2004). Jedoch können beim Erwerb von deklarativem Wissen auch einfache Feedback-Arten lernförderlich sein (vgl. Kulhavy, White, Topp, Chan & Adams, 1985). Von elaboriertem Feedback profitieren den Belegen zufolge vor allem Schüler/innen mit gutem Leseverständnis (Gordijn & Nijhof, 2002) bzw. Personen, die alleine lernen (Krause, Stark & Mandl, 2004). Bei einfachen Aufgaben bietet das Feedback *Answer until correct (AUC)* gegenüber *Knowledge of correct answer (KCR)* den Vorteil, dass sich die Lernenden länger mit einer falsch gelösten Aufgabe beschäftigen (Krause, 2007). Das Feedback *Knowledge of Performance (KP)* ist lernförderlicher als kein Feedback (Brown, 1932; Symonds & Chase, 1929). Und sofortiges Feedback ist im computerbasierten Lehr-Lern-Setting lernwirksamer als verzögertes Feedback (J. A. Kulik & C. C. Kulik, 1988). Jedoch ist für eine lernförderliche Wirkung von Feedback eine bewusste und aufmerksame Rezeption des Feedbacks (*"mindful reception")* durch die Lernenden erforderlich *(*Bangert-Drowns et al., 1991). Ferner lenkt computervermitteltes, schriftliches Feedback die Aufmerksamkeit hin zur Aufgabe und hat somit eine höhere Wirksamkeit auf die Lernleistung als verbales Feedback, denn ein Feedback wirkt sich nicht oder sogar negativ auf den Lernerfolg aus, wenn es die Aufmerksamkeit weg von der Ebene der Aufgabe und hin zum Selbst lenkt (Kluger & DeNisi, 1996).

Bei den eben dargestellten Ergebnissen handelt es sich vor allem um Ergebnisse zu den verschiedenen Arten des informativen Feedbacks, wovon jedoch, wie oben schon erwähnt, Formen der motivierenden Feedbacks abzugrenzen sind. Zu den motivierenden Feedbacks zählen unter anderem die Attributionstrainings, auf welche im nächsten Kapitel näher eingegangen wird.

1.3.3 Attributionstrainings

Unter einem *Attributionstraining* (oder auch *Reattributionstraining* genannt) versteht man ein Training, welches das Ziel verfolgt ungünstige Ursachenerklärungen (i.e. Attributionen, genaue Erläuterung dazu – siehe unten) zu verändern und dadurch die Entwicklung einer realistischen Einschätzung der Ursachen von Leistungshandlungen zu fördern. Attributionstrainings können zusätzlich auch das Vertrauen in die eigenen Fähigkeiten und den Selbstwert stärken, sowie erlernte Hilflosigkeit abbauen. Ferner soll die Lernmotivation von Lernenden gesteigert und deren Leistungsergebnisse verbessert werden (vgl. Grünke & Castello, 2004; Ziegler & Finsterwald, 2008; Ziegler & Schober, 1997).

Im Folgenden wird nun zu Beginn der Begriff Attribution definiert. Anschließend werden die theoretischen Prämissen und Grundlagen von Attributionstrainings sowie die Techniken zur Veränderung von Attributionen kurz erläutert. Zum Abschluss des Kapitels werden dann Ergebnisse zu Attributionstrainings berichtet.

Attributionen (oder auch *Kausalattributionen* genannt) sind subjektive Ursachenzuschreibungen, die auf (1) personenunabhängige Effekte in der Umwelt, (2) das Handeln anderer Personen oder (3) die eigenen Tätigkeiten bezogen sein können. Im bildungspsychologischen Kontext befasst sich Attributionsforschung in erster Linie mit Erklärungen der Ursachen von Handlungen bzw. Handlungsergebnissen anderer und/oder der eigenen Person (vgl. Dresel, 2004; Möller, 2006; Ziegler & Finsterwald, 2008; Ziegler & Schober, 2001).

Attributionstrainings basieren nach Ziegler und Schober (1997, 2001) auf *drei grundlegenden theoretischen Prämissen*: (1) einer konstruktivistischen Perspektive von Wirklichkeit, (2) der Handlungsrelevanz von Attributionen sowie (3) der Erlernbarkeit von Attributionen.

(1) Aus konstruktivistischer Perspektive wird die Wirklichkeit als eine kognitiv konstruierte Wirklichkeit aufgefasst. Objektivität, subjektunabhängiges Denken und Verstehen sind unmöglich (Gerstenmaier & Mandl, 1995). Somit stehen Personen erhebliche Interpretationsfreiräume für die Generierung von z.B. Attributionen zur

Verfügung, welche jedoch durch objektive Widerstände der Umwelt und gesellschaftliche Definitionen der Wirklichkeit eingegrenzt sind (vgl. Ziegler & Finsterwald, 2008; Ziegler & Schober, 2001).

(2) Für die Handlungsrelevanz von Attributionen bietet die Forschungsliteratur eine Fülle von empirischen Belegen. Zum Beispiel wurde gezeigt, dass Attributionen auf ungenügende Anstrengung bei Misserfolg zu erhöhter Erfolgserwartung für künftige Ereignisse und zu gesteigertem Durchhaltevermögen führen. Während Attributionen auf mangelnde Fähigkeit bei Misserfolg die Erfolgserwartung und das Durchhaltevermögen senken sowie depressive Symptome hervorrufen (vgl. Weiner, 1994; Ziegler & Schober, 2001).

(3) Auch die Erlernbarkeit von Attributionen ist in der Literatur gut belegt. Die wichtigsten Informationsquellen zur Ausbildung von Attributionen sind einerseits eigene Erfahrungen, welche subjektiv verarbeitet werden, und andererseits signifikante Bezugspersonen, deren Erklärungsmuster im Rahmen sozialer Interaktionsprozesse übernommen werden (Ziegler & Schober, 2001). Dies wurde beispielsweise anhand von Zusammenhängen zwischen Attributionen von Schüler/inne/n und deren Eltern nachgewiesen (vgl. Dresel, Schober & Ziegler, 2007; Ziegler & Schober, 1996, 2001). Als weiterer Beleg für die Trainierbarkeit von Attributionen können auch die Ergebnisse der bisherigen Attributionstrainings gewertet werden (Ziegler & Schober, 1997).

Es gibt verschiedene Techniken zur Veränderung von Attributionen (siehe unten). Fast alle Trainingstechniken orientieren sich an Weiners Vierfeldertafel und/oder an Kelleys Kovariationsmodell. Das Vierfelderschema von Weiner (1994; siehe unten) wird insbesondere für die *direkten Attributionstechniken* herangezogen. Während an das Kovariationsmodell von Kelley (1967) vor allem die *indirekten Attributionstechniken* angelehnt sind (Ziegler & Finsterwald, 2008; Ziegler & Schober, 2001). Bei direkten Attributionstechniken wird die Ursache eines Erfolgs oder Misserfolgs direkt benannt, während diese bei indirekten Attributionstechniken erst erschlossen werden muss, nämlich über Hinweise, die eine bestimmte Ursache nahe legen (Ziegler & Schober, 2001). Im Trainingsprogramm VEL wurde die *direkte* attributionale Rückmeldung angewandt, weshalb im Folgenden nur auf das Vier-Felder-Schema

von Weiner (1994) genauer eingegangen wird, nicht aber auf das Kovariationsmodell von Kelly (1967)[7].

Ausgehend von den Überlegungen von Heider (1944, 1958), der der Frage nachging, wie sich Individuen Handlungen und Handlungsergebnisse erklären, entwickelten im Jahre 1971 Weiner und seine Mitarbeiter/innen ein bis heute viel beachtetes Vierfelderschema, in dem sie die Lokations- und Stabilitätsdimension mit je zwei Ausprägungen kombiniert haben (Dresel, 2004). Die Dimension *Lokation* sagt aus, ob die Ursache einer Handlung oder eines Handlungsergebnisses internal der handelnden Person zugeschrieben oder ob sie eher auf äußere (= externale) Bedingungen zurückgeführt wird. Die Dimension *Stabilität* beinhaltet die zwei Ausprägungen variabel und stabil (Ziegler & Schober, 2001), wobei stabil "immer wieder auftretend" bedeutet (Ziegler & Finsterwald, 2008)[8].

Tabelle 4: Zweidimensionales Klassifikationsschema für die wahrgenommenen Ursachen von Leistungsergebnissen nach Weiner (1994)

	Lokation	
Stabilität	*internal*	*external*
stabil	Fähigkeit	Aufgabenschwierigkeit
variabel	Anstrengung, Stimmung, Müdigkeit, Krankheit	Zufall

Die in Tabelle 4 abgebildete *Vierfeldertafel nach Weiner* (1994) dient als Klassifikationsschema für die wahrgenommenen Ursachen von Leistungsergebnissen (Dresel, 2004). Der große *Vorteil* dieses Vierfelderschemas ist, dass die wichtigsten *Ursachen* für Handlungsergebnisse *direkt benannt* sind (Ziegler & Schober, 2001). Dies hat wesentlich zu dessen Popularität für die Konzeption von Attributionstrainings beigetragen, denn die inhaltliche Benennung der Attributionen

[7] genaue Erläuterungen dazu siehe Stiensmeier-Pelster & Heckhausen (2006) und Ziegler & Schober (2001)

[8] Zwei weitere Dimensionen, mit welchen das Vierfelderschema in Folge erweitert wurde, sind die Kontrollierbarkeit und die Globalität (vgl. Stiensmeier-Pelster & Heckhausen, 2006; Weiner, 1994; Ziegler & Finsterwald, 2008), auf diese wird aber hier nicht näher eingegangen, da sie für die vorliegenden Arbeit nicht weiter von Bedeutung sind.

bietet eine *kognitive Entlastung* im Vergleich zu indirekten Attributionstechniken (vgl. Ziegler & Finsterwald, 2008).

Nach der Darstellung der Klassifikation der Attributionen nach Weiner (1994) werden nun folgend aus der Forschungsliteratur Aussagen zu (1) *günstigen* bzw. (2) *ungünstigen Attributionen* abgeleitet.

(1) Wie durch viele empirische Studien belegt wurde, sollten *Erfolge bevorzugt auf internale (Fähigkeit, Anstrengung) Ursachen* attribuiert werden, sodass sich Personen als Verursacher/in von Handlungsergebnissen wahrnehmen können. Indessen sollten *Misserfolge auf variable Ursachen (Anstrengung, Zufall)* attribuiert werden, da sie ansonsten als unveränderlich erscheinen (vlg. Ziegler & Finsterwald, 2008; Ziegler & Schober, 2001).
Zusätzlich ist es auch wichtig zwischen *motivationsförderlichen und selbstwert-förderlichen Wirkungen von Attributionen* zu unterscheiden. Attributionen auf variable Ursachen haben motivationsförderliche Konsequenzen. Internale Attributionen von Erfolgen und externale Attributionen von Misserfolgen haben selbstwertdienliche Konsequenzen (Ziegler & Schober, 2001). So ist beispielsweise die Rolle von Attributionen auf die Fähigkeit (internal/stabil) nach Erfolgen umstritten, da diese zwar selbstwertförderlich, aber nicht oder nur wenig motivationsförderlich sind (Heller, 2004; Ziegler & Schober, 2001). Dies wurde zum Teil in einer Studie von Dresel (2004) belegt. Es zeigte sich, dass die Darbietung von Feedback mit Fähigkeitsattributionen (internal/stabil) nach Erfolg zu einer kurzfristigen Selbst-wertsteigerung führte, aber sowohl die Anstrengungsbereitschaft als auch die Leistung verringerte. Allerdings zeigte sich auch, dass alleiniges Feedback auf die Anstrengung (internal/variabel) bei Erfolg zum Teil negative Folgen nach sich ziehen kann. Am geeignetsten erwies sich nach Erfolgen ein kombiniertes Anstrengungs- und Fähigkeitsfeedback, in der eben genannten Reihenfolge (vgl. Finsterwald, 2005; genaue Details dazu siehe Dresel, 2004).
Ferner sei noch auf die Wichtigkeit der *Veridikalität von Attributionsrückmeldungen* hingewiesen. Demnach sollten attributionale Feedbacks realistisch sein, da sie die Grundlage für realitätsangemessenes, effektives Handeln darstellen (Dresel, 2004; Ziegler & Schober, 2001).

(2) Als ungünstige Attributionen werden in der Literatur *externale Attributionen (Aufgabenleichtigkeit, Zufall) bei Erfolg* bezeichnet. Sowie *internal/stabile Attributionen (mangelnde Fähigkeit) bei Misserfolg* (vgl. Grünke & Castello, 2008; Heller, 2004; Robertson, 2000). *Externale Attributionen (Aufgabenschwierigkeit, Zufall) bei Misserfolg* sind zwar selbstwertdienlich, aber nur wenig motivationsförderlich und sollten nur zum Selbstwertschutz bei unsicheren Personen eingesetzt werden (Heller, 2004).

Im Folgenden werden nun die *Techniken zur Veränderung des Attributionsstils*, die in Attributionstrainings verwendet werden, dargestellt. Diese können nach Ziegler und Schober (2001) grob in zwei Kategorien unterteilt werden: (1) *Modellierungstechniken*, welche auf der Theorie des Modelllernens von Bandura basieren (z.B. Bandura, 1977, zitiert nach Ziegler & Schober, 1997, S. 33) und (2) *Kommentierungstechniken*, welche charakteristisch über die Kommentierung von Handlungsergebnissen im Sinne erwünschter Attributionen wirken (Ziegler & Schober, 2001). (1) Bei den *Modellierungstechniken* werden von einem Modell (z.B. Trainer/in oder Lehrkraft) entweder (a) die *erwünschten Attributionen demonstriert* (stellvertretend für die Personen, deren Attributionsstil geändert werden soll), oder (b) die *Personen über die erwünschten Attributionen aufgeklärt*. (2) Die drei wichtigsten Formen von *Kommentierungstechniken* bei Attributionstrainings sind nach Ziegler und Schober (2001) (a) das *verbale attributionale Feedback*, (b) das *schriftliche attributionale Feedback* sowie (c) die *operante Methode*. Bei verbalen bzw. schriftlichen Attributionstrainings werden die Handlungen oder Handlungsergebnisse von Personen mit einem mündlichen Kommentar bzw. einer schriftlichen Rückmeldung versehen, welche/r die erwünschten Attributionen beinhaltet oder nahe legen soll. Bei der operante Methode werden die Personen selektiv für erwünschte Attributionen verstärkt (vgl. Ziegler & Schober, 2001).

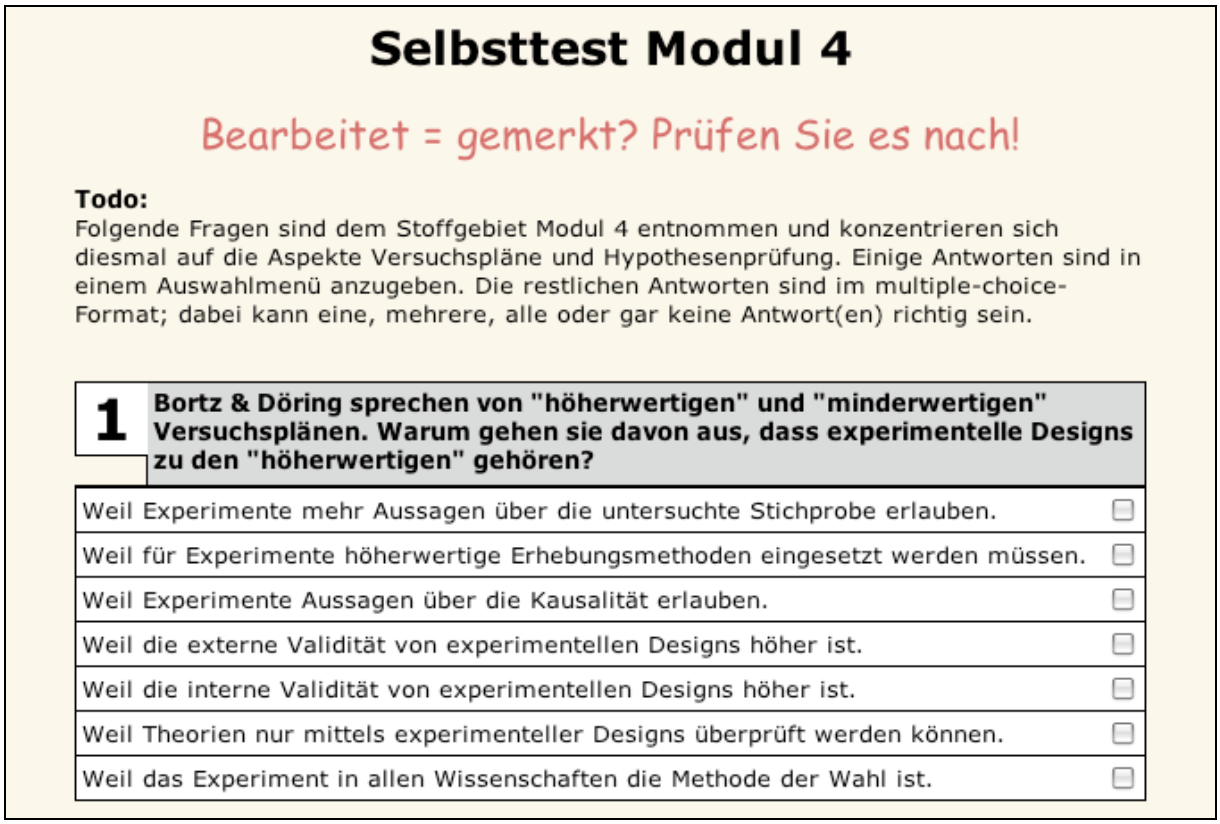
Laut einer Metaanalyse von Finsterwald (1999) sind Attributionstrainings geeignet den Attributionsstil insbesondere hinsichtlich der Dimensionen Fähigkeiten (d = 1.22) und Anstrengung (d = 0.87) substanziell zu verbessern. Die Ergebnisse zeigen weiters eine starke Erhöhung der Selbstwirksamkeit (d = 1.19) sowie einen kleinen Effekt in der Steigerung der schulischen Erfolgserwartung (d = 0.28). Ferner war eine Verbesserung der Schulleistung (d = 0.39) feststellbar, welche auch langfristig

nachweisbar war. Eine differenzielle Wirksamkeit von Attributionstrainings bezogen auf das Geschlecht konnte nicht nachgewiesen werden, was bedeutet, dass sie für beide Geschlechter gleichermaßen effektiv sind (Finsterwald, 2005; Ziegler & Finsterwald, 2008). Attributionale Feedbacks nach Misserfolg auf nicht ausreichende bzw. falsche Lernstrategien (internal/variabel) erwiesen sich als sehr positiv, während Attributionsrückmeldungen auf externale Faktoren (Aufgabenschwierigkeit, Zufall) nach Misserfolgen ungeeignet für die Verbesserung des Attributionsstils waren (Ziegler & Finsterwald, 2008).
Nach einer Metaanalyse von Robertson (2000) sind direkte Rückmeldungen meist wirksamer als indirekte. In kleineren Gruppen sind die Trainingsergebnisse meist besser als in größeren. Und es hat sich als besonders effektiv erwiesen, Attributionstrainings mit Lernstrategien zu kombinieren (vgl. Grünke & Castello, 2004; Robertson, 2000).

Nachdem nun die einzelnen Elemente Selbsttests, Feedback und Attributions-trainings dargestellt wurden, deren Bedeutsamkeit betont und ihre Wirksamkeit belegt wurde, wird nun im Folgenden die Kombination dieser Elemente im Rahmen des Trainingsprogramms VEL beschrieben.

1.3.4 Selbsttests und attributionales Feedback in VEL

Im Trainingsprogramm VEL wurde für jedes der zehn Module eine Lernkontrolle im Multiple-Choice-Format mit jeweils sechs bis elf Aufgaben programmiert, welche die Studierenden nach Freischaltung jederzeit fakultativ zur Kontrolle ihres Lern-fortschrittes nutzen konnten. Es handelte sich hierbei um Lernkontrollen, welche vollständig online abliefen, d.h. die Bearbeitung der Fragen erfolgte ausschließlich am Bildschirm (Schober & Spiel, 2005). Abbildung 3 zeigt einen Screenshot der ersten Frage des Selbsttests aus dem vierten Modul.



Selbsttest Modul 4

Bearbeitet = gemerkt? Prüfen Sie es nach!

Todo:
Folgende Fragen sind dem Stoffgebiet Modul 4 entnommen und konzentrieren sich diesmal auf die Aspekte Versuchspläne und Hypothesenprüfung. Einige Antworten sind in einem Auswahlmenü anzugeben. Die restlichen Antworten sind im multiple-choice-Format; dabei kann eine, mehrere, alle oder gar keine Antwort(en) richtig sein.

1 **Bortz & Döring sprechen von "höherwertigen" und "minderwertigen" Versuchsplänen. Warum gehen sie davon aus, dass experimentelle Designs zu den "höherwertigen" gehören?**

- ☐ Weil Experimente mehr Aussagen über die untersuchte Stichprobe erlauben.
- ☐ Weil für Experimente höherwertige Erhebungsmethoden eingesetzt werden müssen.
- ☐ Weil Experimente Aussagen über die Kausalität erlauben.
- ☐ Weil die externe Validität von experimentellen Designs höher ist.
- ☐ Weil die interne Validität von experimentellen Designs höher ist.
- ☐ Weil Theorien nur mittels experimenteller Designs überprüft werden können.
- ☐ Weil das Experiment in allen Wissenschaften die Methode der Wahl ist.

Abbildung 3: Screenshot des Selbsttests zu Modul 4 vor der Durchführung

Die Studierenden sollten die richtigen Lösungen der Multiple-Choice-Fragen markieren, wobei eine, mehrere, alle oder gar keine Antwort/en richtig sein konnten. Nach der Eingabe des Codes bzw. der Matrikelnummer lieferte ein Mausklick auf den Button „Auswerten“ für alle Studierende (Track A und Track B) ein unmittelbares, automatisch generiertes Feedback. Dieses beinhaltete einerseits die im Selbsttest erreichte Punktezahl (*Knowledge of performance*) und andererseits eine Auflistung, welche Fragen richtig und welche falsch beantwortet wurden (*Knowledge of result*). Die Studierenden konnten den Selbsttest jederzeit abbrechen und so oft sie wollten wiederholen (*Answer until correct*). Im Trainingsprogramm VEL wird das Feedback nach dem Verständnis der kognitiven und der konstruktivistischen Schule als Information und als Angebot an die Studierenden aufgefasst (vgl. Kapitel 1.3.2 – Feedback).

Die Studierenden der “Intensiv-Variante“ (Track A) erhielten zusätzlich zur oben beschriebenen Rückmeldung ein ergebnisabhängiges individuelles attributionales Feedback (siehe Abbildung 4).

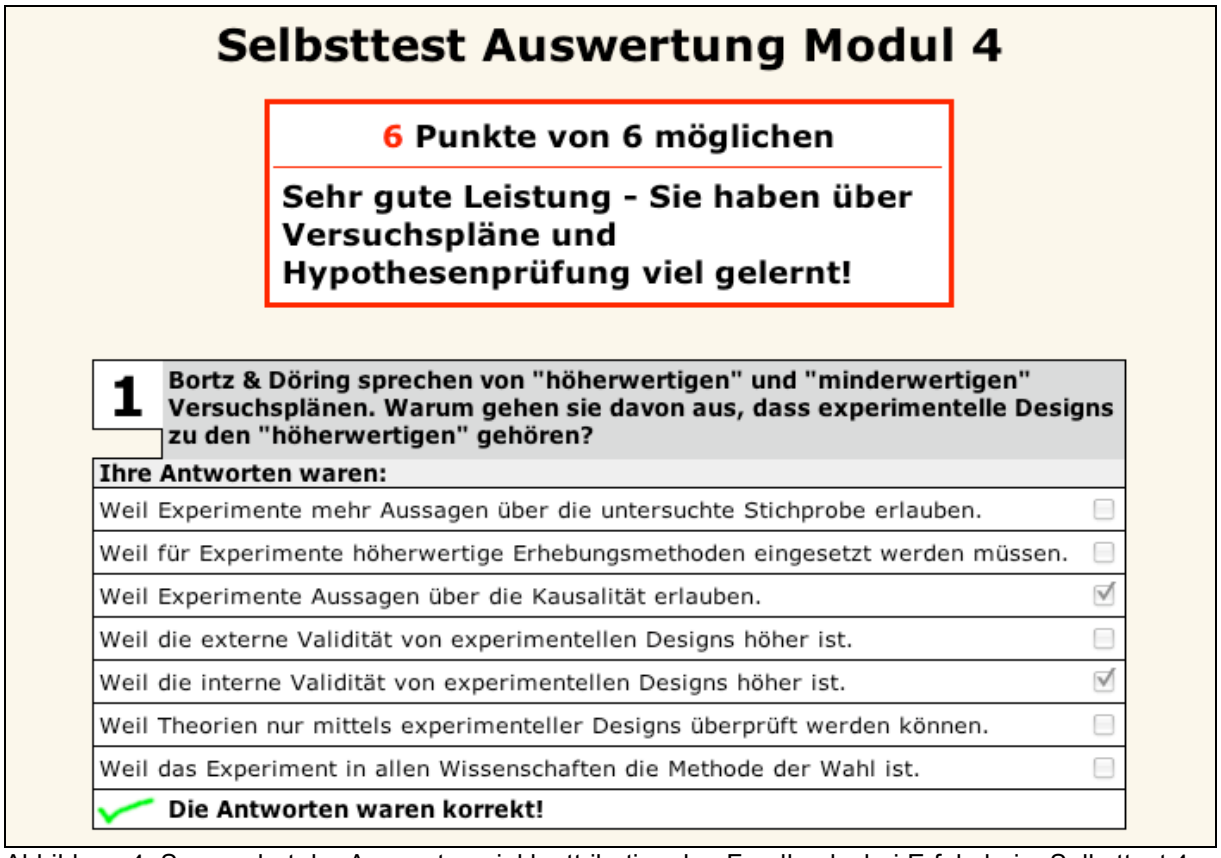


Abbildung 4: Screenshot der Auswertung inkl. attributionalen Feedbacks bei Erfolg beim Selbsttest 4

Die Grundlage für diese attributionalen Feedbacks lieferte nach Schober und Spiel (2005) das Reattributionstraining von Ziegler und Schober (1997, 2001, siehe dazu Kapitel 1.3.3 – Attributionstrainings) und die Ergebnisse von Dresel (2004).

Die Tabelle 5 zeigt eine Übersicht, wann welche Art des attributionalen Feedbacks dargeboten wurde. Zum Beispiel erhielt ein/e Student/in, der/die mehr als 90% der Aufgaben im Selbsttest richtig löste, eine Rückmeldung zu einem großen Erfolg inkl. einem attributionalen Feedback, randomisiert entweder bezogen auf die Anstrengung oder auf die Fähigkeit. Sämtliche ergebnisabhängige attributionale Rückmeldungen aller zehn Selbsttests befinden sich im Anhang 5.2.

Tabelle 5: Übersicht der attributionalen Rückmeldungen in Abhängigkeit der Prozent der richtig gelösten Aufgaben inkl. Beispiele

Prozent der richtig gelösten Aufgaben[9]	**Kategorien und Art des attributionalen Feedbacks**	**Beispiele**
≥ 90 %	Großer Erfolg / randomisierte Vorgabe von Anstrengung und Fähigkeit	Man merkt, Sie haben sich intensiv mit dem Thema auseinandergesetzt.
		Sie beherrschen das offenbar ohne Schwierigkeiten.
		Super, Sie waren bei Ihren Antworten sehr genau.
		Fein, Sie kennen sich offenbar schon recht gut aus.
< 90 %	Erfolg / randomisierte Vorgabe von Anstrengung und Fähigkeit	Wenn Sie über die Aufgaben noch etwas genauer nachdenken, werden Sie sicher alle lösen können.
		Im Wesentlichen haben Sie die wichtigsten Inhalte von Modul 1 verstanden. Wenn Sie noch etwas genauer sind, wird es sicher perfekt.
< 70 %	Tendenzieller Erfolg / v.a. Anstrengung und Fähigkeit; z.T. Aufgabenschwierigkeit	Lesen Sie gezielt nach und probieren Sie es noch mal, Sie schaffen dann sicher noch mehr richtige Lösungen!
		Die Aufgaben sind wirklich nicht so leicht. Nutzen Sie Ihre Fehler, um gezielt Wissenslücken zu schließen!
< 40 %	Misserfolg / Anstrengung; Verweis auf Fähigkeitspotential; z.T. Aufgaben-schwierigkeit	Sie sollten die Literatur noch mal genau durchgehen, dann werden Sie die Aufgaben besser lösen können.
		Sie schaffen sicher mehr, wenn Sie der Sache mehr Aufmerksamkeit widmen!
		Die Aufgaben haben schon ihre Tücken, lesen Sie in der Literatur zu Modul 1 nach und versuchen Sie es noch mal!

Nachdem die Selbsttests und das schriftliche attributionale Feedback des Trainings-programms VEL erläutert wurden, welche im Rahmen dieser Diplomarbeit analysiert und evaluiert wurden, wird nachfolgend eine kurze Einführung zum Thema Evaluation von E-Learning gegeben. Ferner werden Ergebnisse der Evaluation zum Projekt VEL berichtet, sofern sie für die vorliegende Arbeit relevant sind.

[9] Einschränkend ist anzumerken, dass die Einteilung der Prozente der richtig gelösten Aufgaben nicht gänzlich auf die Module 7 und 8 angewendet werden kann: Bei Modul 7 wurden bei 91% richtig gelöster Aufgaben Rückmeldungen zu Erfolg anstatt zu großen Erfolg gemacht. Bei Modul 8 wurden bei 43% richtig gelöster Aufgaben attributionale Feedbacks zu Misserfolg/Anstrengung gegeben und bei 71% Rückmeldungen zu tendenzieller Erfolg anstatt Erfolg (siehe Anhang 5.2).

1.4 Evaluation von E-Learning

In Anbetracht der rasch zunehmenden Zahl an E-Learning-Angeboten wird in der Literatur häufig betont, dass auch die Bedeutung der Evaluation dieser neuen Lernumgebungen zwecks Qualitätssicherung zunimmt, denn die Verbreitung von E-Learning bedeutet nicht nur eine Chance zur didaktischen Verbesserung sondern birgt, vor allem aufgrund noch fehlender Erfahrungen, zunächst auch Risiken (vgl. Kokavecz, Lammers & Holling, 1999; Niegemann et al., 2004; Owston, 2000; Tergan, 2004).

Unter dem Begriff *Evaluation* wird im Allgemeinen die Bewertung des Nutzens oder Wertes eines Gegenstandes verstanden. Solche Evaluationsgegenstände können z.B. Programme, Projekte oder Maßnahmen sein (vgl. Bortz & Döring, 2006; Deutsche Gesellschaft für Evaluation [DeGEval], 2001; Wottawa & Thierau, 1998). *Wissenschaftliche Evaluation* (oder auch *Evaluationsforschung* genannt) ist ein Teilbereich der angewandten Forschung, welche systematisch mit den Methoden der empirischen Forschung die Wirksamkeit von z.B. E-Learning-Angeboten überprüft (vgl. Bortz & Döring, 2006; DeGEval, 2001; Spiel, 2005; Spiel & Strohmeier, 2006).
Evaluationen von E-Learning befassen sich beispielsweise mit Inhalten, die die Auswahl, didaktische Konzeption und Darstellung der Lerninhalte sowie die Usability (z.B. Layout, Bedienungskomfort) betreffen, aber auch Aspekte der Integration in das Gesamtlernsetting (z.B. die Integration eines E-Learning-Programms in Präsenzlehrveranstaltungen). Evaluationsfragestellungen zu E-Learning können zum Beispiel lauten: Entspricht ein E-Learning-Angebot den allgemeinen Standards und hinsichtlich seines Entwicklungstandes dem State-of-the-Art der Instruktionsgestaltung? Ist ein bestimmtes E-Learning-Angebot für die Praxis nützlich (z.B. ist es sinnvoll einsetzbar) und können mit ihm die intendierten Wirkungen (wie z.B. Einstellungs-/Motivationsänderung oder Wissenserwerb) erzielt werden (vgl. Niegemann et al., 2004; Tergan, 2004)? Owston (2000) empfiehlt speziell für die Evaluation von E-Learning einerseits eine verstärkte Kombination von quantitativen und qualitativen Methoden und andererseits die Verwendung und Analyse der unzähligen Log-Dateien[10].

[10] Log-Dateien werden verwendet um bestimmte Vorgänge im Computer zu protokollieren (Winkler, 2006).

1.4.1 Evaluation im VEL

Im Projekt VEL war eine komplexe Evaluationsforschung inkludiert. Diese beinhaltete einerseits eine formative Evaluation[11] der Online-Module und andererseits eine summative Evaluation[12], welche die Wirksamkeit des Programms VEL prüfen sollte, wobei der Schwerpunkt hierbei auf der Zielerreichung der beiden ersten Lehrziele, nämlich der Vermittlung von Fachwissen (Lehrziel 1) und der Vermittlung von Lernkompetenz (Lehrziel 2) lag (vgl. Schober, Reimann et al., 2006; Spiel et al., 2007).

Evaluationen können nach *Kirkpatrick* (1998) auf verschiedenen Ebenen angesetzt sein bzw. stattfinden. Er unterscheidet in seinem Ansatz zur Evaluation von Programmen *vier hierarchische*[13] *Ebenen*. (1) Bei einer Evaluation auf der ersten Ebene, der *Ebene der Reaktion (reaction),* wird untersucht, wie die Teilnehmer/innen auf ein Programm reagieren, es geht demnach um die Zufriedenheit mit der Maßnahme. (2) Bei einer Evaluation auf der zweiten Ebene, der *Ebene des Lernens (learning)*, wird das Ausmaß erfasst, in welchem Personen als unmittelbare Folge der Teilnahme an einer Intervention ihre Einstellungen verändern, ihr Wissen erweitern und/oder ihre Fertigkeiten steigern. (3) Bei Evaluationen auf der dritten Ebene, der *Ebene des Verhaltens (behavior)*, geht es um die Erfassung des Transfers von dem im Programm Gelernten auf Alttagssituationen. Dies wird realisiert durch die Messung der Verhaltensänderung der Teilnehmer/innen. (4) Bei Evaluationen auf der vierten Ebene, der *Ebene der Ergebnisse (results)*, werden Veränderungen, welche auf einer höheren Ebene (des Systems bzw. der Organisation) durch eine Maßnahme erzielt werden, erfasst (vgl. Atria, Reimann & Spiel, 2006; Details dazu siehe Kirkpatrick, 1998). Die Evaluationsfragestellungen im Projekt VEL sind hauptsächlich auf der ersten und zweiten Ebene nach Kirkpatrick (1998) angesiedelt.

[11] Im Rahmen einer formative Evaluation (oder auch begleitende Evaluation genannt) werden regelmäßig Zwischenergebnisse erstellt mit dem Ziel, eine laufende Maßnahme zu modifizieren oder zu verbessern (vgl. Bortz & Döring, 2006; Wottawa & Thierau, 1998).

[12] Bei einer summativen Evaluation wird zusammenfassend die Qualität und Wirksamkeit einer durchgeführten Maßnahme beurteilt. Die Hypothesenprüfung erfolgt typischerweise nachdem die Intervention abgeschlossen ist (vgl. Bortz & Döring, 2006; Wottawa & Thierau, 1998).

[13] Dies bedeutet, dass Evaluationsfragestellungen höherer Ebenen die Beantwortung der Fragestellungen darunter liegender Ebenen voraussetzen. Es ist aber durchaus legitim und in der Praxis zumeist auch der Fall, dass aufgrund eingeschränkter Ressourcen nur Fragen der unteren Ebenen untersucht werden, da der Aufwand und die Komplexität der Evaluationen mit der Höhe der Ebenen zunimmt (Atria, Reimann & Spiel, 2006).

Detaillierte Beschreibungen des Evaluationsmodells befinden sich in Kapitel 2.1.1 – Untersuchungsdesign.

1.4.2 Evaluationsergebnisse zu VEL

Im Folgenden werden nun exemplarisch einige der zahlreichen Ergebnisse der Evaluationsforschung zum Programm VEL berichtet.

Für die *formative Evaluation* wurden die Studierenden der "Intensiv-Variante" (Track A) über einen modulintegrierten Online-Fragebogen (über die Online-Seite *Ihre Meinung*) zu jedem Modul befragt (Schober, Reimann et al., 2006). Im Durchschnitt beteiligten sich 60-75% der Teilnehmer/innen an der Bewertung der Module, wobei ein abnehmendes Engagement jeweils gegen Ende des Semesters – im Sommersemester deutlicher als im Wintersemester – zu verzeichnen war (Schober & Spiel, 2005). Die Studierenden sollten unter anderem angeben wie lange sie das jeweilige Modul bearbeitet hatten. Im Durchschnitt beanspruchten die Module 6-8 Stunden. Bis auf die beiden Ausnahmen Modul 7 und 8, für welche durchschnittlich ca. 11 Stunden aufgewendet wurden (Schober & Spiel, 2005). Die Ursache für den deutlich höheren Zeitaufwand für diese Module liegt vermutlich darin, dass in diesen Modulen englische Literatur zu bearbeiten sowie statistische Auswertungen durchzuführen waren (Spiel et al., 2007). Weiters sollten die Studierenden die einzelnen Komponenten der Module (z.B. Selbsttest) nach subjektiver Wichtigkeit rangreihen. Im gesamten Durchschnitt (d.h. über alle Studierenden und über alle Module hinweg) wurde die Modulkomponente Task als am wichtigsten beurteilt, gefolgt von der Bearbeitung der Literatur. Die Komponenten Selbsttest, Feedback und SRL-Exercise lagen im mittleren Bereich (Schober & Spiel, 2005).

Weiters wurden im Rahmen der Evaluation vom Trainingsprogramm VEL 20 Studierende der "Intensiv-Variante" (Track A) mittels leitfadenorientiertem Interviews befragt (Schober & Spiel, 2005). Anhand der erfragten Daten konnten drei verschiedene Modulbearbeitungstypen der Studierenden herausgearbeitet werden: (1) Die Musterstudent/inn/en (25%), die sofort nach Modulbeginn das Modul bearbeiteten, alle Inhalte lasen, die Literatur genauestens studierten und reflektierten. (2) Die Gemütlichen (35%), welche die Module nicht sofort nach

Modulbeginn ansahen, die ersten drei Modulkomponenten (vgl. Kapitel 1.2.4 – Aufbau von VEL) meist nur überflogen, die Literatur mittelmäßig studierten und nicht reflektierten. (3) Die dritte Gruppe, ist jene der Wankelmütigen (30%), diese riefen manchmal bald, manchmal spät das Modul auf, lasen manchmal die Komponenten des Moduls, manchmal ignorierten sie diese. Die Literatur bearbeiteten und reflektierten sie, je nach Zeitressourcen und Motivation, von genau bis oberflächlich (Grünberger, 2005; Schober & Spiel, 2005).
Der Modulkomponente Selbsttest gegenüber war die Mehrheit der im Interview Befragten (17 von 20) positiv eingestellt und 18 der Befragten beurteilten den Selbsttest als sinnvoll (Grünberger, 2005; Schober & Spiel, 2005). Zur sprachlichen Mitteilung (i.e. das attributionale Feedback) jedoch waren lediglich fünf Befragte positiv eingestellt. Drei der befragten 20 Personen empfanden das schriftliche attributionale Feedback als motivierend (Grünberger, 2005; Schober & Spiel, 2005). Ferner wurden im Interview Unterschiede zwischen Wintersemester (WS) und Sommersemester (SS) genannt. Fünf der 20 befragten Studierenden des Track A klagten über einen Motivationsverlust im Sommersemester, sieben empfanden das Sommer-semester als schwieriger und aufwendiger und acht gaben an, die Module im Sommersemester oberflächlicher bearbeitet zu haben als im Wintersemester. Fünf Personen gaben an, dass es keine Unterschiede in ihrer Modulbearbeitung zwischen den beiden Semestern gab (Grünberger, 2005; Schober & Spiel, 2005).

Wie soeben gezeigt wurde, gibt es schon viele Evaluationsergebnisse zum Projekt VEL. Für die Beantwortung der Fragestellungen wurden subjektive Daten sowie objektiv erhobene Daten herangezogen. Ein wesentliches Ziel dieser Diplomarbeit war es, bzgl. der Modulkomponente Selbsttest inkl. attributionalem Feedback nicht nur die subjektive Einschätzung der Studierenden in die Evaluation mit einzubeziehen, sondern zusätzlich objektive Daten, die laufend gespeicherten Zugriffsdaten der Studierenden auf die Selbsttests, zu analysieren. Somit wurde eine Evaluation auf der Ebene des Lernens realisiert, welche bei Evaluationen von E-Learning häufig verabsäumt wird. Viele E-Learning-Angebote werden häufig gar nicht (Paulus & Strittmatter, 2002) oder nur auf der ersten Ebene nach Kirkpatrick (1998), der Ebene der Reaktion, evaluiert (z.B. Baeßler et al., 2003; Reinmann-Rothmeier, 2003; Söllner, Flöter & Scheibner, 2006). Im nächsten Kapitel werden nun die Ziele und Fragestellungen der vorliegenden Arbeit im Detail beschrieben.

1.5 Ziele und Fragestellungen

Diese Diplomarbeit verfolgte im Wesentlichen zwei Ziele. Das *erste Ziel* war die Analyse der Selbsttestnutzung im Rahmen der Lehrveranstaltung "Forschungsmethoden und Evaluation" und die Überprüfung eines Zusammenhangs der Selbsttestnutzung mit den erreichten Leistungen in den Lehrveranstaltungsprüfungen. Das *zweite Hauptziel* der vorliegenden Arbeit war die summative Evaluation des neu entwickelten schriftlichen attributionalen Feedbacks.

Fragestellungen ad Ziel 1

Die Analyse der Selbsttestnutzung beinhaltete die Ermittlung, wie oft und wann die Selbsttests im Rahmen der Lehrveranstaltung "Forschungsmethoden und Evaluation" (FM&E) durchgeführt worden sind. Die ausformulierten Fragestellungen zur Häufigkeit und zum Zeitpunkt der Selbsttestnutzung lauten:

1a) Wie oft wurden die Selbsttests von den Studierenden durchgeführt?
1b) Wann wurden die Selbsttests von den Studierenden durchgeführt?

Weiters wurde überprüft, ob es in der Durchführungshäufigkeit Unterschiede zwischen den einzelnen Selbsttests, den Semestern und den beiden Varianten von VEL (Track A und Track B) gibt. Folgende Fragestellungen wurden hierfür formuliert:

1c) Gibt es Unterschiede in der Durchführungshäufigkeit der zehn Selbsttests?
1d) Gibt es einen Unterschied in der Durchführungshäufigkeit zwischen den Selbsttests des Wintersemesters und denen des Sommersemesters?
1e) Gibt es einen Unterschied zwischen den Studierenden des Track A und Track B in der Durchführungshäufigkeit der Selbsttests zusammengefasst für das gesamte Studienjahr, zusammengefasst für die beiden Semester bzw. in den einzelnen zehn Selbsttests?

Aufgrund bisheriger Ergebnisse einerseits von der Evaluation des VEL-Programms (z.B. Schober & Spiel, 2005; siehe Kapitel 1.4.2 – Evaluationsergebnisse zu VEL) und andererseits von Befunden zu Dropout bei E-Learning (vgl. Doherty, 2006; Liu,

Gomez, Khan & Yen, 2007) gibt es Hinweise, dass die Durchführungshäufigkeit der Selbsttests im Laufe des Semesters abnimmt und im Sommersemester geringer ist als im Wintersemester. Ferner wird angenommen, dass die Studierenden des Track A die Selbsttests regelmäßiger und somit häufiger durchführen als die Studierenden des Track B, weil sie durch die "Intensiv-Variante" inkl. der Gruppenarbeiten sozial stärker eingebunden sind als die Studierenden des Track B (vgl. Liu et al., 2007).

Weiters war von Interesse, ob und wie sich die Selbsttestnutzung auf die Ergebnisse in den Lehrveranstaltungsprüfungen auswirkt. Hierfür wurde zunächst analysiert, ob es unterschiedliche Typen der Selbsttestnutzung gibt, welche anschließend miteinander verglichen wurden, einerseits bzgl. ihrer Trackzugehörigkeit und andererseits hinsichtlich der erreichten Leistung in den Lehrveranstaltungsprüfungen. Die Fragestellungen dazu lauten:

1f) Gibt es unterschiedliche Typen von Selbsttestnutzer/inne/n?
1g) Gibt es einen Unterschied zwischen Track A und Track B in der Art der Selbsttestnutzung?
1h) Gibt es aufgrund unterschiedlicher Selbsttestnutzung einen Unterschied in der Prüfungsleistung in F&E I bzw. II (gesamt und getrennt nach Reproduktion, Verständnis und Produktion)?

Ergebnisse aus der Forschungsliteratur zu Lerngewohnheiten von Schüler/inne/n zeigen, dass es unterschiedliche Lerntypen gibt, wie z.B. die "Minimalist/inn/en" und "Prüfungslerner/innen" (auch "Saisonarbeiter/innen" genannt), und dass z.B. die Gruppe der "Saisonarbeiter/innen" schlechtere Prüfungsleistungen erzielt als andere Lerntypen (vgl. Haag & Mischo, 2002; Wagner, Schober & Spiel, 2004). Dies ist ein Hinweis, dass es vermutlich auch unterschiedliche Typen von Lerner/inne/n bei Studierenden gibt, welche wiederum auch unterschiedliche Leistungen in den Lehrveranstaltungsprüfungen erzielen.

Fragestellungen ad Ziel 2

Das zweite Hauptziel dieser Diplomarbeit, die Evaluation des attributionalen Feedbacks, beinhaltete die Überprüfung der Wirksamkeit des neu entwickelten schriftlichen attributionalen Feedbacks – das den Studierenden der Versuchsgruppe (Track A) dargeboten wurde, nicht aber der Kontrollgruppe (Track B) – in Bezug auf Veränderungen von Erfolgs- und Misserfolgsattributionen. Daraus lässt sich folgende Fragestellung ableiten:

2a) Verbessert das schriftliche attributionale Feedback die Erfolgs- bzw. Misserfolgsattributionen der Studierenden?

Bei Reattributionstrainings werden oft neben dem vorrangigen, direkten Ziel der Veränderung des Attributionsstils auch indirekte Ziele, wie die positive Veränderungen in den Konstrukten Implizite Persönlichkeitstheorien, Selbstwirksamkeitserwartungen/Erfolgserwartungen, Selbstkonzept des/r Lernenden, Interessen und Hilflosigkeit angestrebt (vgl. Dresel, 2004; Schober, 2002). So wird auch in der vorliegenden Arbeit der Einfluss des schriftlichen attributionalen Feedbacks auf die oben genannten Konstrukte überprüft. Die Fragestellungen dazu lauten:

2) Bewirkt das schriftliche attributionale Feedback eine positive Veränderung in den folgenden Konstrukten:
   - b) Implizite Persönlichkeitstheorien,
   - c) Selbstwirksamkeitserwartungen,
   - d) Selbstkonzept des/der Lernenden,
   - e) Interesse,
   - f) Hilflosigkeit?

2 Empirischer Teil

2.1 Untersuchungsplan

Im Folgenden werden das Untersuchungsdesign, die Stichproben und die verschiedenen Erhebungsinstrumente dargestellt. Hierbei ist zu betonen, dass grundlegende Aspekte der folgenden Kapitel im Kontext des Projekts VEL im Arbeitsbereich Bildungspsychologie und Evaluation entwickelt wurden, wie z.B. der unten erwähnte Fragebogen. Hierauf greift die vorliegende Forschungsarbeit zurück. Die Untersuchungsmethodik der Analyse der Selbsttestnutzung und der Evaluation des attributionalen Feedbacks ist eine Neuentwicklung im Rahmen dieser Diplomarbeit.

2.1.1 Untersuchungsdesign

Es wurden für die Beantwortung der Fragestellungen die Daten der quasi-experimentellen Untersuchung des Projekts VEL herangezogen, die in den Studienjahren 2003/04 und 2004/05 durchgeführt wurde, hierbei wurde ein Zwei-Gruppen-Prätest-Posttest1-Posttest2-Design realisiert. Die Studierenden konnten selbst entscheiden, an welcher Variante, entweder an der “Intensiv-Variante“ (Track A) oder an der “Standard-Variante“ (Track B), sie teilnahmen (vgl. Kapitel 1.2.4 – Aufbau von VEL). Es wurden beide Varianten vor Semesterbeginn angekündigt und ausführlich erläutert. Für die “Intensiv-Variante“ (Track A) war eine schriftliche, individuelle Anmeldung mit Begründung erforderlich (Spiel et al., 2007; Schober, Reimann et al., 2006). Abbildung 5 veranschaulicht die Versuchsanordnung der vorliegenden Forschungsarbeit.

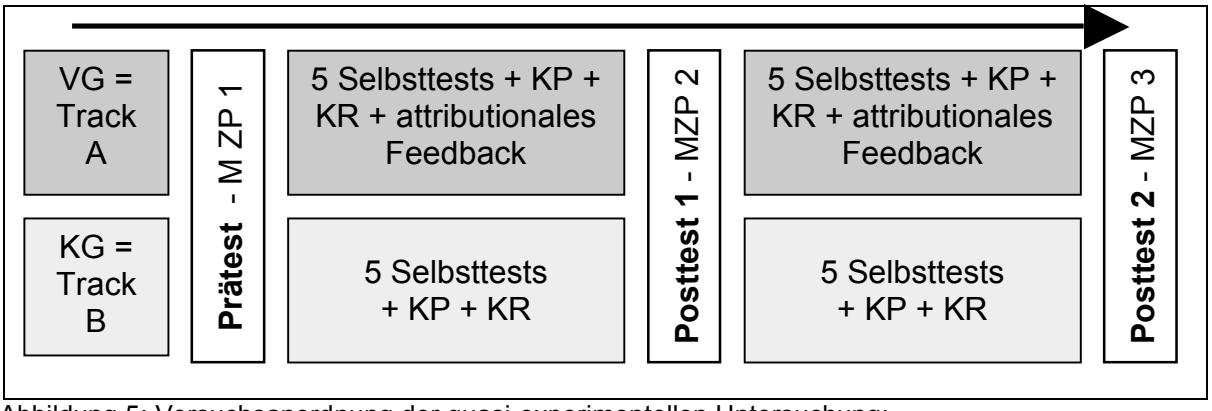


Abbildung 5: Versuchsanordnung der quasi-experimentellen Untersuchung;
Legende: VG = Versuchsgruppe; KG = Kontrollgruppe; MZP = Messzeitpunkt; KP = Knowledge of Performance; KR = Knowledge of Result/Response

Wie oben berichtet, war es nicht möglich die Studierenden randomisiert einer der beiden Gruppen (Track A oder Track B) zuzuteilen, was bei Interventionen im Feld häufig der Fall ist (Spiel et al., 2008). Dies hatte jedoch Folgen für die Wirksamkeitsprüfung von VEL, weil davon auszugehen ist, dass die getroffene Entscheidung für eine der beiden Varianten nicht per Zufall erfolgte, und zu erwarten ist, dass die Effekte des Treatments mit Faktoren konfundiert sind, die ausschlaggebend für die Selbstselektion zur Versuchs- oder Kontrollgruppe sind (Spiel et al., 2007). Für die Evaluation von VEL war für die Adjustierung der Unterschiede, die zwischen Versuchs- und einer nicht randomisierten Kontrollgruppe in der Baseline bestehen, das statistische Matching[14] die Methode der Wahl. Hierfür wurde ein neues Verfahren, die Euclidean Distance Matching (EuM) Prozedur, entwickelt (siehe Spiel et al., 2008). Unter der Verwendung Euklidischer Distanzen wurden Studierende aus Versuchs- und Kontrollgruppe auf Basis der für das Treatment relevanten beobachteten Merkmalen gepaart (Spiel et al., 2007; siehe Kapitel 2.1.2 – Stichproben).

2.1.2 Stichproben

In der vorliegenden Arbeit werden zwei Stichproben für die Beantwortung der Fragestellungen herangezogen. Für die deskriptive Darstellung der Selbsttest-nutzung, z.B. zur Veranschaulichung des Verlaufes der ST-Durchführungshäufigkeit, werden die Daten der Studierenden nur eines Studienjahres, nämlich die des Studienjahres 2003/04, herangezogen (**Stichprobe 1**). Denn aufgrund der unterschiedlichen Tage und Kalenderwochen an denen Ferien und vorlesungsfreie Zeiten innerhalb der Studienjahre stattfinden (dies betrifft vor allem das Sommersemester mit den Oster- und Pfingstferien) würde es zu einer “Verwischung“ der Diagramme der Selbsttestnutzung kommen, wenn man für die Darstellung der Durchführungszeitpunkte die Daten von zwei Jahren zusammenfügen würde.

[14] Unter Matching versteht man das Auswählen von Personen aus der Versuchs- und der Kontrollgruppe dahingehend, dass in den Variablen, von denen man annimmt, dass sie die abhängigen Variablen beeinflussen, keine Unterschiede zwischen den Gruppen bestehen (Howitt & Cramer, 2000; Bortz & Döring, 2006).

Für die Beantwortung der Fragestellungen zur Selbsttestnutzung und Prüfungsleistung mittels Inferenzstatistik sowie zur Evaluation des attributionalen Feedbacks werden die Daten einer **Matching-Stichprobe** (**Stichprobe 2**) herangezogen. Bei dieser Stichprobe besteht die Versuchsgruppe aus den Studierenden, die an der "Intensiv-Variante" (Track A) im Studienjahr 2003/04 bzw. 2004/05 teilgenommen haben. Die Kontrollgruppe wurde aus dem Pool der Studierenden der "Standard-Variante" (Track B) des Studienjahres 2003/04 mit Hilfe der Euclidean Distance Matching (EuM) Prozedur hinsichtlich der Ähnlichkeit zur Versuchsgruppe in den Konstrukten Vorwissen, Interesse, Monitoring, Organisation, Zeit- und Arbeitsmanagement gezogen (siehe Spiel et al., 2008).

Ferner wurde für die inferenzstatistische Absicherung der Anwendbarkeit der Matching-Stichprobe (siehe Kapitel 2.3.3 – Prüfung der ST-Durchführungs-häufigkeits-Gleichheit) eine **"Gesamtstichprobe"** erstellt, welche alle Studierenden der Stichprobe 1 und der Stichprobe 2 sowie die Studierenden des Track B des Studienjahres 2004/05, welche an einem der ersten drei Prüfungsterminen teilgenommen haben, beinhaltet.

2.1.3 Erhebungsinstrumente

Im Folgenden werden die Erhebungsinstrumente vorgestellt. Diese sind (a) die Selbsttests und die Speicherung deren Nutzungsdaten, (b) der Fragebogen zum Selbstregulierten Lernen inkl. der Items für die Erfolgs- und Misserfolgsattributionen sowie (c) die Vorlesungsprüfungen zur Erhebung des Fachwissens.

a) Selbsttests und Speicherung der Nutzungsdaten

Die zehn Selbsttests enthielten jeweils sechs bis 11 Multiple-Choice-Aufgaben. Ein Beispiel einer Multiple-Choice-Aufgabe befindet sich im Kapitel 1.3.4 – Selbsttests und attributionales Feedback in VEL. Die Selbsttests wurden in HTML und PHP programmiert und den Studierenden ab einem bestimmten Zeitpunkt auf dem Server der Universität Wien zur Verfügung gestellt. Der Zugang zu den Selbsttests erfolgte über eine passwortgeschützte Lernplattform (siehe Kapitel 1.2.2 – Technische Umsetzung von VEL). Die Studierenden konnten die Selbsttests über einen

konventionellen Webbrowser jederzeit und so oft sie wollten, online durchführen. Die Speicherung der Selbsttestnutzungsdaten erfolgte mit Betätigung des Buttons “Auswerten“. Gleichzeitig erfolgte mittels PHP-Skript die Auswertung des Selbsttests und anschließende Anzeige des automatisch erstellten (attributionalen) Feedbacks. Die Selbsttestnutzungsdaten der Studierenden wurde in einer *MySQL-Datenbank* auf dem Server der Universität Wien gespeichert und mit der Software *phpMyAdmin* (Version 2.8.2.4) verwaltet. Die genaue Beschreibung der Aufbereitung und Komprimierung der Nutzungsdaten der Selbsttests erfolgt in Kapitel 2.2.1.

b) Fragebogen zum Selbstregulierten Lernen (SRL-Fragebogen)

Als weiteres Erhebungsinstrument diente ein Fragebogen, welcher bei allen Messzeitpunkten (MZP) der Evaluation in den Studienjahren 2003 bis 2005 vorgegeben wurde. Dieser Fragebogen (“SRL-Fragebogen“) umfasst u.a. folgende Bereiche: Implizite Persönlichkeitstheorien, Selbstwirksamkeitserwartungen, Selbstkonzept Lerner/in, Interesse, Hilflosigkeit, Lernstrategien etc.

Nach Schober und Spiel (2005) wurden die meisten Items des SRL-Fragebogens aus bereits publizierten Fragebögen entnommen. Als Basis dienten u.a. Fragebögen wie *Schulbezogene Selbstwirksamkeitserwartung* (Jerusalem & Satow, 1999), der *Fragebogen zum Studieninteresse* (FSI; Schiefele, Krapp, Wild & Winteler, 1993), *Implizite Persönlichkeitstheorien* (Dweck, 1999; Schober, 2002), *Motivationale Orientierung* (Dweck, 1999; Schober, 2002), die *Hilflosigkeitsskala* (HiS; Breitkopf, 1985), das *Lernstrategieinventar für Studentinnen und Studenten* (LASSI – Learning and Study Strategies Inventory; Metzger, Weinstein & Palmer, 2002), *Lernstrategien im Studium* (LIST; Wild, 2000), *Bedingungen Motivierten Selbstgesteuerten Lernens – Instrument zur Erfassung an der Hochschule* (BEMSEL-IHS; Wosnitza, 2000), *Selbstregulation* (Schwarzer, 1999), sowie der *Fragebogen Selbstregulation* (Schmitz, Perels, Bruder & Otto, 2003). Das Antwortformat der Items bestand aus einer sechsstufigen Ratingskala mit den Polen “nie“ (1) und “fast immer“ (6) bzw. “stimme gar nicht zu“ (1) und “stimme völlig zu“ (6) (Schober & Spiel, 2005).

Zur Veranschaulichung der Items befindet sich der SRL-Fragebogen des zweiten Messzeitpunktes des Studienjahres 2003/04 im Anhang 5.3. Ebenso befindet sich im Anhang 5.4 die Tabelle 49 mit einer Übersicht der Skalen des SRL-Fragebogens inkl. Itemanzahl und Reliabilitätskoeffizienten. Ingesamt umfasste der SRL-Fragebogen über 70 Items. Die Skalen des SRL-Fragebogens beinhalteten jeweils von zwei bis sechs Items. Die interne Konsistenz der einzelnen Skalen ist nach Bortz & Döring (2006) niedrig bis mittelmäßig (Cronbachs α = .62 - .84).

Items für die Erfolgs- und Misserfolgsattributionen

Neben den Items zu verschiedenen Aspekten des Selbstregulierten Lernens beinhaltete der SRL-Fragebogen auch zwei Fragen mit freiem Antwortformat hinsichtlich bevorzugter Attributionen in Erfolgs- und Misserfolgssituationen (siehe Abbildung 6 und 7).

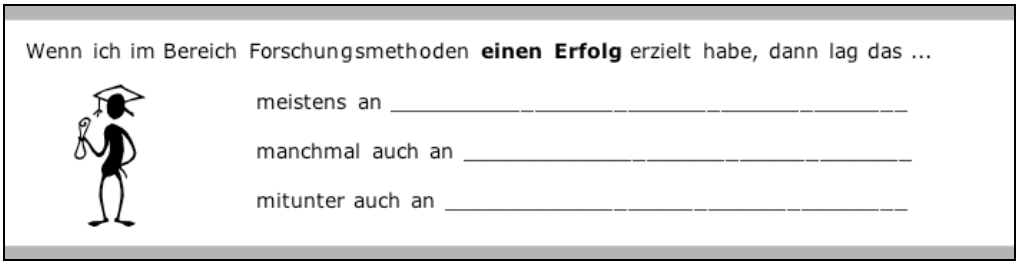


Abbildung 6: Item zur Erhebung der Erfolgsattributionen

Abbildung 7: Item zur Erhebung der Misserfolgsattributionen

Für die Kategorisierung der Antworten der Studierenden auf diese offenen Fragen, wurde ein Kategorienschema herangezogen (siehe Anhang 5.5). Die Güte dieses Kategorienschemas wurde mittels Beurteilerübereinstimmung anhand zweier unabhängiger Rater/innen bestimmt. Der Kappa-Koeffizient nach Cohen (1960, zitiert nach Bortz & Döring, 2002) beträgt .901 (*Näherungsweise T* = 36.29; $p < .001$). Laut Bortz und Döring (2002) erfordert eine gute Übereinstimmung Kappa-Koeffizienten über .70. Im vorliegenden Fall besteht also eine gute Übereinstimmung zwischen den beiden unabhängigen Rater/inne/n, sodass angenommen werden kann, dass es sich hierbei um ein ausgereiftes und nachvollziehbares Kategorienschema handelt.

c) Vorlesungsprüfungen

Das Fachwissen wurde mittels der Vorlesungsprüfungen zu Forschungsmethoden und Evaluation I und II (Posttest 1 und Posttest 2) erhoben. Bei der Konzeption der Prüfungsfragen war darauf geachtet worden, dass die Prüfungsfragen die drei Dimensionen des Wissens "Reproduktion", "Verständnis" und "Produktion"[15], angelehnt an die Lernzieltaxonomie nach Bloom (1956), erfassen.

Zur Vereinheitlichung der Ergebnisse wurden die Prüfungspunkte der gesamten Prüfung bzw. der einzelnen Kategorien ("Reproduktion", "Verständnis" und "Produktion") auf Prozentwerte umgerechnet (*POMP – percent of maximum points*), d.h. es wurde die erreichte Punkteanzahl jedes/r Studierenden in der Prüfung (insgesamt und in der jeweiligen Kategorie) durch die erreichbare Höchstpunkteanzahl (der Prüfung bzw. der jeweiligen Kategorie) dividiert und mit 100 multipliziert. Da sich während der Auswertung der summativen Evaluation im Projekt VEL herausstellte, dass der Mittelwert der maximal erreichten Prüfungspunkte (von Track A und B) des Studienjahres 2004/05 wesentlich unter dem Mittelwert des Studienjahres 2003/04 lag[16], wurden nachträglich die Prozentwerte des Studienjahres 2004/05 auf jene des Studienjahres 2003/04 mit folgender Vorgehensweise normiert: Zunächst wurden die Prüfungspunkte standardisiert, d.h. es wurden z-Werte berechnet nach der Formel: $z2=(x2-\mu2)/\sigma2$ (2 bedeutet hier "des Studienjahres 2", i.e. 2004/05). Anschließend wurden diese Werte normiert auf das Studienjahr 2003/04, mit der Formel: $(z2*\sigma1)+\mu1=x1$ (2 bedeutet wiederum "des Studienjahres 2" und 1 bedeutet "des Studienjahres 1", i.e. 2003/04). Für die Berechnung der Werte (z-Werte, MW, SD) wurden jeweils alle zur Prüfung angetretenen Studierenden (also Track A und B) herangezogen. Die Normierung wurde für WS und SS jeweils für alle Kategorien (Reproduktion, Verständnis, Produktion) und für den Gesamtwert durchgeführt. Vorteil dieser Methode war, dass die Darstellung in Prozentwerten beibehalten werden konnte. Tabelle 6 zeigt die Deskriptivstatistik der Leistungen im Vorwissenstest und in den Prüfungen nach der Normierung am ersten Studienjahr 2003/04, gesamt und nach Kategorien aufgeteilt.

[15] An dieser Stelle sei erwähnt, dass beim 1. Prüfungstermin des SS 2004 zur Kategorie Produktion keine Items vorgegeben wurden, bei allen anderen Erhebungen wurden alle drei Dimensionen des Wissens erfasst.

[16] Die Gründe dafür dürften vermutlich in einer anderen Gestaltung der Prüfungsfragen (diese waren im Studienjahr 2004/05 eindeutiger zum Auswerten) und genaueren Auswertungskriterien liegen.

Tabelle 6: Anzahl der Personen, deren Daten in die vorliegende Deskriptivstatistik eingegangen sind (N), Mittelwerte (M), Standardabweichungen (SD), Mediane (50. Perzentil), 25. und 75. Perzentil, Minima (Min) und Maxima (Max) der Prüfungsleistung (normierten Prozentwerte) der Gesamtpopulation der Studienjahre 2003/04 und 2004/05

				Perzentile				
	N	*M*	*SD*	*25.*	*50.*	*75.*	*Min*	*Max*
Prüfungsleistung gesamt (MZP 2)	658	72.79	10.95	65.24	73.49	80.55	30.12	97.70
Reproduktion (MZP 2)	658	77.84	13.23	69.89	78.45	87.78	27.12	101.26
Verständnis (MZP 2)	658	74.90	11.20	67.32	75.49	83.65	39.39	100.99
Produktion (MZP 2)	658	60.24	20.84	45.99	62.50	75.00	5.00	100.32
Prüfungsleistung gesamt (MZP 3)	485	60.45	15.54	50.00	60.26	71.85	11.28	98.00
Reproduktion (MZP 3)	485	58.26	19.08	46.67	59.13	72.41	-3.90	96.55
Verständnis (MZP 3)	485	62.57	23.02	45.00	61.37	80.26	8.33	108.60
Produktion (MZP 3)	344	51.74	28.29	29.59	54.16	75.45	-3.16	100.00

2.2 Untersuchungsdurchführung

In diesem Kapitel wird die Untersuchungsdurchführung inkl. der Aufbereitung und Komprimierung der Selbsttestdaten und der Zusammenfassung der Attributionen näher erläutert.

Die Tabellen 7 und 8 zeigen die Übersicht der Freischaltung der Module und Selbsttests exemplarisch für das Studienjahr 2003/04 (Tabellen für das Studienjahr 2004/05 - siehe Anhang 5.6). Die Freischaltung der Module und Selbsttests erfolgte für Versuchsgruppe (Track A) und Kontrollgruppe (Track B) parallel.
Die Studierenden wurden in der “Kick-off“-Veranstaltung informiert, dass ihre Daten der Selbsttestnutzung auf einer Datenbank gespeichert, aber streng vertraulich behandelt werden würden und keine Auswirkungen auf das Ergebnis der Vorlesungsprüfungen zu “Forschungsmethoden und Evaluation“ haben würden.

Tabelle 7: Übersicht der Freischaltung der Module und Selbsttests mit Angabe der Kalenderwochen (KW) für das WS 2003/04

43	44	45	46	47	48	49	50	51	52	1	2	3	4
M 1													
S 1													
		M 2											
			S 2										
				M 3									
					S 3								
						M 4		Weihnachtsferien					
							S 4						
										M 5			
										S 5			Prüf.

Tabelle 8: Übersicht der Freischaltung der Module und Selbsttests mit Angabe der Kalenderwochen (KW) für das SS 2004

11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
M 6															
	S 6														
		M 7	Osterferien												
				S 7											
						M 8									
							S 8								
									M 9						
											S 9				
												M 10			
													S 10		Prüf.

Die Selbsttests und das attributionale Feedback sind im Detail in dem Kapitel 1.3.4 – Selbsttest und attributionales Feedback in VEL sowie im Kapitel 2.1.3 – Erhebungsinstrumente beschrieben.

Der Prätest (MZP 1) wurde am Beginn des Wintersemesters (WS) in der “Kick-off“-Veranstaltung durchgeführt (siehe Tabelle 2 in Kapitel 1.2.4 – Aufbau von VEL). Der Posttest 1 (MZP 2) erfolgte anschließend an die Vorlesungsprüfungen zu “Forschungsmethoden und Evaluation I“ und der Posttest 2 (MZP 3) anschließend an die Vorlesungsprüfungen zu “Forschungsmethoden und Evaluation II“ (siehe ebenso Abbildung 5 - Versuchsanordnung).

2.2.1 Aufbereitung und Komprimierung der Daten

Durch die automatische Abspeicherung der Nutzung der Selbsttests entstand eine beinahe unüberschaubar große Menge an Daten, welche speziell aufbereitet und komprimiert werden musste, bevor sie in die statistischen Analysen einbezogen werden konnte. Ebenso mussten die Rohdaten der Attributionen zusammengefasst werden, bevor sie in die inferenzstatistischen Analysen mit eingingen. Für die

Aufbereitung, Komprimierung und Zusammenfassung der Daten wurden die Computer-Programme Excel 2003 und SPSS 11.5 verwendet.

a) Aufbereitung der Selbsttestnutzungsdaten

Wie im Kapitel 2.1.3 – Erhebungsinstrumente schon beschrieben, wurde in der MySQL-Datenbank bei jedem Absenden eines Selbsttests jeweils eine Zeile abgespeichert. Diese Zeile beinhaltete den Code bzw. die Matrikelnummer, welche/r von dem/r Studierenden eingegeben wurde, die IP-Adresse des verwendeten Rechners und die erreichte Punkteanzahl im Selbsttest des jeweiligen Versuches. Weiters wurden das aktuelle Datum und die Uhrzeit, zu der die PHP-Seite an den Server gesendet wurde, sowie die Korrektheit der Antwort auf die jeweilige Multiple-Choice-Frage abgespeichert.

Diese Daten wurden aus der MySQL-Datenbank in Excel-Files abgespeichert. Daraus resultierten 2 x 10 Excel-Files mit jeweils durchschnittlich ca. 2000 Zeilen. Diese Excel-Files wurden zu Beginn speziell aufbereitet. Beispielsweise wurde die Spalte mit dem “Code“ (z.B. b1234567) umgewandelt zu zwei Spalten. Die erste Spalte beinhaltete die Gruppe (Track A bzw. B) als numerischer Wert (a = 1 bzw. b = 2), die zweite Spalte die Matrikelnummer (z.B. 1234567). Anschließend wurden die optimierten Excel-Files ins SPSS eingelesen und als SPSS-Files abgespeichert. So ergaben sich 2 x 10 SPSS-Files mit der gleichen Zeilenanzahl.

Im Original-Datensatz und auch in den daraus generierten SPSS-Files sind pro Person mehrere Zeilen vorhanden, d.h. je ausgefüllten Selbsttest einer Person eine Zeile, da bei jedem Mal Absenden der PHP-Seite eine neue Zeile in der Datenbank abgespeichert wird. Mit der SPSS-Funktion *Umstrukturieren* wurden deshalb neue SPSS-Files generiert, worin jede Person nur noch eine Zeile ausmachte inkl. der komprimierten Informationen aller Zeilen des jeweiligen ursprünglichen SPSS-Files. So beinhaltete eine Zeile nach der Umstrukturierung (1) die Matrikelnummer des/r Studierenden, (2) die Häufigkeit der Selbsttestdurchführung für diesen einen Selbsttest innerhalb eines bestimmten Zeitraums (siehe unten – Weg 1) und (3) aufsteigend gereiht das Datum der Selbsttestdurchführung. Wiederum gab es pro Selbsttest ein SPSS-File (d.h. 2 x 10 umstrukturierte SPSS-Files).

Die wie oben beschrieben umstrukturierten SPSS-Files wurden anschließend für die jeweilige Fragestellung speziell komprimiert (siehe unten im Unterkapitel b) – Komprimierung der SPSS-Files nach bestimmten Kriterien) und im nächsten Schritt mit den dafür speziell vorbereiteten SPSS-Datenfiles der Studierenden des Studienjahres 2003/04 (Stichprobe 1), der Matching-Stichprobe (Stichprobe 2) und des Studienjahres 2004/05 (für die Gesamtstichprobe) zusammengespielt. Die SPSS-Files mit den Daten der Studierenden entstammen großteils der Evaluation des Programms VEL. Diese Files umfassen soziodemografische Daten (wie Geburtsjahr und Geschlecht) sowie alle im Rahmen der summativen Evaluation erhobenen Daten (wie z.B. Fachwissen, Erfolgs- und Misserfolgsattributionen).

Als zweite Quelle für Daten von den Studierenden (vor allem für die Studierenden des Track B des Studienjahres 2004/05) diente die Datenbank der Lernplattform *Blackboard Vista* (siehe Kapitel 1.2.2 – Technische Umsetzung von VEL). Diese beinhaltete Matrikelnummer, Track-Zuordnung und Geschlecht (abgeleitet vom Namen) der Studierenden. Angemerkt soll an dieser Stelle werden, dass nach dem Zusammenspielen der Selbsttest-Datenfiles mit den Studierenden-Datenfiles alle Studierende aus dem Datensatz gefallen sind, welche zwar die Selbsttests durchgeführt hatten, aber entweder nicht identifiziert werden konnten, da sie z.B. einen ungültigen Code eingegeben hatten, oder nicht zu einer der beiden Stichproben (Stichprobe 1 bzw. 2) gehörten, bzw. nicht im Studienjahr 2004/05 waren und bei einem der ersten drei Prüfungsterminen der Lehrveranstaltung “Forschungsmethoden & Evaluation I und II“ angetreten waren.

b) Komprimierung der SPSS-Files nach bestimmten Kriterien

Im Folgenden wird nun die Komprimierung der SPSS-Files im Detail erläutert, diese erfolgte mehrfach, jeweils nach anderen Kriterien. In diesem Kapitel geht es nur um die Methodik, die konkreten Daten werden bei den Ergebnissen berichtet.

(1) Die **Erstellung der Durchführungshäufigkeit** der Selbsttests erfolgte über zwei unterschiedliche Wege. Der **erste Weg**, die **Erstellung der *gesamten* Durchführungshäufigkeit**, wurde mit Hilfe der SPSS-Funktion *Umstrukturierung* realisiert. Bei der Umstrukturierung eines Datenfiles (z.B. ST_M01_0304) wird eine neue Variable gebildet, welche die Anzahl der Fälle enthält, mit denen ein neuer Fall erstellt wird, d.h. diese neue Variable (z.B. n_01_0304) beinhaltet die Information, wie viele Fälle (Zeilen) des ursprünglichen Datenfiles zu einem Fall (einer Matrikelnummer) im neuen Datenfile zusammengefasst wurden. Somit werden alle Durchführungen einer Person pro Selbsttest gezählt, welche in dem ursprünglichen Datenfile als Zeilen abgespeichert waren, unabhängig davon wann sie gemacht wurden, d.h. auch jene nach z.B. dem ersten Prüfungsantritt im WS (i.e. der MZP 2 der Evaluation) bzw. dem ersten Prüfungsantritt im SS (i.e. der MZP 3). Diese "gesamte" Durchführungshäufigkeit wurde für die Beantwortung der Fragestellung 1a (vgl. Kapitel 2.3.4 – Beantwortung der Fragestellungen zu Ziel 1) herangezogen, nicht aber für die Beantwortung der nachfolgenden Fragestellungen. Hierfür wurde ab Fragestellung 1c die "genaue" ST-Durchführungshäufigkeit verwendet. Diese wurde über einen *zweiten Weg* erstellt, nämlich über die Bildung von Kategorien und deren Zusammenfassung (siehe unten im Unterkapitel (4) – Erstellung der *genauen* ST-Durchführungshäufigkeit).

(2) Die **Durchführungshäufigkeiten pro Kalenderwoche** (KW) des Studienjahrs 2003/04 wurden für die Beantwortung der Fragestellung 1b mittels SPSS-Funktion *Zählen* erstellt. Damit wurde eine neue Variable (z.B. KW_4003m01) erstellt, die (pro Zeile) die Information enthält, wie oft der jeweilige Selbsttest (z.B. ST 01) in einer bestimmten Kalenderwoche (z.B. KW 40 im Jahr 2003) durchgeführt worden ist. Da bei SPSS (Version 11.5) eine Kalenderwoche definiert ist von Mittwoch 00:00 Uhr bis Dienstag 23:59 Uhr wurden für das Zählen der Durchführungshäufigkeiten pro KW benutzerdefinierte *Datumszahlen* herangezogen, wobei eine KW definiert wurde von Montag 00:00 Uhr bis Sonntag 23:59 Uhr. Für jede Kalenderwoche (von KW 40 im Jahr 2003 bis KW 27 im Jahr 2005) wurde jeweils eine neue Variable erstellt, und dies für jeden der zehn Selbsttests. Somit wurden 93 Variablen mal zehn Selbsttests erstellt, deren Information für die Verlaufsdiagramme der Selbsttestsnutzung (siehe Kapitel 2.3.4 – Beantwortung der Fragestellungen zu Ziel 1) verwendet wurde.

(3) Weiters wurden die **Durchführungshäufigkeiten nach Kategorien des Durchführungszeitraums** (z.B. Selbsttest *vor der Prüfung* durchgeführt) für beide Studienjahre (2003/04 und 2004/05) erstellt. Hierfür wurde auch die SPSS-Funktion *Zählen* angewendet (siehe oben), jedoch mit anderen *Datumszahlen*. Tabelle 9 zeigt die erstellten Kategorien exemplarisch für das WS. Dieses Vorgehens wurde analog für das SS durchgeführt. Die neu erstellten Kategorien wurden dann im Weiteren für die verschiedenen Herangehensweisen der Erstellung von ST-Nutzungs-Typen unterschiedlich zusammengefasst (siehe unten im Unterkapitel (5) – Erstellung der Selbsttest-Nutzungs-Typen).

Tabelle 9: Übersicht der erstellten Kategorien des Durchführungszeitraumes für das WS

Bezeichnung der Kategorie	**Beschreibung der Kategorie des Durchführungszeitraumes**
im Modul	Diese Kategorie beinhaltet alle Durchführungen des jeweiligen Selbsttests, die vom Tag der Freischaltung des Moduls (00:00 Uhr) bis um 23:59 Uhr des Tages an dem das Modul endete (d.h. der Abgabetermin für die Task), durchgeführt worden sind.
nach dem Modul	Diese Kategorie beinhaltet alle Durchführungen des jeweiligen Selbsttests, die nach dem Ende des jeweiligen Moduls (00:00 Uhr) bis zum 15. Tag vor dem ersten Prüfungstermin (23:59 Uhr) durchgeführt worden sind.
vor dem 1. Prüfungstermin	Diese Kategorie beinhaltet alle Durchführungen des jeweiligen Selbsttests, die vom 14. Tag[17] (00:00 Uhr) vor dem 1. Prüfungstermin bis zum Tag des 1. Prüfungstermins (23:59 Uhr) durchgeführt worden sind.
zwischen 1. und 2. Prüfungstermin[18]	Diese Kategorie beinhaltet alle ST-Durchführungen, die nach dem Tag des 1. Prüfungstermins (00:00 Uhr) bis 15 Tage vor dem 2. Prüfungstermin (23:59 Uhr) durchgeführt worden sind.
vor dem 2. Prüfungstermin	Diese Kategorie beinhaltet alle Durchführungen des jeweiligen Selbsttests, die vom 14. Tag (00:00 Uhr) vor dem 2. Prüfungstermin bis zum Tag des 2. Prüfungstermins (23:59 Uhr) durchgeführt worden sind.
zwischen 2. und 3. Prüfungstermin	Diese Kategorie beinhaltet alle Durchführungen des jeweiligen Selbsttests, die nach dem Tag des 2. Prüfungstermins (00:00 Uhr) bis 15 Tage vor dem 3. Prüfungstermin (23:59 Uhr) durchgeführt worden sind.
vor dem 3. Prüfungstermin	Diese Kategorie beinhaltet alle Durchführungen des jeweiligen Selbsttests, die vom 14. Tag (00:00 Uhr) vor dem 3. Prüfungstermin bis zum Tag des 3. Prüfungstermins (23:59 Uhr) durchgeführt worden sind.

[17] Der definierte Anfang für den Zeitraum "vor der Prüfung" mit 14 Tage vor der Prüfung wurde aufgrund einer Befragung von ca. 20 Studierenden im Rahmen des Diplomandenseminars so veranschlagt, weil die Mehrheit der Studierenden angab, die Kategorie "vor der Prüfung" sei mit 14 Tage zu bemessen.

[18] Die Variablen "zwischen den Prüfungen" wurden nachträglich erstellt, um die ST-Durchführungen, welche in den Zeitraum nach einem Prüfungstermin bis zu 14 Tage vor dem nächsten Prüfungstermin fallen, nicht zu verlieren.

(4) Für die **Erstellung der *genauen* ST-Durchführungshäufigkeit** (i.e. der **zweite Weg** der **Erstellung der Durchführungshäufigkeit** – vgl. oben Unterkapitel (1)) wurde die SPSS-Funktion *Berechnen* in Kombination mit einer *IF-Funktion* angewendet. So wurden neue Variablen (z.B. ng_st_01) erstellt, welche jeweils alle ST-Durchführungshäufigkeiten beinhalten, welche eine Person bis zu ihrem ersten Prüfungs*antritt* (dies kann der 1., 2. oder 3. Prüfungs*termin* sein) des WS bzw. des SS durchgeführt hat, nicht aber jene ST-Durchführungen nach diesem ersten Antritt. Diese "genaue" ST-Durchführungshäufigkeit wurde erstellt, dass nicht ST-Durchführungen in die inferenzstatistische Auswertung mit eingehen, die nach der Datenerhebung für die Evaluation mittels SRL-Fragebogen (dies betrifft den 2. und 3. Messzeitpunkt) erfolgten. Beispielsweise gibt es Personen, die beim ersten Prüfungsantritt, bei welchem sie den SRL-Fragebogen ausgefüllt haben, eine negative Note erhalten haben und sich deshalb ein zweites Mal auf die Prüfung inkl. Durchführung der Selbsttests vorbereiteten. Diese ST-Durchführungen würden dann ebenso in die Auswertung mit einfließen, obwohl sie nach der Erhebung des Posttest stattgefunden haben. Eine Übersicht der Reduzierung der ST-Durchführungs-häufigkeit durch diese Vorgehensweise wird bei den deskriptiven Ergebnissen gegeben (siehe Kapitel 2.3.2 – Deskriptive Ergebnisse zu Ziel 1).

(5) Für die **Erstellung der Selbsttest-Nutzungs-Typen** wurden die sieben Kategorien des Durchführungszeitpunktes (siehe Tabelle 9) wie folgt zusammen-gefasst:

Im ersten Schritt wurde für jeden der zehn Selbsttests die Variable *vor der Prüfung – individuell* erstellt, indem die Variablen *vor dem 1. Prüfungstermin* bis *vor dem 3. Prüfungstermin* und die Variablen *zwischen den 1. und 2. Prüfungsterminen* und *zwischen den 2. und 3. Prüfungsterminen* je Selbsttest mittels einer *IF-Funktion* zusammenfasst wurden (siehe Abbildung 8 – untere Hälfte links). Diese neu erstellten Variablen beinhalteten alle ST-Durchführungen, welche ein/e Student/in vor jenem Prüfungstermin (WS bzw. SS) gemacht hat, zu welchem er/sie das erste Mal angetreten ist (wiederum wurde hierbei der zweite und dritte Prüfungs*antritt* ausgeklammert). Die neuen Variablen umfassten für einen Antritt beim ersten Prüfungstermin den Zeitraum von 14 Tagen und für einen Antritt beim zweiten bzw.

dritten Prüfungstermin umfassten diese den Zeitraum von ca. fünf bzw. 14 Wochen, wobei die Selbsttests meist erst wieder ab ca. vier Wochen vor der Prüfung durchgeführt worden sind (vgl. Abbildungen 10 und 11 in Kapitel 2.3.4 – Beantwortung der Fragstellungen zu Ziel 1).

Die im Anschluss durchgeführte Clusteranalyse zum Auffinden von Typen der Selbsttestnutzung lieferte leider kein zufrieden stellendes Ergebnis. Auch nicht nachdem in einem zweiten Schritt die Variablen *im Modul*, *nach dem Modul* und *vor der Prüfung – individuell* der einzelnen Selbsttests zu Variablen des gesamten Winter- bzw. des gesamten Sommersemesters zusammengefasst wurden (siehe Abbildung 8 – *im Modul ST 1-4*, *nach dem Modul ST 1-4* und *vor der Prüfung – individuell ST 1-4*). Wobei aber für das Wintersemester nur die Variablen der Selbsttests 1 bis 4 und für das Sommersemester die Variablen der Selbsttests 6 bis 9 herangezogen wurden, da es bei den Selbsttests 5 und 10 einerseits die Variable (sprich den Zeitraum) *nach dem Modul* nicht gibt (vgl. Tabelle 7 und 8 – Übersicht der Freischaltung) und andererseits eine klare Trennung zwischen der ST-Durchführung *im Modul* und *vor der Prüfung* nicht möglich ist (vgl. Abbildungen 10 und 11 in Kapitel 2.3.4 – Beantwortung der Fragestellungen zu Ziel 1). Abbildung 8 zeigt exemplarisch die Zusammenfassung der Variabeln der einzelnen Selbsttests zu den Variablen für das gesamte Wintersemester.

Auch ein weiterer Versuch die Clusteranalyse einzusetzen, nachdem in einem dritten Schritt die Variablen *im Modul* und *nach dem Modul* zusammenfasst worden waren zu den Variablen *im Semester* – wiederum zunächst für die einzelnen Selbsttests und dann für das gesamte Winter- bzw. Sommersemester – brachte keine zufriedenstellenden Ergebnisse.

Abbildung 8 veranschaulicht die Zusammenfassung der sieben Kategorien des ST-Durchführungszeitraums für das Wintersemester. Diese Zusammenfassung wurde analog für das Sommersemester durchgeführt.

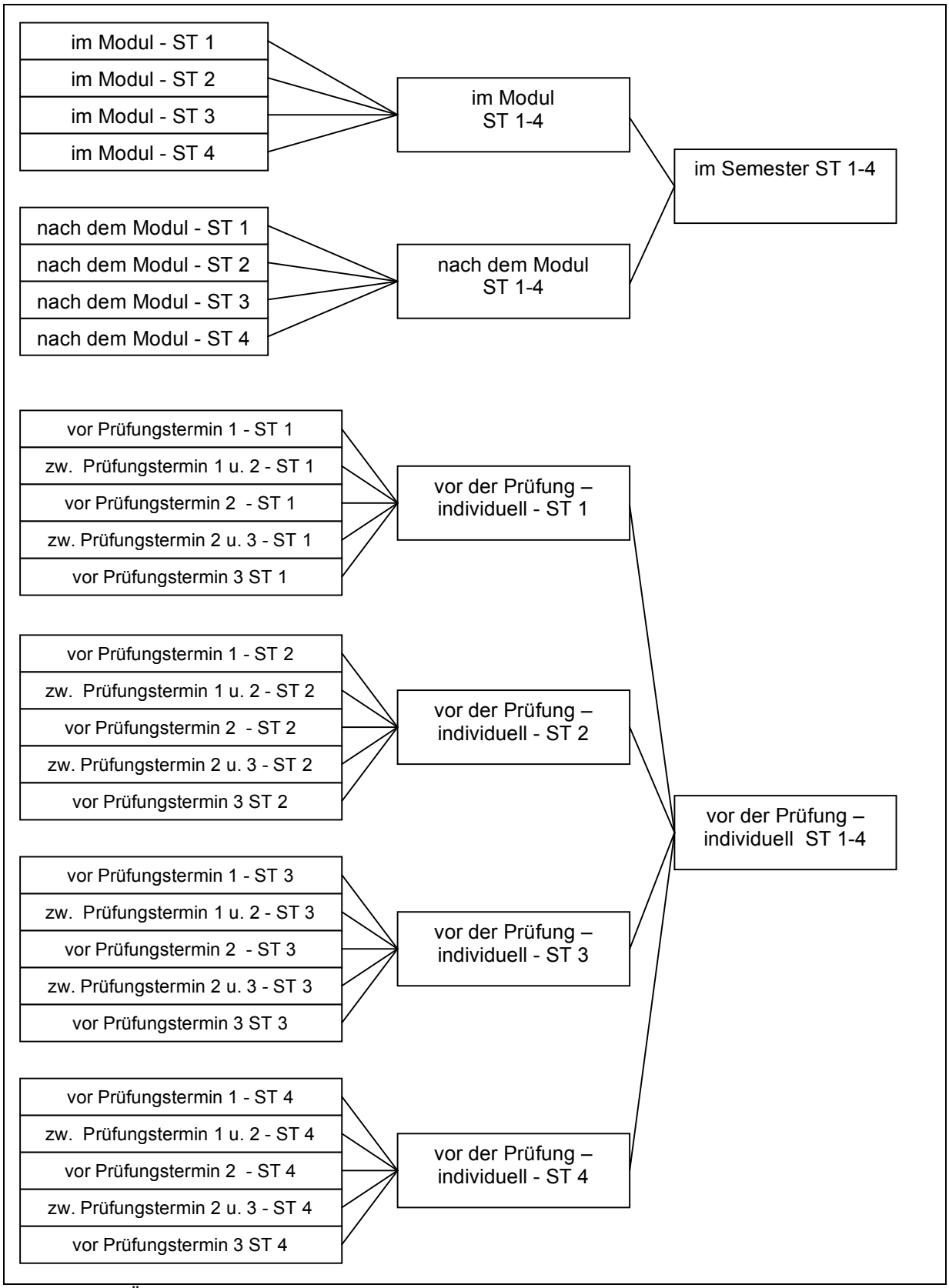


Abbildung 8: Übersicht der Zusammenfassung der sieben Kategorien des ST-Durchführungszeitraums exemplarisch für das WS

Da die durchgeführten Clusteranalysen keine zufriedenstellenden Ergebnisse bzgl. der Typen der Selbsttestnutzung lieferten, wurden mittels theoriegeleitetem Wissen (vgl. dazu Haag & Mischo, 2002; Wagner et al., 2004) Typen der Selbsttestnutzung erstellt, um somit die Fragestellungen bzgl. des Zusammenhangs von Selbsttest-nutzung und Prüfungsleistung beantworten zu können.

Hierfür wurden zunächst die *Durchführungshäufigkeiten* der Variablen *im Semester ST 1-4* (bzw. *im Semester ST 6-9*) und *vor der Prüfung ST 1-4* (bzw. *vor der Prüfung ST 6-9*) folgendermaßen kategorisiert:

- ST gar nicht durchgeführt (0 = 1)
- ST selten durchgeführt (1 bis 10 = 2)
- ST mittel oft durchgeführt (11 bis 20 = 3)
- ST oft durchgeführt (21 bis 30 = 4)
- ST sehr oft durchgeführt (31 bis höchster Wert = 5)

In der Folge, nach optischer Kontrolle der Häufigkeitsverteilungen, wurden die fünf Kategorien weiter zusammengefasst zu zwei Kategorien:

- ST gar nicht bis selten durchgeführt (0 bis 10 = 1)
- ST mittel bis sehr oft durchgeführt (11 bis höchster Wert = 2)

Aus diesen jeweils zwei 2-kategoriellen Variablen (für WS und SS) wurde je eine Variable für WS und eine für das SS mit jeweils vier “Ausprägungen“ (d.h. Kombinationen), den so genannten *ST-Nutzungs-Typen,* gebildet. Tabelle 10 veranschaulicht die Kombinationen der zwei 2-kategoriellen Variablen.

Tabelle 10: Vierfelderschema der Kombinationen der Variablen im Semester und vor der Prüfung inkl. der Benennungen der vier Typen der Selbsttestnutzung

	vor der Prüfung	
im Semester	*gar nicht bis selten*	*mittel bis sehr oft*
gar nicht bis selten	ST-Nicht-Nutzer/innen	Saison-nutzer/innen
mittel bis sehr oft	Mit-lerner/innen	ST-Viel-Nutzer/innen

Diese *Typen der Selbsttestnutzung* wurden, in Anlehnung an die Forschungsergebnisse zu Lerngewohnheiten von Lernenden (siehe Haag & Mischo, 2002; Wagner et al., 2004), wie folgt benannt (vgl. Tabelle 10):

- *ST-Nicht-Nutzer/innen* nutzten die Selbsttests im Semester gar nicht bis selten und auch vor der Prüfung nutzten sie die Selbsttests gar nicht bis selten.
- *Saison-Nutzer/innen* nutzten die Selbsttests im Semester gar nicht bis selten, jedoch nutzten sie die Selbsttests vor der Prüfung mittel bis sehr oft.
- *Mitlerner/innen* nutzten die Selbsttests im Semester mittel bis sehr oft, jedoch vor der Prüfung nutzten sie die Selbsttests gar nicht bis selten.
- *ST-Viel-Nutzer/innen* nutzten die Selbsttests im Semester mittel bis sehr oft und auch vor der Prüfung nutzten sie die Selbsttests mittel bis sehr oft.

(6) Weiters wurden **ST-Durchführungs-Häufigkeits-Typen** erstellt, welche in erster Linie für die inferenzstatistischen Auswertungen der Evaluation des attributionalen Feedbacks gedacht waren. Hierfür wurde die *genaue ST-Durchführungshäufigkeit* folgendermaßen kategorisiert (*umkodiert*):

- ST gar nicht bis sehr selten durchgeführt (0 bis 10 = 1)
- ST selten durchgeführt (11 bis 20 = 2)
- ST mittel oft durchführt (21 bis 30 = 3)
- ST oft bis sehr oft durchgeführt (31 bis höchster Wert = 4)

c) Zusammenfassung der Attributionen

Für die inferenzstatistische Auswertung der Evaluation des attributionalen Feedbacks mussten ebenso die Erfolgs- bzw. Misserfolgsattributionen zusammengefasst werden. Hierfür wurde pro Messzeitpunkt jeweils nur die erstgenannte Attribution herangezogen. Diese von den Studierenden zu den drei Messzeitpunkten (MZP) jeweils erstgenannten Attributionen zu Erfolg und Misserfolg wurden in Anlehnung an das Zweidimensionale Klassifikationsschema für die wahrgenommenen Ursachen von Leistungsergebnissen nach Weiner (1994; siehe Kapitel 1.3.3 – Attributionstrainings) in vier Kategorien zusammengefasst. Tabelle 11 zeigt die Zuordnung der ursprünglichen Kategorien zu den neu erstellten Kategorien.

Tabelle 11: Übersicht der ursprünglichen und neuen Kategorienzuordnung der Attributionen

Ursprüngliche Kategorie		Neue Kategorie	
1	Emotionale Zustände	3	internal-variabel
2	Externale Hilfestellung	4	external-variabel
3	Durchführungscharakteristika 1	4	external-variabel
4	Durchführungscharakteristika 2	3	internal-variabel
5	Anstrengung	3	internal-variabel
6	Konzentration	3	internal-variabel
7	Begabung/ Fähigkeit	1	internal-stabil
8	Zufall	4	external-variabel
9	Externale stabile Faktoren	2	external-stabil
10	Attributionsunsicherheit	5	Attributionsunsicherheit
11	Nicht zuordenbar	6	Nicht zuordenbar

Abschließend werden in Tabelle 12 nochmals alle neu erstellten Variablen, welche für die folgenden Kapitel von Bedeutung sind, inkl. deren Wertebereich und für welche Fragestellung/en diese herangezogen werden, wiedergegeben.

Tabelle 12: Übersicht der neu erstellten Variablen inkl. deren Wertebereich und für welche Fragestellung/en, diese herangezogen werden

Name der Variable	Wertebereich pro Person	für die Fragestellung/en
gesamte ST-Durchführungshäufigkeit	von 0 bis ca. 100 Durchführungen je Selbsttest	1a
ST-Durchführungshäufigkeit pro Kalenderwoche	von 0 bis ca. 100 Durchführungen je Selbsttest	1b
genaue ST-Durchführungshäufigkeit	von 0 bis ca. 100 Durchführungen je Selbsttest	1c bis 1e
Selbsttest-Nutzungs-Typen	1 ST-Nicht-Nutzer/innen 2 Saison-Nutzer/innen 3 Mitlerner/innen 4 ST-Viel-Nutzer/innen	1f bis 1h
ST-Durchführungs-Häufigkeits-Typen (auch *ST-Häufigkeits-Typen* genannt)	1 ST gar nicht bis sehr selten durchgeführt 2 ST selten durchgeführt 3 ST mittel oft durchführt 4 ST oft bis sehr oft durchgeführt	2a bis 2f und Zusatzfragestellung ad Ziel 1
Kategorisierte Erfolgs- bzw. Misserfolgsattributionen	1 internal-stabil 2 external-stabil 3 internal-variabel 4 external-variabel 5 Attributionsunsicherheit 6 Nicht zuordenbar	2a

2.2.2 Statistische Auswertungsverfahren

In der vorliegenden Arbeit wurde für die statistischen Auswertungen einerseits das *Statistic Programm for Social Sciences* (SPSS, Version 11.5) verwendet. Zum anderen kam zur Berechnung der Konfigurationsfrequenzanalyse (KFA) das Programm *Configural Frequency Analysis* (CFA) – a program for 32 bit Windows operating systems (Two-Sample CFA 2000 und First order CFA 2002 - letzte Revision März 2007) von Alexander von Eye (2002) zum Einsatz. Sämtliche Abbildungen wurden im Programm *Microsoft Excel (2003 SP 2* und *2004 for Mac; Version 11.0)* erstellt.

Die Fragestellungen 1a, 1b und 1f wurden mittels Deskriptivstatistik beantwortet, die Abbildungen dafür wurden in *Microsoft Excel* erstellt. Für die Fragestellungen bzgl. der ST-Durchführungshäufigkeit (1c bis 1e und 1h) wurden trotz Verletzung der Voraussetzungen parametrische Verfahren angewendet, da nach Rasch und Guiard (2004) in den meisten Fällen, trotz Voraussetzungsverletzungen, parametrische Verfahren gegenüber non-parametrischen Verfahren vorzuziehen sind. Parametrische Verfahren sind gegenüber Voraussetzungsverletzungen, wie z.B. der Verletzung der Normalverteilung, sehr robust (Rasch & Guiard, 2004). Zusätzlich aber wurden alle signifikanten Ergebnisse – wenn möglich – mit adäquaten non-parametrischen Verfahren abgesichert. Die Fragestellung 1g wurde mittels 2-Stichproben-KFA beantwortet.

Die Beantwortungen der Fragestellungen zum Ziel 2 erfolgten zum Teil mittels KFAs der 1. Ordnung und zum Teil anhand univariater Varianzanalysen mit Messwiederholungen.

Das α-Fehler-Niveau wurde auf 5% festgesetzt. Bei signifikanten Ergebnissen wurde zusätzlich auch der Schätzer der Effektgröße Eta-Quadrat (η^2) angegeben. Effektgrößen wurden nach der Klassifikation von Cohen (1988, 1992; zitiert nach Bortz & Döring, 2006) gedeutet.

2.3 Ergebnisse

Zwecks besserer Übersicht wird die Ergebnisdarstellung aufgeteilt in zwei Blöcke. Zuerst werden zum Ziel 1 der Diplomarbeit, die Analyse der Selbsttestnutzung, die deskriptiven Ergebnisse wiedergegeben und die Fragestellungen beantwortet. Anschließend werden zum Ziel 2, die Evaluation des web-basierten attributionalen Feedbacks, die deskriptiven Ergebnisse dargestellt und die Fragestellungen beantwortet. Vorweg jedoch erfolgt die Stichprobenbeschreibung.

2.3.1 Stichprobenbeschreibung

Insgesamt wurden Selbsttestdaten von **387 Studierenden** (337 Frauen (87.1%) und 50 Männer (12.9%)) aus zwei Studienjahren (229 Personen (59.2%) im Studienjahr 2003/04 und 158 Personen (40.8%) im Studienjahr 2004/05) bearbeitet und komprimiert (= **Gesamtstichprobe**). Davon absolvierten 85 Personen (22.0%) die "Intensiv-Variante" (Track A) und 302 Personen (78.0%) die "Standard-Variante" (Track B).

Wie zuvor schon berichtet (vgl. Kapitel 2.1.2 – Stichproben) wurden für die vorliegende Untersuchung zwei Stichproben herangezogen, welche nun folgend kurz mittels Deskriptivstatistik beschrieben werden.

Die **Stichprobe 1** besteht aus den Studierenden des **Studienjahres 2003/04**, die an den beiden Lehrveranstaltungen "Forschungsmethoden und Evaluation I und II" teilgenommen haben, bei welchen die Matrikelnummer und der Code zusammengeführt werden konnten und die bei allen drei Messzeitpunkten der Evaluation teilgenommen haben (N = 229). Aus dieser Stichprobe wurden insgesamt drei Studierende ausgeschlossen (wegen Track-Wechsel bzw. unrealistischer Daten, d.h. die Selbsttests mehr als 300-mal durchgeführt).

Die Stichprobe 1 umfasst somit **226 Studierende**. Die Studierenden waren im Durchschnitt 24.55 Jahre alt *(SD* = 3.09 Jahre). Das häufigste Alter war 23. Der/die jüngste Teilnehmer/in war 21 und der/die älteste 42 Jahre alt. Der Kolmogorov-Smirnov-Test (K-S-Test) und die optische Prüfung zeigten, dass das Alter der

Stichprobe nicht normalverteilt (*K-S-Z* = 3.19, *p* < .001), sondern linkssteil und rechtsschief ist.

Die Stichprobe 1 besteht aus 197 Frauen (87.2%) und 29 Männern (12.8%). 35 Studierende (15.5%) der Stichprobe 1 wählten die "Intensiv-Variante" (Track A) und 191 (84.5%) absolvierten die "Standard-Variante" (Track B). Tabelle 13 zeigt eine Übersicht, zu welchem Prüfungstermin die Studierenden der Stichprobe 1 angetreten sind. Es wurde jeweils nur der erste Antritt gezählt, auch wenn ein/e Studierende/r mehrere Antritte hatte.

Tabelle 13: Anzahl der angetretenen Personen (nur erster Antritt) pro Prüfungstermin

	Anzahl der Personen	
Erster Prüfungsantritt	*WS*	*SS*
Beim 1. Prüfungstermin angetreten	93 (41.2%)	124 (54.9%)
Beim 2. Prüfungstermin angetreten	89 (39.4%)	76 (33.6%)
Beim 3. Prüfungstermin angetreten	44 (19.5%)	26 (11.5%)
Gesamt	226 (100.0%)	226 (100.0%)

Die **Stichprobe 2** besteht aus **168 Studierenden**. Die Versuchsgruppe umfasst 84 Studierenden (50.0%), die an der "Intensiv-Variante" (Track A) im Studienjahr 2003/04 bzw. 2004/05 teilgenommen haben. Die Kontrollgruppe (n = 84; 50.0%) wurde aus dem Pool der Studierenden der „Standard-Variante" (Track B) des Studienjahres 2003/04 mittels Matching-Verfahrens gezogen (siehe Kapitel 2.1.1 – Untersuchungsdesign und Kapitel 2.1.2 – Stichproben). Die Studierenden der Stichprobe 2 waren im Durchschnitt 25.01 Jahre alt (*SD* = 3.98 Jahre). Das häufigste Alter war 22 Jahre. Der/die jüngste Teilnehmer/in war 21 und der/die älteste 50 Jahre alt. Der K-S-Test und die optische Prüfung zeigten, dass auch das Alter der Stichprobe 2 nicht normalverteilt (*K-S-Z* = 2.79, *p* < .001), sondern linkssteil und rechtsschief ist. Nach optischer Prüfung der Verteilung der ST-Durchführungshäufigkeit wurde in der Stichprobe 2 keine Person ausgeschieden.

Die Stichprobe 2 (**Matching-Stichprobe**) besteht aus 138 Frauen (82.1%) und 30 Männern (17.9%), wobei die Geschlechterverteilung in Versuchs- und Kontrollgruppe gleich ist, d.h. je Track 69 Frauen (82.1%) und 15 Männer (17.9%). Tabelle 14 zeigt,

zu welchem Prüfungstermin die Studierenden der Stichprobe 2 angetreten sind. Es wurde wieder jeweils nur der erste Antritt gezählt, auch wenn ein/e Studierende/r mehrmals angetreten war. Die Stichprobe 2 beinhaltet auch eine/n Studierende/n (0.6%), der/die bei keinem der ersten drei Prüfungstermine im WS angetreten ist und vier Studierende (2.4%), die bei keinem der ersten drei Prüfungstermine im SS angetreten sind.

Tabelle 14: Anzahl der angetretenen Personen (nur erster Antritt) pro Prüfungstermin

	Anzahl der Personen	
Erster Prüfungsantritt	*WS*	*SS*
Beim 1. Prüfungstermin angetreten	97 (57.7%)	106 (63.1%)
Beim 2. Prüfungstermin angetreten	43 (25.6%)	43 (25.6%)
Beim 3. Prüfungstermin angetreten	27 (16.1%)	15 (8.9%)
Bei keinem der ersten drei Prüfungstermine angetreten	1 (0.6%)	4 (2.4%)
Gesamt	168 (100.0%)	168 (100.0%)

2.3.2 Deskriptive Ergebnisse zu Ziel 1

Als Ziel 1 der Diplomarbeit wurde die Analyse der Selbsttestnutzung definiert. Die Selbsttestnutzung wurde operationalisiert einerseits durch die ST-Durchführungshäufigkeit (Wie oft haben die Studierenden die Selbsttests abgesendet?) und andererseits durch die Kombination der ST-Durchführungshäufigkeit und dem ST-Durchführungszeitraum (Haben die Studierenden die Selbsttests im Semester und/oder vor der Prüfung gar nicht bis sehr oft durchgeführt?).

Für die Erstellung der ST-Durchführungshäufigkeit gab es zwei unterschiedliche Vorgehensweisen (“Wege“ – vgl. Kapitel 2.2.1 – Aufbereitung und Komprimierung der Daten). Im Folgenden werden nun diese zwei ST-Durchführungshäufigkeiten, die “gesamte“ und die “genaue“ ST-Durchführungshäufigkeit gegenübergestellt.

Gegenüberstellung der gesamten ST-Durchführungshäufigkeiten und der genauen ST-Durchführungshäufigkeiten

Zur Beantwortung der Fragestellung 1a mittels Deskriptivstatistik wird, wie in Kapitel 2.2.1 – Aufbereitung und Komprimierung der Daten – schon erwähnt, die *gesamte* ST-Durchführungshäufigkeit verwendet (d.h. alle ST-Durchführungen, auch jene nach dem ersten Prüfungsantritt). Für die Beantwortung aller weiteren Fragestellungen (ab Fragestellung 1c) wird jedoch eine genau berechnete ST-Durchführungshäufigkeit, die so genannte *genaue* Durchführungshäufigkeit, herangezogen, welche nur die ST-Durchführungen enthält, welche die Studierenden bis vor ihrem jeweilig ersten Prüfungs*antritt* im WS bzw. im SS durchgeführt haben. Die exakte Vorgehensweise der Erstellung der beiden ST-Durchführungshäufigkeiten ist im Kapitel 2.2.1 – Aufbereitung und Komprimierung der Daten – beschrieben.

Tabelle 15: Übersicht über die gesamte, genaue ST-Durchführungshäufigkeiten und der Differenz in absoluten Zahlen und in Prozent der zehn Selbsttests und gesamt

	gesamte **ST-Durchführungs-häufigkeit**	***genaue*** **ST-Durchführungs-häufigkeit**	**Differenz**
ST 1	2950	2824	126 (4.27%)
ST 2	3174	3026	148 (4.66%)
ST 3	2570	2452	118 (4.59%)
ST 4	2816	2682	134 (4.76%)
ST 5	2992	2765	227 (7.59%)
ST 6	2301	2145	156 (6.78%)
ST 7	3075	2867	208 (6.76%)
ST 8	2148	1989	159 (7.40%)
ST 9	2129	1990	139 (6.53%)
ST 10	2309	2112	197 (8.53%)
ST 1-5	14502	13749	753 (5.19%)
ST 6-10	11962	11103	859 (7.18%)
Gesamt	**26464**	**24 852**	**1612** (**6.09**%)

Tabelle 15 zeigt die *gesamte* und *genaue* ST-Durchführungshäufigkeiten der Gesamtstichprobe (*N* = 387), sowie die Differenzen in absoluten Zahlen und als Prozentangaben für alle 10 Selbsttest und gesamt. Die Selbsttests 1-5 wurden in den Studienjahren 2003/04 und 2004/05 insgesamt 14 502-mal durchgeführt, hingegen erfolgte die ST-Durchführung vor dem ersten Prüfungsantritt der Studierenden 13 749-mal (94.81%). Dies ergibt eine Differenz von 753 (5.19%) ST-Durchführungen. Die Selbsttests 6-10 wurden in den Studienjahren 2003/04 und 2004/05 insgesamt 11 962-mal durchgeführt, die ST-Durchführung erfolgte 11 103-mal (92.82%) vor dem ersten Prüfungsantritt. Die Differenz beträgt 859 (7.18%) ST-Durchführungen.

Die Verteilung der *genauen* Durchführungshäufigkeiten eines jeden einzelnen der zehn Selbsttests wurde mittels K-S-Test auf Normalverteilung geprüft. Es zeigte sich, dass die *genauen* Durchführungshäufigkeiten aller zehn Selbsttests der Gesamtstichprobe nicht normalverteilt (*K-S-Z* = 3.03 – 6.01, $p < .001$) sind. Ebenso sind die *genauen* ST-Durchführungshäufigkeiten zusammengefasst für das WS (*K-S-Z* = 3.17, $p < .001$) und für das SS (*K-S-Z* = 4.42, $p < .001$) nicht normalverteilt. Die optische Prüfung der Verteilungen zeigte, dass alle einzelnen Verteilungen der ST-Durchführungshäufigkeiten linkssteil und rechtsschief sind.

2.3.3 Prüfung der ST-Durchführungshäufigkeits-Gleichheit

Die Matching-Stichprobe (Stichprobe 2), beinhaltet Studierende zweier Studienjahre (vgl. Kapitel 2.1.2 – Stichproben). Um abzusichern, dass diese Stichprobe für die statistische Auswertung herangezogen werden kann, wurde im Vorfeld der eigentlichen Fragestellungen anhand einer Gesamtstichprobe (*N* = 387 – siehe Kapitel 2.1.2 – Stichproben) überprüft, dass es keine Unterschiede in der ST-Durchführungshäufigkeit zwischen den beiden Kohorten (Studienjahr 2003/04 und 2004/05) gibt. Hierfür wurde die *genaue* Durchführungshäufigkeit der Selbsttests (siehe Kapitel 2.2.1 – Unterkapitel (4) Erstellung der genauen ST-Durchführungshäufigkeit) herangezogen, da diese *genaue* Durchführungshäufigkeit auch bei der anschließenden inferenzstatistischen Auswertung der Fragestellungen verwendet wird. Im Folgenden werden nun die wichtigsten Ergebnisse wiedergegeben:

Gibt es Unterschiede in der ST-Durchführungshäufigkeit[19] zwischen den Studierenden des Studienjahres 2003/04 und 2004/05?

Es bestehen keine signifikanten Unterschiede zwischen der ST-Durchführungshäufigkeit im Studienjahr 2003/04 und der ST-Durchführungshäufigkeit im Studienjahr 2004/05 ($t(385) = -0.98$, $p = .328$). Ebenso gibt es keine signifikanten Unterschiede zwischen den beiden Studienjahren hinsichtlich der ST-Durchführungshäufigkeit im WS ($t(385) = -1.19$, $p = .234$), der ST-Durchführungshäufigkeit im SS ($t(268.84) = -0.45$, $p = .651$) und der ST-Durchführungshäufigkeit in den zehn einzelnen Selbsttests ($t(267.93 – 385) = (-1.20) – (0.48)$, $p = .229 – .864$). Tabelle 16 zeigt die zugehörigen Mittelwerte und Standardabweichungen.

Tabelle 16: Mittelwerte (M) und Standardabweichungen (SD) der Durchführungshäufigkeiten der zehn Selbsttests, des WS (ST 1-5), des SS (ST 6-10) und des gesamten Studienjahres (ST 1-10) jeweils für 2003/04 und 2004/05 und die Gruppengröße (n).

	Studienjahr 2003/04		**Studienjahr 2004/05**	
	n = 229		*n* = 158	
Selbsttest	*M*	*SD*	*M*	*SD*
1	7.10	7.30	7.59	6.15
2	7.48	9.76	8.30	10.14
3	5.98	7.26	6.85	7.59
4	6.25	8.47	7.92	18.33
5	6.91	10.45	7.48	9.22
6	5.69	7.88	5.32	7.08
7	7.11	10.92	7.84	14.82
8	5.20	7.42	5.05	9.77
9	4.89	6.77	5.51	8.63
10	5.03	9.19	6.08	12.64
WS	33.72	35.81	38.14	35.90
SS	27.92	32.62	29.80	44.65
Gesamt	61.65	59.29	67.94	66.04

[19] Zwecks Vereinfachung und besseren Verständlichkeit wird ab nun die *genaue* Durchführungshäufigkeit der Selbsttests nur noch als *ST-Durchführungshäufigkeit* bezeichnet. Die *gesamte* Durchführungshäufigkeit wird weiterhin extra deklariert.

Gibt es Unterschiede in der ST-Durchführungshäufigkeit zwischen Studienjahr 2003/04 und 2004/05 bei den Studierenden des Track A bzw. bei den Studierenden des Track B?

Es gibt auch keine signifikanten Unterschiede innerhalb des Track A zwischen der ST-Durchführungshäufigkeit im Studienjahr 2003/04 und der ST-Durchführungshäufigkeit im Studienjahr 2004/05 ($t(83) = -0.19$, $p = .854$). Ebenso bestehen keine signifikanten Unterschiede innerhalb des Track B zwischen der ST-Durchführungshäufigkeit im Studienjahr 2003/04 und der ST-Durchführungshäufigkeit im Studienjahr 2004/05 ($t(300) = -1.08$, $p = .283$).
Tabelle 17 zeigt die Mittelwerte und Standarbabweichungen der ST-Durchführungshäufigkeiten der beiden Kohorten (2003/04 und 2004/05) getrennt nach Trackzugehörigkeit.

Tabelle 17: Mittelwerte (M) und Standardabweichungen (SD) der ST-Durchführungshäufigkeiten Studienjahre (ST 1-10) 2003/04 und 2004/05 getrennt nach Track A und Track B inkl. der zugehörigen Gruppengröße (n)

	Track A			**Track B**		
Kohorte	*n*	*M*	*SD*	*n*	*M*	*SD*
Studienjahr 2003/04	35	61.29	46.49	194	61.71	61.42
Studienjahr 2004/05	50	63.52	60.07	108	69.99	68.81

Da es weder signifikante Unterschiede zwischen den beiden Studienjahren (2003/04 und 2004/05), noch zwischen den beiden Studienjahren innerhalb der beiden Tracks in der *genauen* Durchführungshäufigkeit der Selbsttests gibt, kann für die inferenzstatistische Auswertung der Fragestelllungen die Matching-Stichprobe herangezogen werden. Denn obwohl diese Stichprobe Studierende zweier unterschiedlicher Studienjahre beinhaltet, unterscheiden sich diese nicht in der Durchführungshäufigkeit der Selbsttests. So beruhen die Ergebnisse der Beantwortung der Fragestellungen auf den Daten der Matching-Stichprobe ($N = 168$) und den *genauen* Durchführungshäufigkeiten der Selbsttests, mit Ausnahme der Fragestellungen 1a und 1b (vgl. Kapitel 2.2.1 – Aufbereitung und Komprimierung der Daten), welche im nächsten Kapitel erläutert werden.

2.3.4 Beantwortung der Fragestellungen zu Ziel 1

Im Folgenden werden die Fragestellungen des Zieles 1 der Diplomarbeit, i.e. die Analyse der Selbsttestnutzung und die Überprüfung eines Zusammenhangs der Selbsttestnutzung mit den Ergebnissen in den Lehrveranstaltungsprüfungen zu FM&E, beantwortet. Die beiden ersten Fragestellungen (1a und 1b) werden mittels Deskriptivstatistik beantwortet. Die Ergebnisse beruhen auf den Daten der Stichprobe 1 (siehe Kapitel 2.3.1 – Stichprobenbeschreibung). Da hierbei ein Überblick gegeben werden soll, wie oft und wann die Selbsttests genutzt worden sind, wurden für die so genannte *gesamte* Durchführungshäufigkeit der Selbsttests alle Selbsttestdurchführungen des Studienjahres 2003/04 (von der KW 40 im Jahr 2003 bis zur KW 27 im Jahr 2005) herangezogen, d.h. auch jene ST-Durchführungen, die ein/e Studierende/r vor seinem/ihrem zweiten bzw. dritten Prüfungsantritt durchgeführt hat. Für die Beantwortung aller weiteren Fragestellungen wurden jedoch wieder die *genauen* Durchführungshäufigkeiten der Selbsttests (vgl. Kapitel 2.2.1 – Aufbereitung und Komprimierung der Daten) herangezogen.

1a) Wie oft wurden die Selbsttests von den Studierenden durchgeführt?

Die zehn Selbsttests wurden von den 226 Studierenden des Studienjahres 2003/04 (Stichprobe 1) insgesamt 14 806-mal durchgeführt. Davon wurden die Selbsttests des WS (ST 1-5) 8 142-mal (55.0%) und die Selbsttests des SS (ST 6-10) 6 664-mal (45.0%) absolviert.

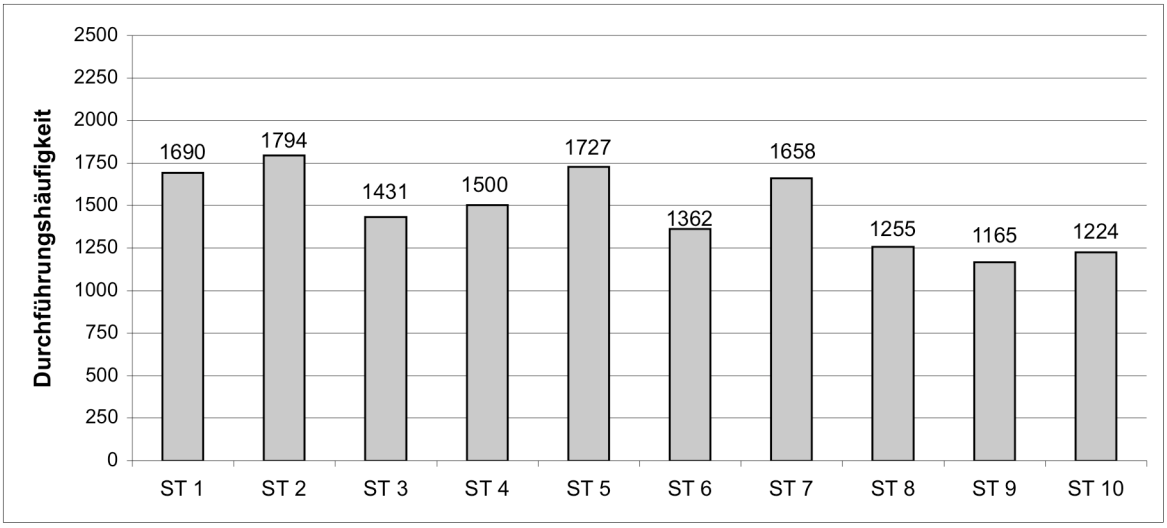


Abbildung 9: Durchführungshäufigkeiten der zehn Selbsttests

Abbildung 9 gibt einen Überblick, wie oft jeder der zehn Selbsttests von den 226 Studierenden der Stichprobe 1 im Studienjahr 2003/04 durchgeführt wurde. Wobei hier nochmals darauf hingewiesen wird, dass die Selbsttests eine unterschiedliche Anzahl von Multiple-Choice-Aufgaben haben (siehe Kapitel 2.1.3 – Erhebungsinstrumente). Tabelle 18 zeigt Mittelwerte, Standardabweichungen, Mediane, Modalwerte, Minima und Maxima der Durchführungshäufigkeiten der zehn Selbsttests, des Wintersemesters 2003/04 (ST 1-5), des Sommersemesters 2004 (ST 6-10) und des gesamten Studienjahres 2003/04 (ST 1-10).

Tabelle 18: Mittelwerte (M), Standardabweichungen (SD), Mediane (50. Perzentile), 25. und 75. Perzentile, Modalwerte (Mod), Minima (Min) und Maxima (Max) der Durchführungshäufigkeiten der zehn Selbsttests, des WS 2003/04 (ST 1-5), des SS 2004 (ST 6-10) und des gesamten Studienjahres 2003/04 (ST 1-10)

			Perzentile					
Selbsttest	*M*	*SD*	*25.*	*50.*	*75.*	*Mod*	*Min*	*Max*
1	7.48	7.34	3	5	10	2+3	0	49
2	7.94	9.72	2	5	10	2	0	59
3	6.33	7.35	2	4	8	1	0	42
4	6.64	8.95	1	3	8	2	0	62
5	7.64	11.01	1	4	9	1	0	78
6	6.03	7.77	1	4	8	0	0	72
7	7.34	10.51	1	4	8	0	0	64
8	5.55	7.76	1	3	6	0	0	57
9	5.15	6.87	1	3	7	0	0	39
10	5.42	9.77	1	3	6	0	0	99
WS 03/04	36.03	35.78	11.75	24.50	48.25	11	0	232
SS 2004	29.49	33.48	7.75	18.50	38	0	0	217
Gesamt 2003/04	65.51	59.11	26	47.50	82	2	1	296

Die Überprüfung auf Normalverteilung mittels K-S-Test zeigte, dass die *gesamten* Durchführungshäufigkeiten aller zehn Selbsttests der Stichprobe 1 nicht normalverteilt ($K\text{-}S\text{-}Z$ = 2.43 – 4.36, $p < .001$) sind. Die optische Prüfung der Verteilungen zeigte, dass alle zehn Verteilungen der ST-Durchführungshäufigkeit linkssteil und rechtsschief sind.

1b) Wann wurden die Selbsttests von den Studierenden durchgeführt?

Im Folgenden wird nun der Verlauf der Selbsttestnutzung nach Kalenderwochen (d.h. ST-Durchführungshäufigkeit pro Kalenderwoche) veranschaulicht. Es sei darauf hingewiesen, dass in den folgenden beiden Abbildungen die *Häufigkeiten* der ST-Durchführung wiedergegeben sind, wodurch die Kurven durch einzelne Ausreißer (d.h. durch Personen, welche die Selbsttests sehr oft durchgeführt haben) verzerrt sein könnten. Da es sich hierbei aber um eine deskriptive Veranschaulichung des Verlaufs handelt wurde die Gefahr der Verzerrung als vernachlässigbar eingeschätzt. Abbildung 10 zeigt die Durchführungshäufigkeit pro Kalenderwoche für die fünf Selbsttests des WS 2003/04 von Kalenderwoche (KW) 42 im Jahr 2003 bis KW 27 im Jahr 2004.

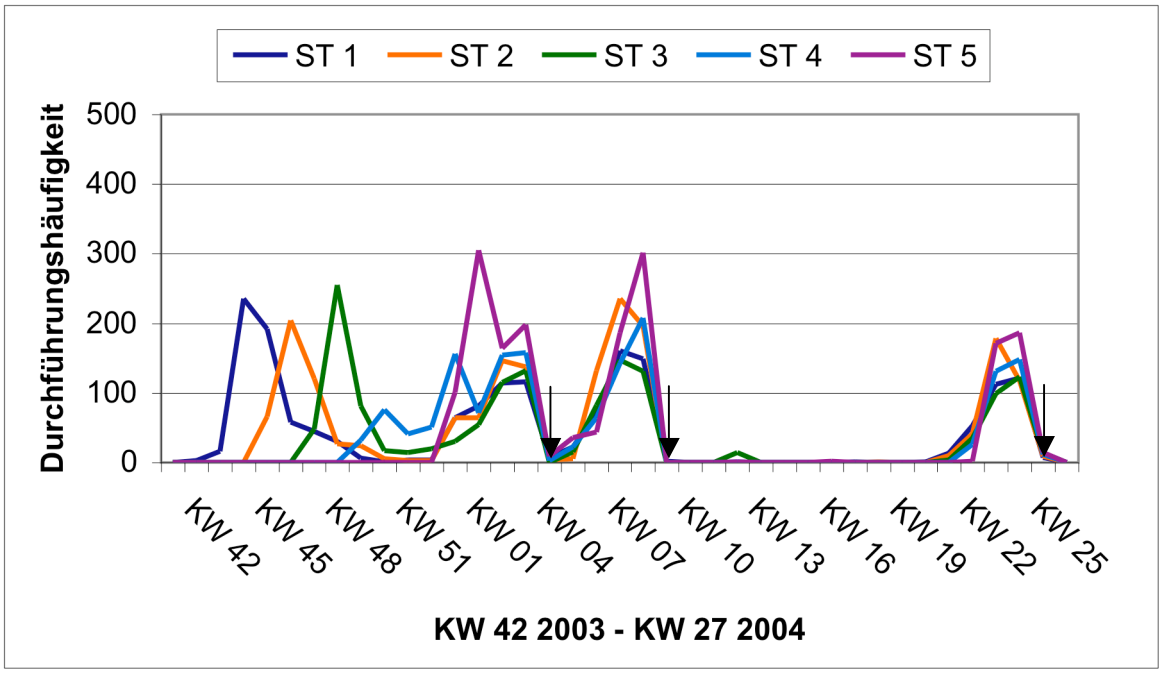


Abbildung 10: Durchführungshäufigkeiten der 5 Selbsttests im WS 2003/04 nach Kalenderwochen (KW); *Anmerkung:* Die Pfeile deuten die Zeitpunkte der Prüfungstermine an.

Wie Abbildung 10 veranschaulicht, wurde der Selbsttest 1 (ST 1) am häufigsten in der Kalenderwoche (KW) 45 (235-mal) und KW 46 (192-mal), d.h. während des Moduls 1, durchgeführt, danach flacht die Kurve der Selbsttestnutzung des ST 1 relativ rasch ab und steigt dann ca. vier Wochen vor dem jeweiligen Prüfungstermin wieder an bis auf ca. 110 bis 160 Durchführungen in den zwei letzten Wochen vor dem jeweiligen Prüfungstermin. Der Selbsttest 2 (ST 2) wurde in der KW 47, das ist während des Moduls 2, 204-mal durchgeführt. Insgesamt wurde der ST 2 von Studierenden während des Moduls (KW 45-47) 270-mal und nach Modul-Ende bis

zur KW 53 im Jahr 2003 182-mal durchgeführt. Auch der ST 2 wurde vor den Prüfungsterminen wieder öfters durchgeführt, wobei er drei Wochen vor dem zweiten Prüfungstermin (KW 7-9) am häufigsten durchgeführt wurde (insgesamt 563-mal). Beim Selbsttest 3 (ST 3) war der Höhepunkt der Selbsttest-Nutzung im Modul (KW 49) mit 255 Durchführungen. Die Durchführungshäufigkeit nahm aber rasch wieder ab auf ca. 20 Durchführungen pro KW nach dem Modul. Vor den Prüfungsterminen wurde der ST 3 jeweils ca. 130- bis 150-mal in den letzten beiden Wochen vor dem jeweiligen Prüfungstermin durchgeführt. Die Kurven der Selbsttests 1-3 zeigen, dass diese Selbsttests hauptsächlich einerseits während des Moduls und andererseits vor der Prüfung durchgeführt wurden. Bei Selbsttest 4 und 5 zeigt sich jedoch ein anderer Verlauf. Selbsttest 4 (ST 4) wurde von den Selbsttests des WS am seltensten während des Moduls durchgeführt (324-mal von KW 49 bis KW 1). Dies liegt höchstwahrscheinlich an der Tatsache, dass dieses Modul inkl. ST in den Weihnachtsferien zur Bearbeitung zur Verfügung stand. Bei Selbsttest 5 kommt es zu einer Art “Verschmelzung“ der Durchführung im Modul und vor dem ersten Prüfungstermin. Der ST 5 erreichte den Höhepunkt seiner Durchführung in der KW 2 (305-mal), diese KW fällt noch in den Bearbeitungszeitraum von Modul 5, liegt aber auch zwei Wochen vor dem ersten Prüfungstermin, sodass hier eine Unterscheidung von Bearbeitung im Modul bzw. vor der Prüfung nicht mehr vorgenommen werden kann. Insgesamt wurden die fünf Selbsttests des Wintersemesters am häufigsten (2 424-mal) vor dem ersten Prüfungstermin (KW 1-4) und am seltensten (1 571-mal) vor dem dritten Prüfungstermin (KW 22-25) durchgeführt. Vor dem zweiten Prüfungstermin (KW 6-9) wurden die fünf Selbsttests insgesamt 2 348-mal durchgeführt.

Abbildung 11 zeigt den Verlauf der Selbsttestnutzung nach Kalenderwochen für die fünf Selbsttests des SS 2004 von KW 10 im Jahr 2004 bis KW 05 im Jahr 2005. Wie zu erkennen ist, wurden die Selbsttests des Sommersemesters 2004 (ST 6-10) im Vergleich zu den Selbsttests im Wintersemester 2003/04 während des jeweiligen Moduls seltener durchgeführt (von ca. 65-mal den ST 8 bis 180-mal den ST 9). Weiters wurden die fünf Selbsttests des Sommersemesters, insgesamt 2 940-mal vor dem ersten Prüfungstermin des Sommersemesters (KW 23-26) und insgesamt 2 042-mal vor dem zweiten Prüfungstermin (KW 38-41) durchgeführt. Die fünf

Selbsttests 6-10 wurden vor dem dritten Prüfungstermin (KW 1-4 im Jahr 2005) insgesamt am seltensten (594-mal) durchgeführt.

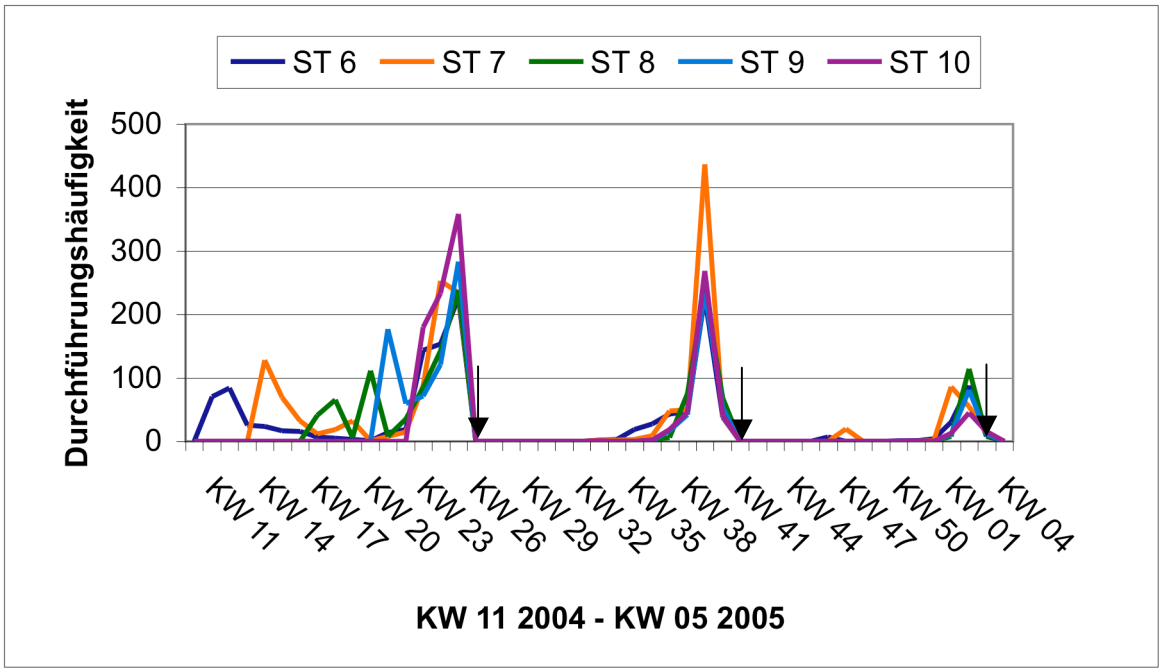


Abbildung 11: Durchführungshäufigkeiten der 5 Selbsttest im SS 2004 nach Kalenderwochen (KW); *Anmerkung:* Die Peile deuten die Zeitpunkte der Prüfungstermine an.

Die Abbildung 12 zeigt die Mittelwerte der Durchführungshäufigkeit aller Selbsttests pro Kalenderwoche getrennt nach Track A (n = 35) und Track B (n = 191). Da in der Stichprobe 1 unterschiedlich viele Personen in Track A und Track B sind, wurden für die Darstellung des Verlaufs Mittelwerte gewählt, sodass die ST-Durchführungshäufigkeit pro Kalenderwoche zwischen Track A und B vergleichbar wird, obwohl es auch bei Mittelwerten durch Ausreißer zu Verzerrungen kommen kann. Diese wurden aber in Kauf genommen, da die Abbildung wiederum zur Veranschaulichung des Verlaufs dient und dieser Verlauf durch Mediane oder Modalwerte nicht hinreichend wiedergegeben werden kann. Wie in Abbildung 12 zu sehen ist, wurden zu Beginn des Wintersemesters 2003/04 (KW 43-51) die Selbsttests im Durchschnitt von den Studierenden des Track A häufiger durchgeführt als von den Studierenden des Track B. Weiters ist zu erkennen, dass im Sommersemester 2004 die Studierenden des Track B die Selbsttests während der Module 6-8 (i.e. KW 11-19) im Durchschnitt relativ selten durchführen (M = 0.13 – 0.55; SD = 0.00 – 2.87), während die Studierenden des Track A mindestens in einer Kalenderwoche während des jeweiligen Moduls die Selbsttests im Durchschnitt häufiger als einmal absolviert haben (KW 12: M = 1.31, SD = 4.64; KW 15: M = 1.29, SD = 3.16; KW 19: M = 1.37, SD = 5.09).

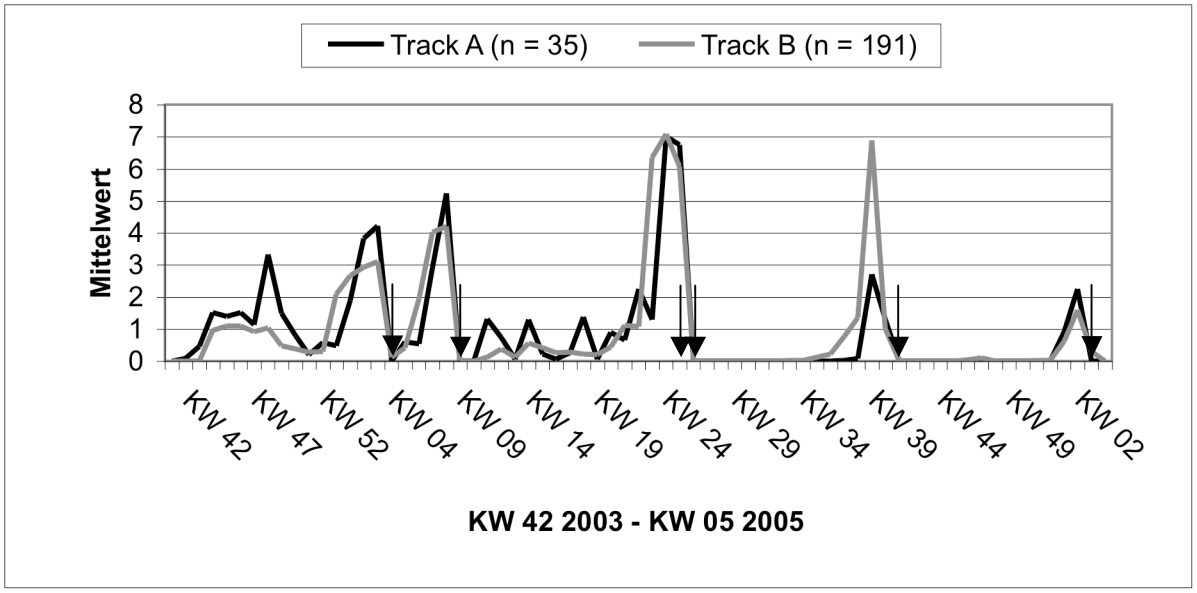


Abbildung 12: Verlaufskurve der Mittelwerte der Durchführungshäufigkeit gesamt für alle zehn ST je Kalenderwoche (KW) getrennt für Track A und Track B für das Studienjahr 2003/04; *Anmerkung:* Die Pfeile deuten die Zeitpunkte der Prüfungstermine an.

Sämtliche folgenden Fragestellungen der Diplomarbeit werden nun, wie zuvor schon beschrieben, anhand der Daten der Matching-Stichprobe (Stichprobe 2) beantwortet und es wird durchgehend die *genaue* Durchführungshäufigkeit der Selbsttests verwendet. Auch die *genauen* ST-Durchführungshäufigkeiten der Matching-Stichprobe (Stichprobe 2) sind wiederum durchgängig nicht normalverteilt ($K\text{-}S\text{-}Z$ = 2.23 – 3.47, p < .001), sondern die Verteilungen sind linkssteil und rechtschief.

1c) Gibt es Unterschiede in der Durchführungshäufigkeit der zehn Selbsttests?

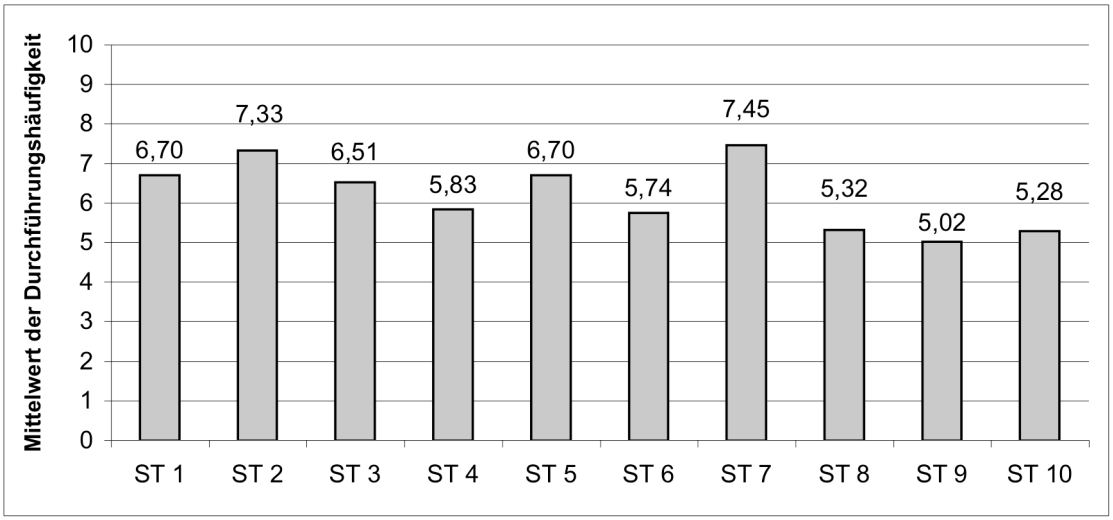


Abbildung 13: Mittelwerte der Durchführungshäufigkeiten der zehn Selbsttests

In Abbildung 13 sind die Mittelwerte der Durchführungshäufigkeiten der zehn Selbsttests dargestellt. Die Überprüfung mittels *Varianzanalyse für abhängige Gruppen*[20] ob es Unterschiede in den Durchführungshäufigkeiten der zehn Selbsttests gibt, zeigte ein signifikantes Ergebnis ($F(5.05;843.26) = 3.25$, $p = .006$, $\eta^2 = .019$). Mittels Post-hoc-Tests *Paarweise Vergleiche inkl. Anpassung für Mehrfachvergleiche nach Bonferroni* wurde festgestellt, dass sich die Durchführungshäufigkeit von ST 7 signifikant von der des ST 8 (*Mittlere Differenz* = 2.14, *Standardfehler* = 0.63, $p = .041$) und der des ST 9 (*Mittlere Differenz* = 2.44, *Standardfehler* = 0.58, $p = .002$) unterscheidet. Zwischen allen anderen Selbsttests gibt es keine signifikanten Unterschiede in der Durchführungshäufigkeit. Tabelle 19 zeigt die Mittelwerte, Standardabweichungen, Mediane, Quartile sowie Minima und Maxima der zehn Selbsttests.

Tabelle 19: Mittelwerte (M), Standardabweichungen (SD), Mediane (50. Perzentil), 25. und 75. Perzentile, Minima (Min), Maxima (Max) der Durchführungshäufigkeiten der zehn Selbsttests

			Perzentile				
Selbsttest	*M*	*SD*	*25.*	*50.*	*75.*	*Min*	*Max*
1	6.70	6.72	2	4.5	9	0	47
2	7.33	9.89	2	4	7.75	0	59
3	6.51	7.76	2	4	8	0	42
4	5.83	6.45	1	4	8	0	31
5	6.70	10.21	2	4	7	0	78
6	5.74	6.81	1	4	8	0	43
7	7.45	10.87	1	4	8	0	53
8	5.32	8.40	1	3	6	0	57
9	5.02	7.32	0	3	6	0	36
10	5.28	8.22	0	3	6	0	48

Die signifikanten Ergebnisse wurden mittels non-parametrischer Verfahren, d.h. *Friedman-Test* ($Chi^2(9) = 74.63$, $p < .001$) und *Vorzeichentests (i.e. Binomialtests) inkl. Bonferroni-Alpha-Adjustierung* überprüft. Hierbei zeigte sich, dass nur der signifikante Unterschied zwischen ST 7 und ST 9 abgesichert werden konnte ($z = -3.698$, $p < .001$), nicht aber der Unterschied zwischen ST 7 und ST 8 ($z = -2.200$, $p = .028$).

[20] Neben der Voraussetzung der Normalverteilung war zusätzlich auch die Voraussetzung der Sphärizität verletzt, wie der Mauchly-Test zeigt ($\chi^2(44) = 385.15$, $p < .001$), weshalb die Korrektur der Freiheitsgrade nach Greenhouse-Geisser ($\varepsilon = .56$) angewendet wurde.

Zusammenfassend kann man sagen, dass der Selbsttest 7 signifikant häufiger durchgeführt worden ist als der Selbsttest 9. Alle anderen Selbsttests unterscheiden sich nicht signifikant in ihrer Durchführungshäufigkeit.

1d) Gibt es einen Unterschied in der ST-Durchführungshäufigkeit zwischen den Selbsttests des Wintersemesters und denen des Sommersemesters?

Ob es einen Unterschied in der ST-Durchführungshäufigkeit zwischen WS und SS gibt, wurde mittels *T-Test für abhängige Gruppen* überprüft. Es gibt keine signifikanten Unterschiede in der Durchführungshäufigkeit der Selbsttests des WS und der Durchführungshäufigkeit der Selbsttests des SS ($t(167) = 1.58$, $p = .116$). Tabelle 20 zeigt die Mittelwerte, Standardabweichungen, Quartile, Minima und Maxima der ST-Durchführungshäufigkeiten des WS, des SS und zusätzlich für das gesamte Studienjahr, d.h. die Durchführungshäufigkeiten aller zehn Selbsttests zusammengefasst.

Tabelle 20: Mittelwerte (M), Standardabweichungen (SD), Mediane (50. Perzentil), 25. und 75. Perzentile, Minima (Min), Maxima (Max) der ST-Durchführungshäufigkeiten des WS (ST 1-5), des SS (ST 6-10) und für das gesamte Studienjahr (ST 1-10)

			Perzentile				
Selbsttests	*M*	*SD*	*25.*	*50.*	*75.*	*Min*	*Max*
1 – 5 (WS)	33.07	34.08	12	23.5	40.25	0	230
6 – 10 (SS)	28.81	34.88	6	18	33	0	175
1 – 10 (Gesamt)	61.88	59.45	25.25	44	71.75	1	294

1e) Gibt es einen Unterschied zwischen den Studierenden des Track A und Track B in der ST-Durchführungshäufigkeit zusammengefasst für das gesamte Studienjahr, zusammengefasst für die beiden Semester bzw. in den einzelnen zehn Selbsttests?

Die Fragestellung, ob es Unterschiede zwischen den Studierenden des Track A und Track B in der ST-Durchführungshäufigkeit zusammengefasst für das gesamte Studienjahr gibt, wurde mittels *T-Test für unabhängige Gruppen* überprüft. Es gibt keine signifikanten Unterschiede ($t(166) = 0.155$; $p = .877$) zwischen den Teil-

nehmer/innen des Track A (*M* = 62.60, *SD* = 54.93) und den Teilnehmer/innen des Track B (*M* = 61.17, *SD* = 63.98) in der ST-Durchführungshäufigkeit zusammengefasst für das gesamte Studienjahr.

Weiters wurde mittels *Univariater Varianzanalyse mit Messwiederholungen*[21] überprüft, ob es Unterschiede zwischen den Studierenden des Track A und des Track B in der Durchführungshäufigkeit der Selbsttests des WS und des SS gibt. Hierbei zeigte sich, dass es keinen signifikanten Einfluss der Trackzugehörigkeit auf die Durchführungshäufigkeit der Selbsttests des WS und des SS gibt ($F(1;166) = 0.024$, $p = .877$). D.h. es gibt keine Unterschiede in der ST-Durchführungshäufigkeit der Studierenden des Track A und des Track B, weder im WS (Track A: *M* = 31.87, *SD* = 25.47; Track B: *M* = 34.27, *SD* = 41.06) noch im SS (Track A: *M* = 30.73, *SD* = 36.80; Track B: *M* = 26.89, *SD* = 32.95).

Wiederum mittels *Univariater Varianzanalyse mit Messwiederholungen*[22] wurde ferner überprüft, ob es Unterschiede zwischen den Studierenden des Track A und den Studierenden des Track B in der Durchführungshäufigkeit der zehn einzelnen Selbsttests gibt. Die Ergebnisse zeigen, dass es keinen signifikanten Einfluss der Trackzugehörigkeit auf die ST-Durchführungshäufigkeit der zehn Selbsttests gibt ($F(1;166) = 0.024$, $p = .877$). D.h. es gibt keine Unterschiede zwischen den Studierenden des Track A und des Track B in der Durchführungshäufigkeit bei allen zehn Selbsttests. Tabelle 21 zeigt die Mittelwerte und Standardabweichungen der Durchführungshäufigkeiten der zehn einzelnen Selbsttests, sowie zusammengefasst für das WS, das SS und das gesamte Studienjahr, getrennt nach Track A und Track B.

[21] Neben der Verletzung der Voraussetzung der Normalverteilung ist auch die Voraussetzung der Varianzenhomogenität bei den ST-Durchführungshäufigkeiten des WS verletzt, wie der Levene-Test zeigt ($F(1;166) = 5.103$, $p = .025$).

[22] Wiederum ist neben der Verletzung der Voraussetzung der Normalverteilung auch die Voraussetzung der Varianzenhomogenität bei einigen der ST-Durchführungshäufigkeiten verletzt, und zwar wie der Levene-Test zeigt bei den Selbsttests 1, 2, 4, 5 und 8 ($F(1;166) = 4.040 – 5.432$, $p = .021 – .046$).

Tabelle 21: Mittelwerte (M) und Standardabweichungen (SD) der Durchführungshäufigkeiten der einzelnen Selbsttests, sowie zusammengefasst für das WS, das SS und das gesamten Studienjahr, getrennt nach Track A und Track B

	Track A		**Track B**	
Selbsttest	*M*	*SD*	*M*	*SD*
1	6.77	4.67	6.62	8.32
2	6.81	7.88	7.85	11.58
3	6.70	7.64	6.32	7.93
4	5.46	5.41	6.20	7.36
5	6.12	7.05	7.29	12.62
6	6.37	7.77	5.12	5.67
7	7.75	11.32	7.15	10.47
8	5.92	9.19	4.71	7.53
9	5.19	8.14	4.85	6.44
10	5.50	7.86	5.06	8.61
WS	31.87	25.47	34.27	41.06
SS	30.73	36.80	26.89	32.95
Gesamt	62.60	54.93	61.17	63.98

Zusammenfassend kann man festhalten, dass es keine signifikanten Unterschiede zwischen den Studierenden des Track A und den Studierenden des Track B in der Durchführungshäufigkeit weder zusammengefasst für das ganze Studienjahr, noch für die beiden Semester (WS und SS), noch für die einzelnen zehn Selbsttests gibt.

1f) Gibt es unterschiedliche Typen von Selbsttestnutzer/inne/n?

Wie schon im Kapitel 2.2.1 – Aufbereitung und Komprimierung der Daten – geschildert, lieferte die Clusteranalyse keine zufriedenstellende Cluster/Typen, weshalb mittels Literaturstudium und theoriegeleitetem Wissen vier Typen der Selbsttest-Nutzung erstellt wurden. Die detaillierte Beschreibung der Erstellung der Typen befindet sich im Kapitel 2.2.1 – Aufbereitung und Komprimierung der Daten.

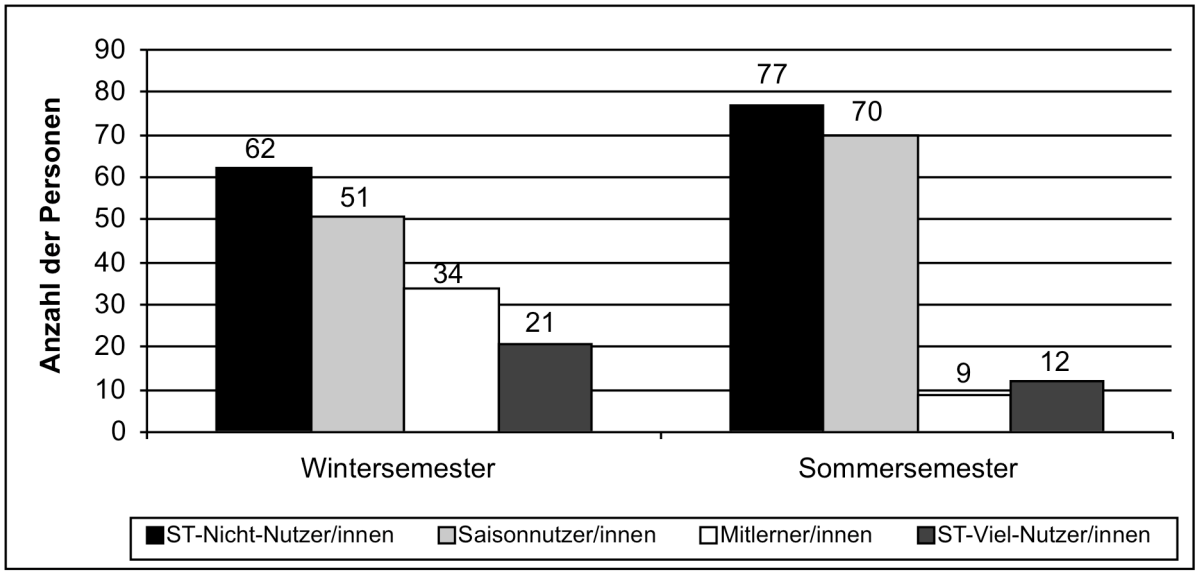


Abbildung 14: Anzahl der Personen im jeweiligen ST-Nutzungs-Typ für das WS und SS

Abbildung 14 zeigt die Verteilung der Studierenden in den vier Typen der Selbsttestnutzung getrennt für das Winter- und Sommersemester. Im WS haben 62 Studierende (36.9%) die Selbsttests im Semester und vor der Prüfung gar nicht bis selten durchgeführt haben (ST-Nicht-Nutzer/innen), im SS sind 77 Studierende (45.8%) ST-Nicht-Nutzer/innen. 21 Studierende (12.5%) verwendeten die Selbsttests im WS mittel bis sehr oft und auch vor der Prüfung des WS führten diese Studierenden die Selbsttests mittel bis sehr oft durch (ST-Viel-Nutzer/innen), im SS sind 12 Personen (7.1%) ST-Viel-Nutzer/innen. Saisonnutzer/innen werden jene Studierenden genannt, welche die Selbsttests im Semester gar nicht bis selten nutzten, jedoch vor der Prüfung die Selbsttest mittel bis sehr oft durchführten. Von diesem ST-Nutzungs-Typ gibt es im WS 51 Studierende (30.4%) und im SS 70 Personen (41.7%). Die Studierenden, welche im Semester die Selbsttests mittel bis sehr oft nutzen, aber vor der Prüfung diese gar nicht bis sehr selten durchführen, werden in dieser Arbeit als Mitlerner/innen bezeichnet. Im WS waren 34 Personen Mitlerner/innen (20.2%) und neun Studierende (5.4%) im SS.

1g) Gibt es einen Unterschied zwischen Track A und Track B in der Art der Selbsttestnutzung?

Abbildung 15 zeigt die Verteilung der Studierenden in den vier Typen der Selbsttestnutzung getrennt nach Track A und Track B für das Winter- und Sommersemester.

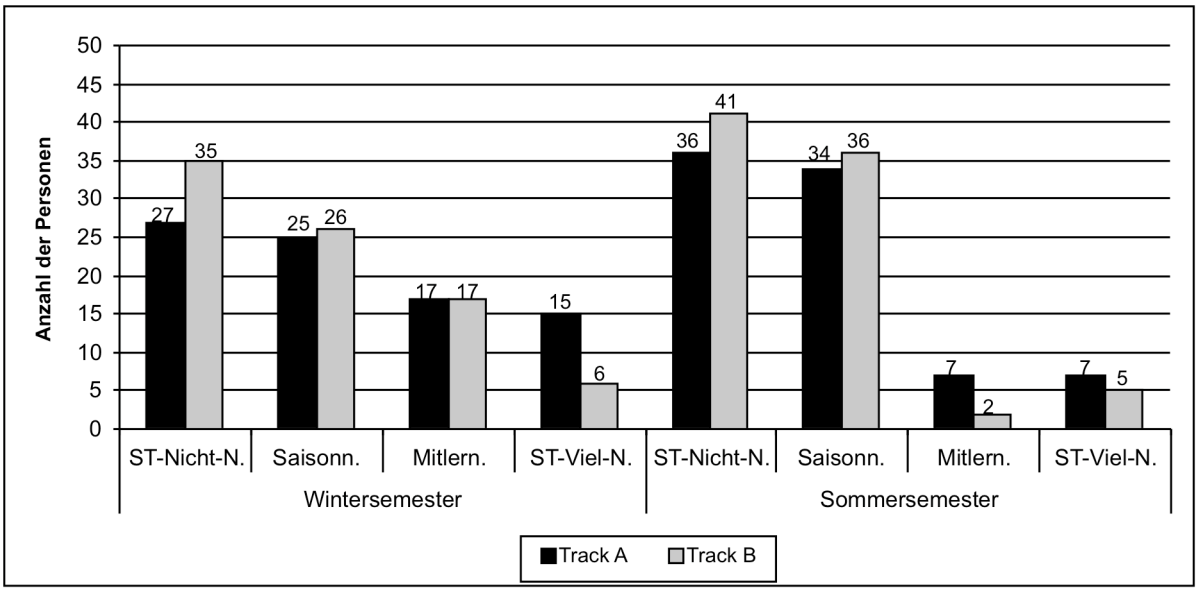


Abbildung 15: Anzahl der Personen im jeweiligen ST-Nutzungs-Typ getrennt nach Track A und B für das WS und SS; *Legende:* ST-Nicht-N. = ST-Nicht-Nutzer/innen; Saisonn. = Saisonnutzer/innen; Mitlern. = Mitlerner/innen; ST-Viel-N. = ST-Viel-Nutzer/innen

Dieser Fragestellung wurde mittels einer *2-Stichproben-KFA* nachgegangen, hierbei wurde geschaut, ob die Studierenden des Track A und Track B in den vier ST-Nutzungs-Typen unterschiedlich vertreten sind, jeweils für das WS und für das SS. Obwohl diese Fragestellung auch mit einer Kreuztabelle beantwortet werden könnte, wurde die 2-Stichproben-KFA gewählt, da hierbei zusätzlich unmittelbar feststellbar ist, ob es Unterschiede zwischen Track A und B hinsichtlich der Verteilung in einem der vier ST-Nutzungs-Typen gibt. Tabelle 22 zeigt im Detail die Ergebnisse der 2-Stichproben-KFAs für WS und SS inkl. der Häufigkeiten der Konfigurationen.

Tabelle 22: Ergebnisse der 2-Stichproben-KFA für das WS und das SS

Konfiguration	**Wintersemester**			**Sommersemester**		
NT-A/B	*f*	*Statistik*	*p*	*f*	*Statistik*	*p*
11	27			36		
12	35	1.253	.263	41	.384	.536
21	25			34		
22	26	.000	1.000	36	.024	.876
31	17			7		
32	17	.037	.848	2	1.787	.171
41	15			7		
42	6	3.483	.062	5	.090	.765

Legende: NT = ST-Nutzungs-Typ (ST-Nicht-Nutzer/innen = 1; Saisonnutzer/innen = 2; Mitlerner/innen = 3; ST-Viel-Nutzer/innen = 4) A/B = Trackzugehörigkeit (Track A = 1; Track B = 2)

Wie aus Tabelle 22 abzulesen ist, gibt es weder im WS noch im SS einen signifikanten Unterschied zwischen Track A und Track B hinsichtlich der Zugehörigkeit zu einen ST-Nutzungs-Typen, d.h. in allen ST-Nutzungs-Typen im WS sowie im SS sind die Studierenden des Track A und die Studierenden des Track B in etwa gleich verteilt. *Anders formuliert*: Es gibt keinen signifikanten Unterschied zwischen den Studierenden des Track A und den Studierenden des Track B in der Art der Selbsttestnutzung.

1h) Gibt es aufgrund unterschiedlicher Selbsttestnutzung einen Unterschied in der Prüfungsleistung in F&E I bzw. II (gesamt und getrennt nach Reproduktion, Verständnis und Produktion)?

Diese Fragestellung wurde spezifiziert: *Gibt es einen Unterschied zwischen den vier Selbsttest-Nutzungs-Typen in der Prüfungsleistung zu FM&E I und II (gesamt und getrennt nach Reproduktion, Verständnis und Produktion)?*

Als *unabhängige Variablen* wurden einerseits die *ST-Nutzungs-Typen WS und SS* mit jeweils vier Ausprägungen (ST-Nicht-Nutzer/innen, Saisonnutzer/innen, Mitlerner/innen, ST-Viel-Nutzer/innen) verwendet. Und andererseits, weil die Versuch- bzw. Kontrollbedingung ebenfalls einen Einfluss auf das Leistungsergebnis in den Prüfungen haben könnte wurde als zweite unabhängige Variable die *Trackzugehörigkeit* (Track A und Track B) mit eingeschlossen, um etwaige Interaktionseffekte zwischen Selbsttestnutzung und Trackzugehörigkeit aufzudecken. Es wird jedoch bewusst darauf verzichtet, die einzelnen Ergebnisse der Trackzugehörigkeit zu berichten, da die Untersuchung des Einfluss der Trackzugehörigkeit auf die Leistungen im Vorwissenstest und in den Prüfungen zu FM&E nicht Ziel dieser Forschungsarbeit war.

Als *abhängige Variablen* wurden die *Leistungen* der Studierenden (d.h. *die erreichten normierten Prozentwerte* – siehe dazu Kapitel 2.1.3 – Erhebungsinstrumente) *in den Prüfungen FM&E I und II* herangezogen. Die Prüfungsleistungen wurden in die inferenzstatistischen Analysen einerseits als *Gesamtwert* einbezogen und

andererseits unterteilt in die drei Kategorien *Reproduktion, Verständnis und Produktion.*[23]

Abbildung 16 zeigt die Mittelwerte der erreichten Leistung (gesamt - normierte Prozentwerte) in den Prüfungen zu FM&E I und II, getrennt nach den vier Typen der Selbsttestnutzung.

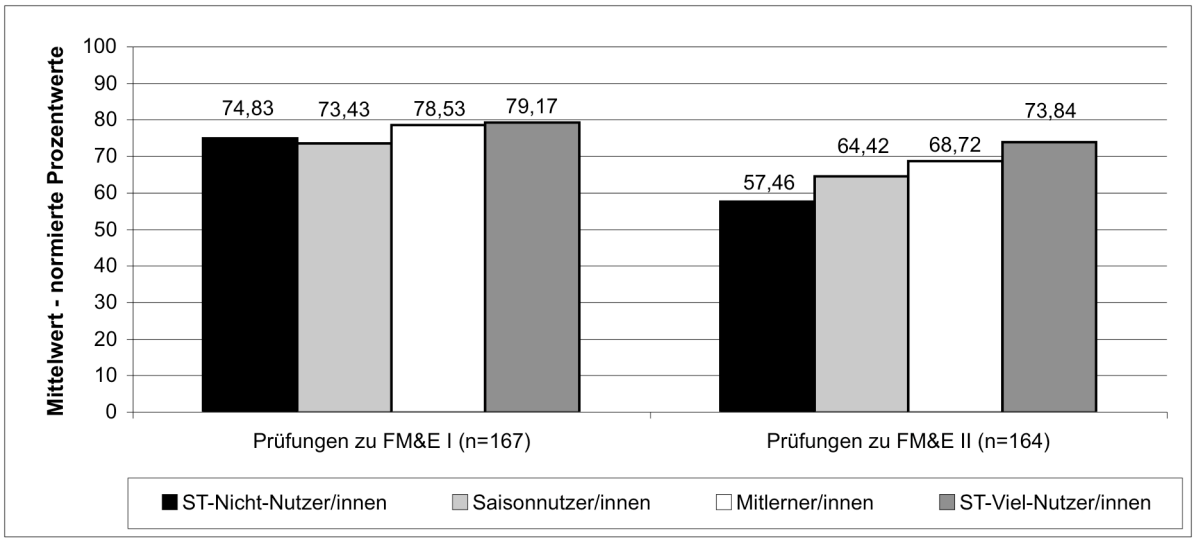


Abbildung 16: Mittelwerte (MW) der erreichten normierte Prozentwerte (gesamt) bei den Prüfungen zu FM&E I und II und Anzahl der angetretenen Personen (n) getrennt nach ST-Nutzungs-Typen

Im ersten Schritt wurde anhand zweier *univariater Varianzanalysen für unabhängige Gruppen* überprüft, ob es Unterschiede zwischen den vier *ST-Nutzungs-Typen* in der *Prüfungsleistung gesamt* bei FM&E I und II gibt. Die Ergebnisse zeigen, dass es keine signifikanten Unterschiede zwischen den vier ST-Nutzungs-Typen des WS in der Prüfungsleistung in FM&E I gibt (*Mittlere Differenz* = 0.64 – 5.74, *Standardfehler* = 1.78 – 2.61, p = .116 – .996), obwohl die Tests der Zwischensubjekteffekte ein signifikantes Ergebnis zeigen ($F(3;159) = 3.15$, p = .026; η^2 = .056). Jedoch gibt es einen signifikanten Unterschied zwischen den vier ST-Nutzungs-Typen des SS in der Prüfungsleistung in FM&E II ($F(3;156) = 5.88$, p = .001; η^2 = .102), und zwar haben ST-Viel-Nutzer/innen signifikant (*Mittlere Differenz* = 16.38, *Standardfehler* = 4.62, p = .007) eine durchschnittlich höhere Gesamtleistung in den Prüfungen zu FM&E II

[23] Die Voraussetzungsprüfung der Normalverteilung zeigte, dass die erreichten normierten Prozentwerte der Prüfungen FM& E I und II im Gesamt relativ normalverteilt sind (obwohl durch die optische Kontrolle der Verteilungen sichtbar wurde, dass es meist viele "Zinken" in den Verteilungen gibt), nicht aber in den einzelnen Kategorien Reproduktion, Verständnis und Produktion, welche rechtssteil und linksschief sind. Die Vorraussetzung der Varianzenhomogenität wurde mittels Levene-Tests überprüft und ist für die folgenden Varianzanalysen erfüllt.

(*M* = 73.84, *SD* = 15.13, *n* = 12) als ST-Nicht-Nutzer/innen (*M* = 57.46, *SD* = 15.07, *n* = 73). Dieses signifikante Ergebnis wurde mittels *Kruskal-Wallis-Test* abgesichert ($\chi^2(3) = 17.77$, $p < .001$). *Anzumerken* ist hier jedoch, dass die vier Gruppen der ST-Nutzer/innen ungleich groß sind (siehe dazu Abbildung 14). Es konnte keine signifikante Interaktion zwischen und den ST-Nutzungs-Typen und der Trackzugehörigkeit festgestellt werden, weder bei der Gesamtleistung in den Prüfungen zu FM&E I ($F(3;159) = 1.44$, $p = .235$) noch bei der Gesamtleistung in den Prüfungen zu FM&E II ($F(3;156) = 0.25$, $p = .862$). Tabelle 23 gibt die Mittelwerte, Standardabweichung und Gruppengrößen der eben berichteten Ergebnisse wieder.

Tabelle 23: Mittelwerte (MW) und Standardabweichungen (SD) der erreichten normierte Prozentwerte bei den Prüfungen zu FM&E I und II und Anzahl der Personen (n) der vier ST-Nutzungs-Typen des WS bzw. des SS und zusätzlich getrennt nach Track A und B

ST-Nutzungs-Typ	**Trackzu-gehörigkeit**	**WS**			**SS**		
		MW	*SD*	*n*	*MW*	*SD*	*n*
ST-Nicht-Nutzer/innen	Track A	76.56	10.05	27	57.29	15.00	33
	Track B	73.46	9.24	34	57.60	15.32	40
	Gesamt	**74.83**	9.65	61	**57.46**	15.07	73
Saison-nutzer/innen	Track A	75.00	11.76	25	64.97	12.56	34
	Track B	71.92	9.44	26	63.89	15.47	36
	Gesamt	**73.43**	10.65	51	**64.42**	14.04	70
Mit-lerner/innen	Track A	76.60	9.10	17	67.21	17.95	7
	Track B	80.46	6.28	17	74.00	8.49	2
	Gesamt	**78.53**	7.94	34	**68.72**	16.11	9
ST-Viel-Nutzer/innen	Track A	78.40	8.11	15	76.01	13.62	7
	Track B	81.09	6.29	6	70.80	18.20	5
	Gesamt	**79.17**	7.58	21	**73.84**	15.13	12
Gesamt	Track A	76.43	10.00	84	62.99	14.99	81
	Track B	74.96	9.19	83	61.52	15.80	83
	Gesamt	75.70	9.61	167	62.25	15.38	164

Im zweiten Schritt wurde mittels zwei *multivariaten Varianzanalysen*[24] überprüft, ob es Unterschiede zwischen den *ST-Nutzungs-Typen* in den *drei Kategorien Reproduktion, Verständnis und Produktion* gibt[25].

Die Ergebnisse zeigen, dass es keine signifikanten Unterschiede zwischen den vier ST-Nutzungs-Typen im WS gibt, weder bei der Kategorie *Reproduktion* ($F(3;159) = 2.40$, $p = .070$), noch bei der Kategorie *Verständnis* ($F(3;159) = 1.33$, $p = .266$). Und obwohl es in der Kategorie *Produktion* ein signifikantes Ergebnis hinsichtlich der ST-Nutzungs-Typen gibt ($F(3;159) = 3.09$, $p = .029$), konnte dieses nicht mittels *Post-hoc-Tests* für die vier ST-Nutzungs-Typen im Einzelnen abgesichert werden (*Mittlere Differenz* = 2.13 – 13.48, *Standardfehler* = 3.50 – 5.13, $p = .051 – .962$). Weiters gibt es keine signifikanten Interaktionseffekte zwischen der ST-Nutzungs-Typen und der Trackzugehörigkeit im WS ($F(3;159) = 0.40 – 1.15$, $p = .330 – .755$).

Im SS zeigte sich in der Kategorie *Reproduktion* ein signifikanter Haupteffekt in Bezug auf die ST-Nutzungs-Typen ($F(3;83) = 3.42$, $p = .021$), welcher aber im Einzelnen nicht abgesichert werden konnte (*Mittlere Differenz* = 3.05 – 27.87, *Standardfehler* = 3.87 – 14.46, $p = .082 – .994$). Ebenso zeigte sich für die Kategorie *Produktion* ein signifikanter Haupteffekt hinsichtlich ST-Nutzungs-Typen ($F(3;83) = 2.79$, $p = .045$), welcher aber wiederum im Einzelnen nicht abgesichert werden konnte (*Mittlere Differenz* = 2.59 – 32.25, *Standardfehler* = 6.02 – 22.51, $p = .111 – .999$). Bei der Kategorie *Verständnis* gibt es keinen signifikanten Haupteffekt hinsichtlich ST-Nutzungs-Typen ($F(3;83) = 0.25$, $p = .865$), aber einen signifikanten Interaktionseffekt zwischen der ST-Nutzungs-Typen und der Trackzugehörigkeit im SS ($F(2;83) = 5.77$, $p = .004$), welcher mittels Univariater Varianzanalyse abgesichert werden konnte ($F(3;156) = 2.76$, $p = .044$). Wobei aber nochmals darauf hingewiesen

[24] Die für multivariate Varianzanalysen zusätzlich geltende Voraussetzung der Korrelation der abhängigen Variablen wurde überprüft und ist erfüllt (siehe Korrelationsmatrizen – Tabelle 53 und Tabelle 54 im Anhang 5.7). Auch die Voraussetzung der Gleichheit der Kovarianzenmatrizen wurde mittels Box-M-Test geprüft und ist für das WS ($F(42;6325) = 1.15$, $p = .229$) sowie auch für das SS ($F(18;17931) = 0.53$, $p = .947$) erfüllt.

[25] Anzumerken ist an dieser Stelle, dass für die multivariate Auswertung des SS ein reduzierter Datensatz von insgesamt 90 Personen (anstatt 168 Studierenden) verwendet werden musste, weil beim 1. Prüfungstermin des SS 2004 keine Items zur Kategorie Produktion vorgegeben wurden. Dies führt dazu, dass einzelne Zellen in der multivariaten Varianzanalyse nicht bzw. sehr gering besetzt sind. Aufgrund dieses Sachverhalts wurden signifikante Ergebnisse der Multivariaten Varianzanalyse mittels Univariater Varianzanalysen und – wenn möglich – mittels Kruskal-Wallis-Test überprüft.

wird, dass die einzelnen Zellenbesetzungen sehr unterschiedlich und zum Teil sehr klein sind (vgl. Tabelle 24).

Tabelle 24: Mittelwerte (MW), Standardabweichungen (SD) und Anzahl der Personen pro Zelle (n) für den Interaktionseffekt ST-Nutzungs-Typen und Trackzugehörigkeit in der Kategorie Verständnis im SS aus der univariaten Varianzanalyse

ST-Nutzungs-Typ	**Trackzugehörigkeit**	*MW*	*SD*	*n*
ST-Nicht-Nutzer/innen	Track A	68.57	19.21	33
	Track B	55.96	25.43	40
Saisonnutzer/innen	Track A	65.71	18.45	34
	Track B	69.49	21.42	36
Mitlerner/innen	Track A	69.00	25.98	**7**
	Track B	87.50	3.54	**2**
ST-Viel-Nutzer/innen	Track A	70.69	23.03	**7**
	Track B	83.00	18.23	**5**

Bei den Kategorien Reproduktion und Produktion konnten auch im SS keine signifikanten Interaktionseffekte festgestellt werden (Reproduktion: $F(2;83) = 0.89$, $p = .415$; Produktion: $F(2;83) = 0.49$, $p = .614$).

Zusammengefasst kann gesagt werden, dass es einen signifikanten Unterschied zwischen den ST-Viel-Nutzer/inne/n und den ST-Nicht-Nutzer/inne/n in der Prüfungsleistung zu FM&E II gibt. ST-Viel-Nutzer/innen haben signifikant durchschnittlich höhere normierte Prozentwerte in den Prüfungen FM&E II gesamt als ST-Nicht-Nutzer/innen. Ansonst konnten keine signifikanten Unterschiede zwischen den vier ST-Nutzungs-Typen festgestellt werden, weder in den Prüfungen zu FM&E I gesamt, noch in den einzelnen Kategorien Reproduktion, Verständnis und Produktion der Prüfungen FM&E I und II. Interaktionseffekte zwischen den vier ST-Nutzungs-Typen und der Trackzugehörigkeit waren alle nicht signifikant, bis auf die Ausnahme der Kategorie Verständnis bei den Prüfungen FM&E II, welche aber angesichts der sehr ungleichen und teilweise extrem kleinen Zellenbesetzungen nur mit Vorsicht zu interpretieren sind und deshalb mit folgenden zusätzlichen Analysen überprüft werden.

2.3.5 Zusätzliche Ergebnisse zu Ziel 1

Um die signifikanten Unterschiede zwischen den beiden ST-Nutzungs-Typen (ST-Nicht-Nutzer/innen und ST-Viel-Nutzer/innen) in der Prüfungsleistung in FM&E II gesamt (siehe Fragestellung 1h) bzw. die signifikante Interaktion bei der Kategorie Verständnis in den Prüfungen FM&E II abzusichern, wurden weitere inferenzstatistische Analysen durchgeführt. Hierfür wurden die Variablen ST-Durchführungs-Häufigkeits-Typen (kurz: ST-Häufigkeits-Typen) des WS und des SS herangezogen, für welche die Studierenden hinsichtlich ihrer ST-Durchführungshäufigkeit im WS bzw. im SS in je vier Gruppen unterteilt wurden. Für diese Unterteilung wurde nur die ST-Durchführungshäufigkeit herangezogen, unabhängig davon wann die Selbsttests genutzt wurden, d.h. der Zeitpunkt der Selbsttestdurchführung fand hierbei keine Berücksichtigung (siehe Kapitel 2.2.1 – Aufbereitung und Komprimierung der Daten). Die Variablen ST-Häufigkeits-Typen (WS und SS) wurden ursprünglich für die Analyse des Ziels 2 erstellt. Die Abbildung 17 veranschaulicht die Verteilung der Studierenden in den vier ST-Häufigkeits-Typen für das WS und das SS getrennt nach Track A und Track B.

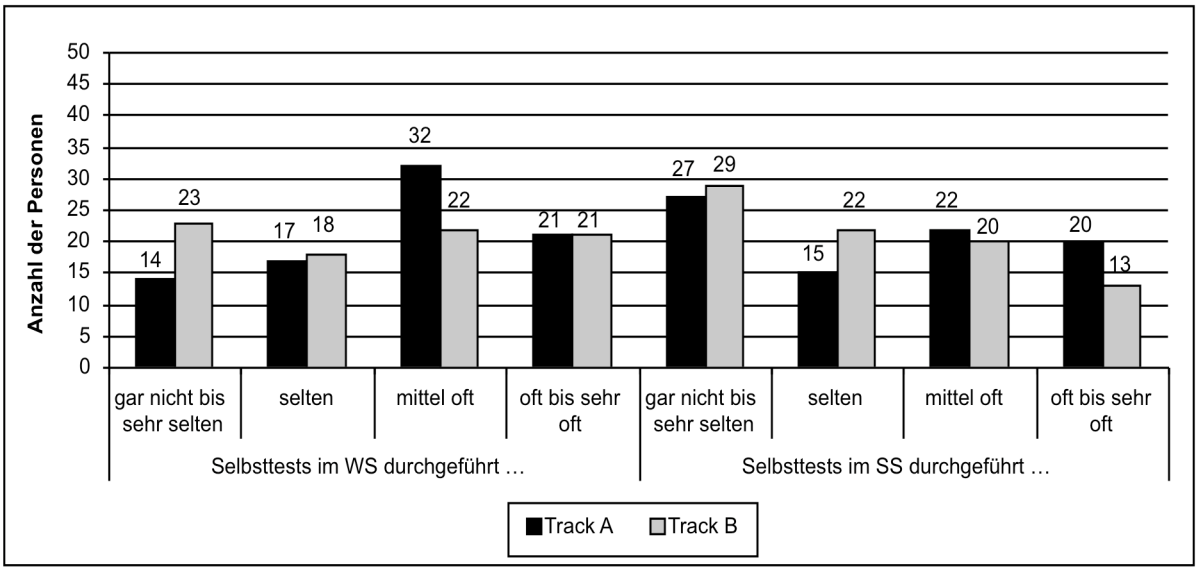


Abbildung 17: Verteilung der Studierenden in den vier ST-Häufigkeits-Typen für das WS und SS gesamt und getrennt nach Trackzugehörigkeit

Die Tabelle 25 zeigt die Verteilung der Studierenden in den vier ST-Häufigkeits-Typen für das WS und das SS, und zusätzlich aufgetrennt nach Track A und Track B. Es gibt keine signifikanten Unterschiede der beobachteten zu den erwarteten

Häufigkeiten in den einzelnen Zellen, d.h. es gibt keine signifikanten Unterschiede in der Häufigkeit der Studierenden in den verschiedenen Gruppen.

Tabelle 25: Verteilung der Studierenden in den vier ST-Häufigkeits-Typen für das WS und SS gesamt und getrennt nach Trackzugehörigkeit, sowie die Ergebnisse der χ^2-*Tests*.

ST-Häufigkeits-Typen	**WS**			**SS**		
Selbsttests durchgeführt...	**Track**			**Track**		
	A	**B**	**Gesamt**	**A**	**B**	**Gesamt**
gar nicht bis sehr selten	14	23	**37** (22.0%)	27	29	**56** (33.3%)
selten	17	18	**35** (20.8%)	15	22	**37** (22.0%)
mittel oft	32	22	**54** (32.2%)	22	20	**42** (25.0%)
oft bis sehr oft	21	21	**42** (25.0%)	20	13	**33** (19.7%)
Gesamt	84	84	**168** (100.0%)	84	84	**168** (100.0%)
$\chi^2(3)$	4.07		5.19	2.98		7.19
p	.254		.158	.395		.066

Gibt es einen Unterschied zwischen den vier Selbsttest-Häufigkeits-Typen in der Prüfungsleistung zu FM&E I und II (gesamt und getrennt nach Reproduktion, Verständnis und Produktion)?

Als *unabhängige Variablen* wurden einerseits, wie so eben berichtet, die ST-Häufigkeits-Typen herangezogen und andererseits die Trackzugehörigkeit. Die *abhängige Variablen* waren wiederum die *Leistungen* der Studierenden *in den Prüfungen FM&E I und II (gesamt* und unterteilt in die Kategorien *Reproduktion, Verständnis und Produktion*) [26].

Bei den Prüfungen zu FM&E I gibt es keine signifikanten Unterschiede zwischen den vier ST-Häufigkeits-Typen in der Prüfungsleistung gesamt ($F(3;159) = 0.86$, $p = .462$) bzw. in den einzelnen Kategorien (Reproduktion: $F(3;159) = 2.14$, $p = .097$;

[26] Die Vorraussetzung der Varianzenhomogenität wurde mittels Levene-Tests überprüft und ist für die folgenden Varianzanalysen wiederum erfüllt (Für die Ergebnisse der Voraussetzungsprüfung der Normalverteilung siehe bei Fragestellung 1h). Auch die Voraussetzung der Gleichheit der Kovarianzenmatrizen für die multivariaten Varianzanalysen wurde mittels Box-M-Tests geprüft und ist erfüllt.

Verständnis: $F(3;159) = 1.10$, $p = .350$; Produktion: $F(3;159) = 0.94$, $p = .423$). Weiters gibt es auch keine signifikanten Interaktionseffekte zwischen ST-Häufigkeits-Typen des WS und der Trackzugehörigkeit weder in der Gesamtleistung der Prüfungen zu FM&E I ($F(3;159) = 0.36$, $p = .782$), noch in den einzelnen Kategorien (Reproduktion: $F(3;159) = 0.46$, $p = .711$; Verständnis: $F(3;159) = 0.54$, $p = .659$; Produktion: $F(3;159) = 0.08$, $p = .969$).

Abbildung 18 veranschaulicht die Mittelwerte der erreichten Leistung (gesamt - normierte Prozentwerte) in den Prüfungen zu FM&E I und II, getrennt nach den vier ST-Häufigkeits-Typen.

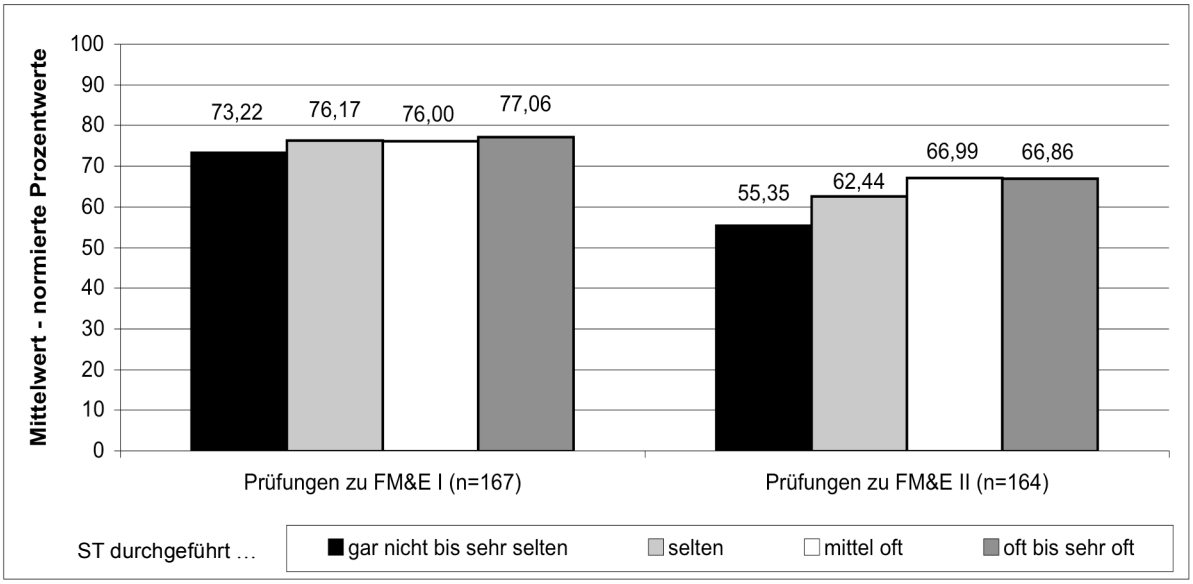


Abbildung 18: Mittelwerte der erreichten normierte Prozentwerte (gesamt) in den Prüfungen zu FM&E I und II und Anzahl der angetretenen Personen (n) getrennt nach ST-Häufigkeits-Typen.

Bei den Prüfungen zu FM&E II gibt es einen signifikanten Unterschied zwischen den vier ST-Häufigkeits-Typen in der Prüfungsleistung gesamt ($F(3;156) = 6.13$, $p = .001$; $\eta^2 = .105$). Studierende, welche die Selbsttests gar nicht bis sehr selten durchgeführt haben, haben signifikant niedrigere durchschnittliche normierte Prozentwerte als Studierende, die die Selbsttests mittel oft (*Mittlere Differenz* = 11.64, *Standardfehler* = 3.03, $p = .003$) bzw. oft bis sehr oft (*Mittlere Differenz* = 11.51, *Standardfehler* = 3.25, $p = .007$) durchgeführt haben. Es gibt aber keine signifikanten Interaktionseffekte zwischen den ST-Häufigkeits-Typen und der Trackzugehörigkeit ($F(3;156) = 1.66$, $p = .178$) in der Leistung gesamt in den Prüfungen zu FM&E II.

Tabelle 26 zeigt die Mittelwerte, Standarbabweichung und Zellenbesetzungen der beiden univariaten Varianzanalyen.

Tabelle 26: Mittelwerte (MW), Standardabweichungen (SD) erreichte normierte Prozentwerte bei den Prüfungen zu FM&E I und II und Anzahl der Personen (n) der vier ST-Häufigkeits-Typen des WS bzw. des SS und zusätzlich getrennt nach Track A und B

ST-Häufig-keits-Typ	**Trackzu-gehörigkeit**	**WS**			**SS**		
		MW	*SD*	*n*	*MW*	*SD*	*n*
ST *gar nicht bis sehr selten* durchgeführt	Track A	75.10	9.30	14	55.34	14.36	24
	Track B	72.02	7.79	22	55.36	15.36	28
	Gesamt	**73.22**	8.42	36	**55.35**	14.76	52
ST *selten* durchgeführt	Track A	75.57	11.73	17	64.88	14.71	15
	Track B	76.74	9.70	18	60.77	14.76	22
	Gesamt	**76.17**	10.59	35	**62.44**	14.68	37
ST *mittel oft* durchgeführt	Track A	76.93	10.83	32	63.79	13.42	22
	Track B	74.65	9.64	22	70.50	14.98	20
	Gesamt	**76.00**	10.33	54	**66.99**	14.41	42
ST *oft bis sehr oft* durchgeführt	Track A	77.27	8.00	21	69.87	14.51	20
	Track B	76.85	9.43	21	62.23	14.68	13
	Gesamt	**77.06**	8.64	42	**66.86**	14.83	33
Gesamt	Track A	76.43	10.00	84	62.99	14.99	81
	Track B	74.96	9.19	83	61.52	15.80	83
	Gesamt	75.70	9.61	167	62.25	15.38	164

Die Ergebnisse der *multivariaten Varianzanalyse* zeigen, dass es zwischen den ST-Häufigkeits-Typen bei den Prüfungen zu FM&E II signifikante Unterschiede bei den Kategorien *Reproduktion* ($F(3;82) = 3.23$, $p = .027$; $\eta^2 = .106$) und *Produktion* ($F(3;82) = 4.42$, $p = .006$; $\eta^2 = .139$), nicht aber bei der Kategorie *Verständnis* ($F(3;82) = 1.05$, $p = .377$) gibt. Diese signifikanten Ergebnisse der multivariaten Varianzanalyse wurden mittels univariater Varianzanalysen abgesichert. Wie die Ergebnisse der univariaten Varianzanalysen bestätigen, haben Studierende, welche die Selbsttests gar nicht bis sehr selten durchgeführt haben, bei den Prüfungen zu FM&E II in der Kategorie *Reproduktion* signifikant durchschnittlich niedrigere normierte Prozentwerte als Studierende, welche die Selbsttests mittel oft (*Mittlere Differenz* = 12.49, *Standardfehler* = 3.68, $p = .011$) bzw. oft bis sehr oft

(*Mittlere Differenz* = 15.27, *Standardfehler* = 3.95, *p* = .002) durchgeführt haben. In der Kategorie *Produktion* der Prüfungen zu FM&E II haben Studierenden, welche die Selbsttests oft bis sehr oft durchgeführt haben signifikant durchschnittlich höhere normierte Prozentwerten als Studierende, die die Selbsttests gar nicht bis sehr selten (*Mittlere Differenz* = 25.60, *Standardfehler* = 7.52, *p* = .012) durchgeführt haben. Ferner konnten bei den Prüfungen zu FM&E II keine signifikanten Interaktionseffekte zwischen den ST-Häufigkeits-Typen und der Trackzugehörigkeit festgestellt werden, weder bei der Kategorie *Reproduktion* ($F(3;82) = 0.74$, $p = .531$), noch bei der Kategorie *Verständnis* ($F(3;82) = 2.10$, $p = .106$) und auch nicht bei der Kategorie *Produktion* ($F(3;82) = 1.31$, $p = .277$). Tabelle 27 zeigt die Mittelwerte, Standarbabweichung und Zellenbesetzungen der univariaten Varianzanalyen der Kategorien *Reproduktion, Verständnis* und *Produktion.*

Tabelle 27: Mittelwerte (MW), Standardabweichungen (SD) erreichte normierte Prozentwerte bei den Prüfungen zu FM&E II in den Kategorien Reproduktion, Verständnis und Produktion und Anzahl der Personen (n) der vier ST-Häufigkeits-Typen des SS und zusätzlich getrennt nach Track A und B

ST-Häufig-keits-Typ	**Track**	**Reproduktion**			**Verständnis**			**Produktion**		
		MW	*SD*	*n*	*MW*	*SD*	*n*	*MW*	*SD*	*n*
ST *gar nicht bis sehr selten* durchgeführt	A	45.47	18.52	24	67.21	17.99	24	50.79	22.67	18
	B	52.90	17.48	28	55.42	24.66	28	41.78	25.05	15
	Ges.	**49.47**	18.18	52	**60.86**	22.43	52	**46.70**	23.84	33
ST *selten* durchgeführt	A	56.90	17.41	15	71.60	20.48	15	77.32	17.62	7
	B	57.74	17.92	22	64.55	22.57	22	47.78	34.69	8
	Ges.	**57.40**	17.48	37	**67.40**	21.74	37	**61.56**	31.11	15
ST *mittel oft* durchgeführt	A	57.94	15.89	22	69.06	17.53	22	64.16	28.07	13
	B	66.38	18.50	20	75.00	22.90	20	69.88	20.79	9
	Ges.	**61.96**	17.49	42	**71.89**	20.23	42	**66.50**	24.97	22
ST *oft bis sehr oft* durchgeführt	A	69.49	17.96	20	63.42	23.06	20	73.91	24.94	14
	B	57.44	18.27	13	66.03	25.14	13	68.52	42.41	6
	Ges.	**64.74**	18.77	33	**64.45**	23.55	33	**72.29**	30.09	20
Gesamt	A	56.90	19.33	81	67.59	19.53	81	63.93	25.78	52
	B	58.14	18.36	83	64.22	24.50	83	53.92	31.08	38
	Ges.	57.53	18.80	164	65.88	22.18	164	59.70	28.41	90

Legende: ST = Selbsttests; Ges. = Gesamt

Abbildung 19 veranschaulicht die erreichten durchschnittlichen Leistungen der vier ST-Häufigkeits-Typen in den einzelnen Kategorien *Reproduktion, Verständnis* und *Produktion* der Prüfungen zu FM&E II.

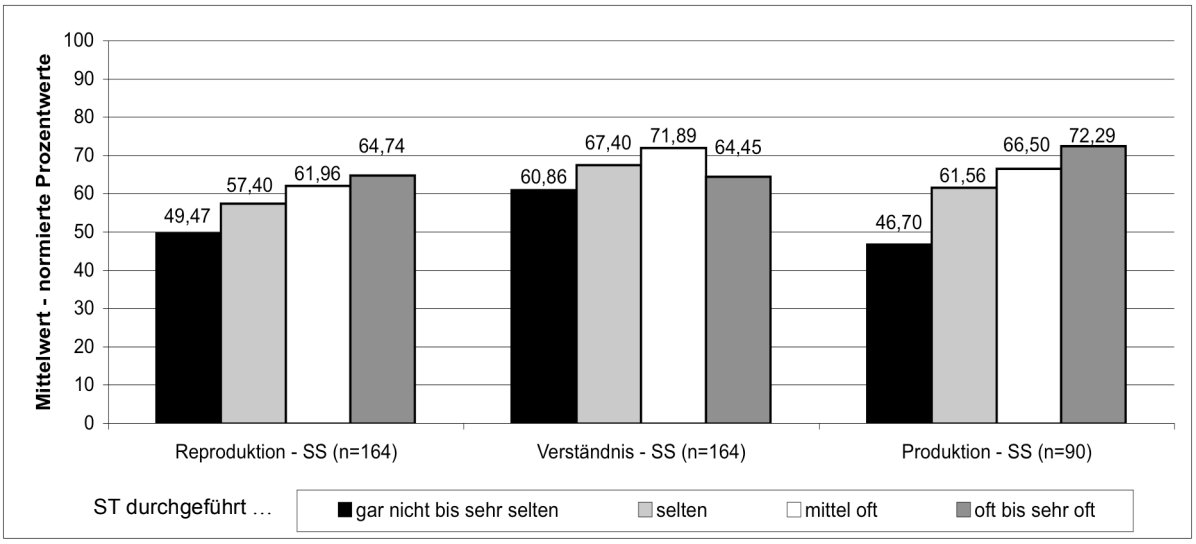


Abbildung 19: Mittelwerte der erreichten normierten Prozentwerte in den Kategorien Reproduktion, Verständnis und Produktion der Prüfungen zu FM&E II und Anzahl der angetretenen Personen (n) getrennt nach ST-Häufigkeits-Typen

Zusammengefasst kann man festhalten, dass in der Prüfungsleistung zu FM&E I zwischen den ST-Häufigkeits-Typen keine signifikanten Unterschiede gibt, weder in der Gesamtleistung, noch in den einzelnen Anforderungskategorien. In den Prüfungen zu FM&E II gibt es signifikante Unterschiede zwischen den ST-Häufigkeits-Typen. Studierende, welche die Selbsttests gar nicht bis sehr selten durchgeführt haben, haben durchschnittlich niedriger normierte Prozentwerte in den Prüfungen gesamt und in den Kategorien Reproduktion und Produktion als Studierende, die die Selbsttest oft bis sehr oft durchgeführt haben. Weiters haben auch Studierenden, welche die Selbsttests oft durchgeführt haben signifikant durchschnittlich höhere normierte Prozentwerte in den Prüfungen gesamt und in der Kategorie Reproduktion als Studierende, die die Selbsttests gar nicht bis sehr selten durchgeführt haben. Es gibt aber keine signifikanten Interaktionseffekte zwischen den ST-Häufigkeits-Typen und der Trackzugehörigkeit, weder in den Prüfungsleistungen gesamt, noch in den einzelnen Kategorien der Prüfungen FM&E I und II.

2.3.6 Deskriptive Ergebnisse zu Ziel 2

In diesem Kapitel werden nun die von den Studierenden in den Erhebungen genannten Attributionen deskriptiv dargestellt, wobei sich alle folgenden Ausführungen, sowie die späteren Analysen und Ergebnisse auf die *erst*genannten als angenommene dominante Attributionen beziehen. Es werden im Folgenden immer zuerst die Tabellen und Erläuterungen der Erfolgsattributionen dargestellt und im Anschluss daran die entsprechenden Ausführungen und Tabellen für die Misserfolgsattributionen.

Die Häufigkeiten der ursprünglichen zwölf Kategorien befinden sich aus Platzgründen im Anhang 5.8 (Tabelle 55 bis 56). Ebenso die Tabellen mit den Häufigkeiten der Kategorien Erfolgsattributionen bzw. Misserfolgsattributionen getrennt nach Track A und B (Tabelle 57 und 58). Abbildung 20 veranschaulicht die Häufigkeiten der erstgenannten *Erfolgsattributionen* der drei Messzeitpunkte (MZP) zusammengefasst in die vier Kategorien nach Weiner (1994; vgl. Kapitel 1.3.3 – Attributionstrainings).

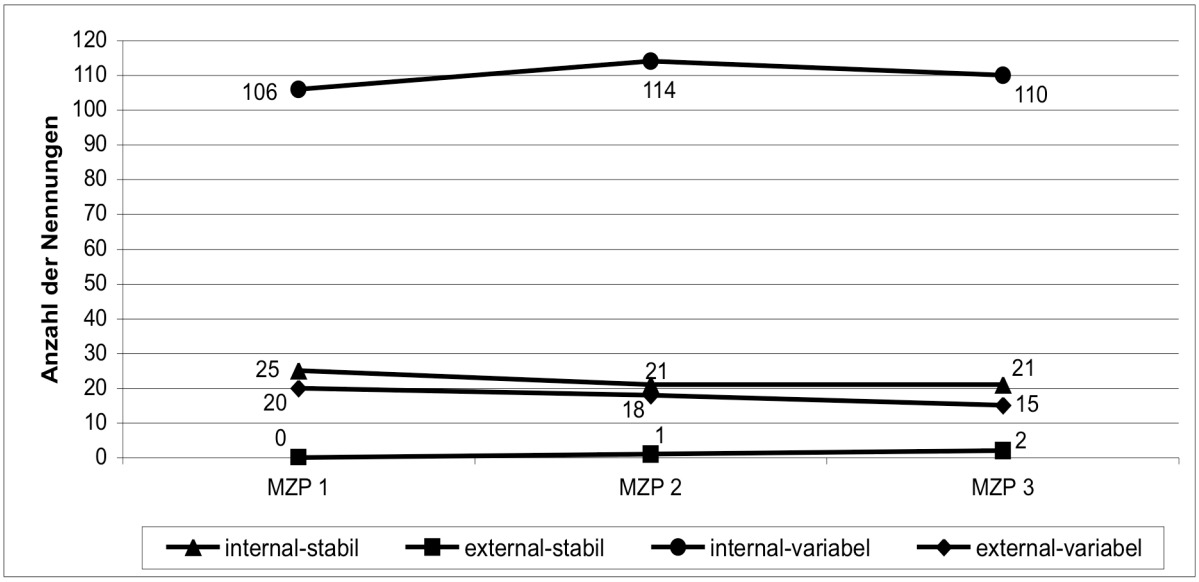


Abbildung 20: Häufigkeiten der Erfolgsattributionen (jeweils die erste Nennung) der drei MZP

Wie in Abbildung 20 und Tabelle 28 zu erkennen ist, wurde die Kategorie external-stabil beim MZP 1 nicht genannt. Und auch beim MZP 2 und 3 wurde diese Kategorie nur ein- bzw. zweimal genannt. Im Gegensatz dazu wurden Erfolgsattributionen der Kategorie internal-variabel durchschnittlich 110-mal genannt.

Die Kategorie external-variabel wurde 25-mal beim MZP 1, 18-mal beim MZP 2 und 15-mal beim MZP 3 genannt. Die Kategorie internal-stabil wurde beim ersten Messzeitpunkt 25-mal und bei den beiden letzten MZP je 21-mal genannt. Tabelle 28 zeigt nochmals die Häufigkeiten und zusätzlich die Prozente der erstgenannten *Erfolgsattributionen* aller drei Messzeitpunkte der vier Kategorien.

Tabelle 28: Häufigkeiten der erstgenannten Erfolgsattributionen der drei MZP

Erfolgsattributionen	**MZP 1**	**MZP 2**	**MZP 3**
internal-stabil	25 (15.3%)	21 (12.5%)	21 (12.7%)
external-stabil	0 (0.0%)	1 (0.6%)	2 (1.2%)
internal-variabel	106 (65.0%)	114 (67.9%)	110 (66.7%)
external-variabel	20 (12.3%)	18 (10.7%)	15 (9.1%)
nicht zuordenbar	12 (7.4%)	14 (8.3%)	17 (10.3%)
Gesamt	163 (100.0%)	168 (100.0%)	165 (100.0%)
Fehlend	5	0	3

Abbildung 21 veranschaulicht die Häufigkeiten der erstgenannten *Misserfolgs-attributionen* der drei Messzeitpunkte wiederum zusammengefasst in die vier Kategorien nach Weiner (1994).

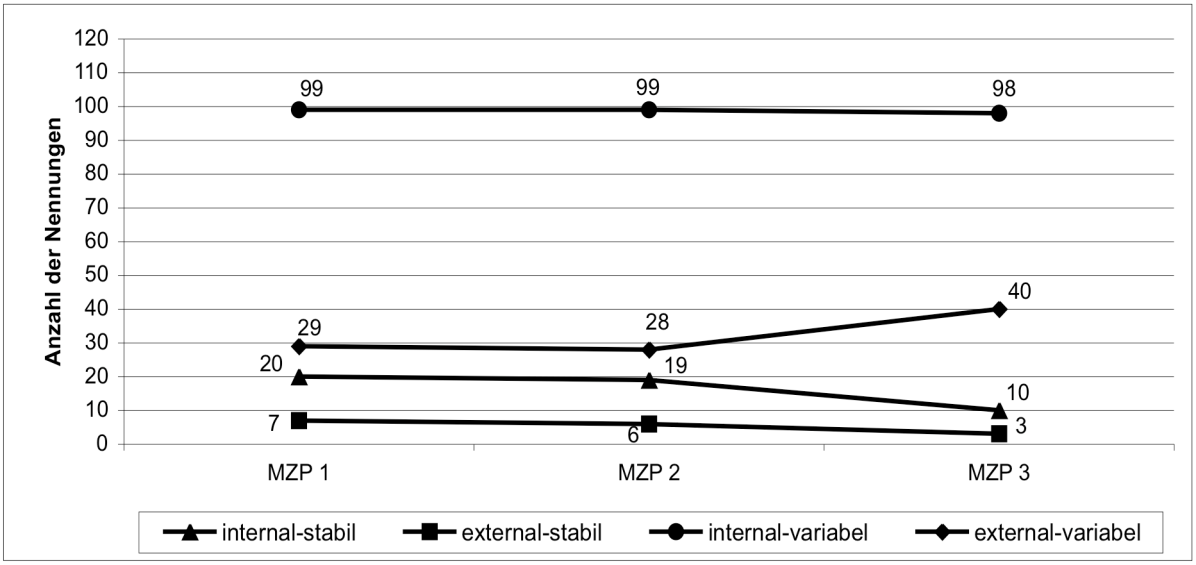


Abbildung 21: Häufigkeiten der Misserfolgsattributionen (jeweils die erste Nennung) der drei MZP

Wie aus Abbildung 21 und Tabelle 29 abzulesen ist, wurde die Misserfolgs-attributionen, welche zur Kategorie internal-stabil gezählt werden, beim ersten und zweiten MZP 20- bzw. 19-mal genannt, beim dritten Messzeitpunkt zehnmal. Die

Kategorie external-stabil wurde sieben-, sechs- und dreimal genannt. Während die Kategorie internal-variabel bei allen drei Messzeitpunkten je ca. 99-mal genannt wurde. Die Kategorie external-variabel wurde beim ersten und zweiten Messzeitpunkt 29- bzw. 28-mal genannt und beim dritten Messzeitpunkt 40-mal. Tabelle 29 zeigt wiederum die Häufigkeiten und die dazugehörigen Prozente der erstgenannten *Misserfolgsattributionen* aller drei Messzeitpunkte der vier Kategorien.

Tabelle 29: Häufigkeiten der erstgenannten Misserfolgsattributionen der drei MZP

Misserfolgsattributionen	**MZP 1**	**MZP 2**	**MZP 3**
internal-stabil	20 (12.2%)	19 (11.5%)	10 (6.1%)
external-stabil	7 (4.3%)	6 (3.6%)	3 (1.8%)
internal-variabel	99 (60.4%)	99 (60.0%)	98 (60.1%)
external-variabel	29 (17.7%)	28 (17.0%)	40 (24.5%)
nicht zuordenbar	9 (5.5%)	13 (7.9%)	12 (7.4%)
Gesamt	164 (100.0%)	165 (100.0%)	163 (100.0%)
Fehlend	4	3	5

Folgenden Tabellen 30 bis 33 zeigen die Veränderungen der erstgenannten Attributionen im Detail. Tabelle 30 zeigt die Veränderungen der Erfolgsattributionen von MZP 1 auf MZP 2.

Tabelle 30: Häufigkeiten und Prozente der Veränderungen der erstgenannten Erfolgsattributionen von MZP 1 auf MZP 2

Erfolgs-attributionen		**MZP 2**				
		internal-stabil	external-stabil	internal-variabel	external-variabel	nicht zu-ordenbar
MZP 1	internal-stabil	10 (6.1%)	0 (0.0%)	11 (6.7%)	1 (0.6%)	3 (1.8%)
	external-stabil	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	internal-variabel	8 (4.9%)	0 (0.0%)	80 (49.1%)	13 (8.0%)	5 (3.1%)
	external-variabel	2 (1.2%)	1 (0.6%)	12 (7.4%)	3 (1.8%)	2 (1.2%)
	nicht zu-ordenbar	1 (0.6%)	0 (0.0%)	6 (3.7%)	1 (0.6%)	4 (2.5%)

Wie in Tabelle 30 zu erkennen ist, kam es bei 97 Nennungen (59.5%) in der Kategorienzugehörigkeit zu keiner Veränderung von ersten MZP auf zweiten MZP. Je 14 Studierende (8.6%) nannten beim MZP 1 internale bzw. externale, und dann beim MZP 2 externale bzw. internale Erfolgsattributionen. 22 Attributionen (13.5%) waren bei einem oder bei beiden MZP den Kategorien nicht zuordenbar. Tabelle 31 zeigt die nun Veränderungen der erstgenannten Erfolgsattributionen von MZP 2 auf MZP 3.

Tabelle 31: Häufigkeiten und Prozente der Veränderungen der erstgenannten Erfolgsattributionen von MZP 2 auf MZP 3

Erfolgs-attributionen		**MZP 3**				
		internal-stabil	external-stabil	internal-variabel	external-variabel	nicht zu-ordenbar
MZP 2	internal-stabil	9 (5.5%)	0 (0.0%)	9 (5.5%)	0 (0.0%)	3 (1.8%)
	external-stabil	0 (0.0%)	0 (0.0%)	1 (0.6%)	0 (0.0%)	0 (0.0%)
	internal-variabel	9 (5.5%)	2 (1.2%)	86 (51.2%)	10 (6.1%)	6 (3.6%)
	external-variabel	2 (1.2%)	0 (0.0%)	13 (7.9%)	2 (1.2%)	1 (0.6%)
	nicht zu-ordenbar	1 (0.6%)	0 (0.0%)	1 (0.6%)	3 (1.8%)	7 (4.2%)

Zu keiner Veränderung in der Kategorienzugehörigkeit von MZP 2 auf MZP 3 kam es bei 104 Nennungen von Erfolgsattributionen (63.0%). 16 Personen (9.7%), welche beim MZP 2 eine externale Erfolgsattribution nannten, gaben beim MZP 3 internale Erfolgsattributionen. Umgekehrt nannten 12 Studierende (7.3%) beim MZP 2 internale und dann beim MZP 3 externale Erfolgsattributionen.

Wie Tabelle 32 zeigt, kam es bei den 92 Nennungen der Misserfolgsattributionen (56.1%) zu keiner Veränderung in der Kategorienzugehörigkeit von MZP 1 auf MZP 2, wovon 71 Nennungen (43.8%) internal-variabel sind (siehe Tabelle 32). 26 Studierende (15.9%), welche beim MZP 1 diverse andere Misserfolgsattributionen nannten, gaben beim MZP 2 internal-variable Misserfolgsattributionen an. Umgekehrt nannten 23 Personen (14.0%) beim MZP 1 internal-variable Misserfolgsattributionen und beim MZP 2 nannten sie Misserfolgsattributionen anderer Kategorien.

Tabelle 32: Häufigkeiten und Prozente der Veränderungen der erstgenannten Misserfolgsattributionen von MZP 1 auf MZP 2

Misserfolgs-attributionen		**MZP 2**				
		internal-stabil	external-stabil	internal-variabel	external-variabel	nicht zu-ordenbar
MZP 1	internal-stabil	8 (4.9%)	0 (0.0%)	10 (6.2%)	2 (1.2%)	0 (0.0%)
	external-stabil	1 (0.6%)	0 (0.0%)	2 (1.2%)	3 (1.9%)	0 (0.0%)
	internal-variabel	8 (4.9%)	2 (1.2%)	71 (43.8%)	13 (8.0%)	4 (2.5%)
	external-variabel	1 (0.6%)	1 (0.6%)	11 (6.8%)	10 (6.2%)	6 (3.7%)
	nicht zu-ordenbar	0 (0.0%)	3 (1.9%)	3 (1.9%)	0 (0.0%)	3 (1.9%)

In Tabelle 33 werden nun die Veränderungen der erstgenannten Misserfolgsattributionen von MZP 2 auf MZP 3 wiedergegeben. Die Nennungen von 89 Studierenden (54.6%) ändern sich von MZP 2 auf MZP 3 nicht in der Kategorienzugehörigkeit, wovon 67 Nennungen (41.6%) internal-variabel sind. 30 Studierende (18.4%) gaben bei MZP 3 internal-variable Misserfolgsattributionen an, obwohl sie bei MZP 2 diverse andere Misserfolgsattributionen nannten. Umgekehrt nannten 26 Personen (16.0%) internal-variable Misserfolgsattribution beim MZP 2, jedoch bei MZP 3 gaben sie Misserfolgsattributionen anderer Kategorien an.

Tabelle 33: Häufigkeiten und Prozente der Veränderungen der erstgenannten Misserfolgsattributionen von MZP 2 auf MZP 3

Misserfolgs-attributionen		**MZP 3**				
		internal-stabil	external-stabil	internal-variabel	external-variabel	nicht zu-ordenbar
MZP 2	internal-stabil (1)	4 (2.5%)	0 (0.0%)	9 (5.6%)	3 (1.9%)	2 (1.2%)
	external-stabil (2)	0 (0.0%)	1 (0.6%)	1 (0.6%)	3 (1.9%)	1 (0.6%)
	internal-variabel (3)	4 (2.5%)	2 (1.2%)	67 (41.6%)	20 (12.4%)	3 (1.9%)
	external-variabel (4)	2 (1.2%)	0 (0.0%)	15 (9.3%)	11 (6.8%)	0 (0.0%)
	nicht zu-ordenbar (6)	0 (0.0%)	0 (0.0%)	5 (3.1%)	2 (1.2%)	6 (3.7%)

2.3.7 Beantwortung der Fragestellungen zu Ziel 2

In diesem Kapitel werden nun die Fragestellungen zum Ziel 2 der Diplomarbeit, die Evaluation des web-basierten attributionalen Feedbacks, beantwortet.

2a) Verbessert das schriftliche attributionale Feedback, welches den Studierenden der Versuchsgruppe (Track A) dargeboten wurde, nicht aber der Kontrollgruppe (Track B), die Erfolgs- bzw. Misserfolgsattributionen der Studierenden?

Diese Fragestellung wurde mittels einer *Konfigurationsfrequenzanalyse (KFA) der 1. Ordnung* geprüft. Hierfür wurden im ersten Schritt die vier Kategorien der Erfolgs- bzw. Misserfolgsattributionen in Bezugnahme auf die Ergebnisse der Forschungsliteratur (vgl. Dresel, 2004; Grünke & Castello, 2008; Robertson, 2000; Ziegler & Finsterwald, 2008; Ziegler & Schober, 2001; siehe dazu auch Kapitel 1.3.3 – Attributionstrainings) bewertet. Tabelle 34 dokumentiert die Bewertung der vier Kategorien der Attributionen.

Tabelle 34. Bewertung der vier Kategorien der Attributionen nach Erfolg und nach Misserfolg

		Bewertung			
Kategorien		*Erfolgsattributionen*		*Misserfolgsattributionen*	
1	internal-stabil	1	Positiv	2	Negativ
2	external-stabil	2	Negativ	2	Negativ
3	internal-variabel	1	Positiv	1	Positiv
4	external-variabel	2	Negativ	2	Negativ
5	Attributionsunsicherheit	9	Missing	9	Missing
6	Nicht zuordenbar	9	Missing	9	Missing

Im Anschluss daran wurden die bewerteten Attributionen des ersten (bzw. zweiten) Messzeitpunkt (+/-), die bewerteten Attributionen des zweiten (bzw. dritten) Messzeitpunkt (+/-), die vier ST-Häufigkeits-Typen (1/2/3/4) und die Trackzugehörigkeit (A/B) miteinander kombiniert. Dies ergibt 2 x 2 x 4 x 2 = 32 Kombinationen. Da jedoch einzelne Kombinationen zu kleinen Häufigkeiten aufwiesen (d.h. zu viele Konfigurationen mit der Häufigkeit Null; $fe < 0.5$), wurden die vier ST-Häufigkeits-

Typen zu zwei Gruppen (1/2) zusammengefasst, in (1) Studierende, welche die Selbsttests gar nicht bis selten durchführt haben und (2) Studierende, welche die Selbsttests mittel oft bis sehr oft durchgeführt haben. Anschließend wurden wiederum Kombinationen (2 x 2 x 2 x 2 = 16) der vier Variablen erstellt – von 1111 bis 2222. Beispielsweise bedeutet die Zugehörigkeit zur Konfiguration 1111, dass eine Person zum ersten (bzw. zweiten) Messzeitpunkt eine positive Attribution angegeben hat, ebenso zum zweiten (bzw. dritten) Messzeitpunkt, sie hat die Selbsttests gar nicht bis selten durchgeführt und war in der Versuchgruppe. Die Konfigurationen 2222 im Gegensatz dazu bedeutet, dass den Studierenden zum ersten (bzw. zweiten) Messzeitpunkt eine negative Attribution angegeben wurde, ebenso zum zweiten (bzw. dritten) Messzeitpunkt, die Selbsttests wurden mittel oft bis sehr oft durchgeführt und es handelt sich um Studierende der Kontrollgruppe.

Im nächsten Schritt wurde a priori definiert, welche Konfigurationen belegen, dass das schriftliche attributionale Feedback zu einer Verbesserung der Attribution führt. Als *Erfolg des attributionale Feedbacks* wird gewertet, wenn die *Konfiguration 2121 überzufällig vorkommen* würde, *nicht aber* die *Konfigurationen 2122 und 2111*. Die Konfiguration *2121* bedeutet, dass zum ersten (bzw. zweiten) Messzeitpunkt eine negative Erfolgs- bzw. Misserfolgsattribution angegeben wurde, jedoch beim zweiten (bzw. dritten) Messzeitpunkt wurde eine positive Erfolgs- bzw. Misserfolgsattribution genannt, die Selbsttests wurden mittel oft bis sehr oft durchgeführt und das attributionale Feedback wurde dargeboten. Die Konfiguration *2122* bedeutet, dass wiederum zum ersten (bzw. zweiten) Messzeitpunkt eine negative Erfolgs- bzw. Misserfolgsattribution genannt wurde, aber eine positive Erfolgs- bzw. Misserfolgsattribution beim zweiten (bzw. dritten) Messzeitpunkt angegeben wurde, auch die Selbsttests wurden mittel oft bis sehr oft durchgeführt, jedoch wurde kein schriftliches attributionales Feedback dargeboten. Die Konfiguration *2111* beinhaltet auch die positive Veränderung bei der Nennung der Attributionen, jedoch wurden die ST gar nicht bis selten durchgeführt, d.h. auch das attributionale Feedback wurde gar nicht bis selten dargeboten.

Es wurden vier *KFAs der 1. Ordnung* durchgeführt inkl. *Lehmachers Test mit Kuechenhoff-Kontinuitätskorrektur* und *Alpha-Adjustierung nach Bonferroni* (α = .05/3 = .017). Tabelle 35 und 36 veranschaulichen die Ergebnisse.

Tabelle 35: Ergebnisse der beiden KFA – 1. Ordnung der Erfolgsattributionen für WS und SS

	Wintersemester				**Sommersemester**			
Pearson χ^2 *(df),* sig.	21.89 (11), p = .025				8.112 (11), p = .703			
	Häufigkeiten		Statistiken		Häufigkeiten		Statistiken	
Konfiguration	o	e	z	*p(z)*	o	e	z	*p(z)*
2121	5	4.74	-0.135	.446	6	3.74	1.060	.145
2122	3	4.06	-0.326	.372	3	3.30	0.127	.449
2111	1	3.72	-1.342	.090	3	4.89	-0.766	.222

Legende: erste Stelle der Konfiguration ist (2) negative Attribution zum ersten (bzw. zweiten) MZP; zweite Stelle die positive Attribution zum zweiten (bzw. dritten) MZP; dritte Stelle beinhaltet die ST-Durchführungshäufigkeit (1 = gar nicht bis selten; 2 = mittel oft bis sehr oft); vierte Stelle ist die Trackzugehörigkeit (1 = Track A; attributionales Feedback wurde dargeboten; 2 = Track B; attributionales Feedback wurde nicht dargeboten)

Tabelle 36: Ergebnisse der beiden KFA – 1. Ordnung der Misserfolgsattributionen für WS und SS

	Wintersemester				**Sommersemester**			
Pearson χ^2 *(df),* sig.	21.99 (11), p = .024				19.02 (11), p = .060			
	Häufigkeiten		Statistiken		Häufigkeiten		Statistiken	
Konfiguration	o	e	z	*p(z)*	o	e	z	*p(z)*
2121	5	9.56	-1.615	.053	6	7.26	-0.332	.370
2122	7	8.91	-0.576	.282	4	7.05	-1.133	.129
2111	5	7.11	-0.712	.238	7	8.84	-0.549	.292

Legende: erste Stelle der Konfiguration ist (2) negative Attribution zum ersten (bzw. zweiten) MZP; zweite Stelle die positive Attribution zum zweiten (bzw. dritten) MZP; dritte Stelle beinhaltet die ST-Durchführungshäufigkeit (1 = gar nicht bis selten; 2 = mittel oft bis sehr oft); vierte Stelle ist die Trackzugehörigkeit (1 = Track A; attributionales Feedback wurde dargeboten; 2 = Track B; attributionales Feedback wurde nicht dargeboten)

Wie Tabelle 35 und 36 zeigen, gibt es bei den Erfolgs- und bei den Misserfolgsattributionen für das WS und für das SS keine überzufällige Besetzung der Konfiguration 2121. Dies bedeutet, es konnte keine Verbesserung der Attributionen durch das schriftliche attributionale Feedback nachgewiesen werden, weder im Wintersemester noch im Sommersemester.

2) Bewirkt das schriftliche attributionale Feedback, welches den Studierenden der Versuchsgruppe (Track A) dargeboten wurde, nicht aber der Kontrollgruppe (Track B), eine positive Veränderung in den Konstrukten (b) Implizite Persönlichkeitstheorien (IPT), (c) Selbstwirksamkeitserwartungen, (d) Selbstkonzept Lerner/in, (e) Interesse und (f) Hilflosigkeit?

Diese Fragestellung wurde spezifiziert: Gibt es einen Unterschied zwischen den vier ST-Häufigkeits-Typen des Track A in der Veränderung der oben genannten Konstrukten, dahingehend, dass Studierende, welche die Selbsttests seltener durchgeführt haben auch niedrigere positive Veränderungen aufweisen als Studierende, die die Selbsttests öfters durchgeführt haben? Und gibt es diese Unterschiede nicht zwischen den vier ST-Häufigkeits-Typen des Track B? Demzufolge wird als Erfolg des attributionalen Feedbacks gewertet, wenn es einen signifikanten Haupteffekt der ST-Häufigkeits-Typen und einen signifikanten Interaktionseffekt zwischen ST-Häufigkeits-Typen und Trackzugehörigkeit gibt. Dies bedeutet, dass es bei Studierenden, welche die Selbsttests seltener durchgeführt haben, keine Unterschiede zwischen Track A und Track B gibt, während es bei den Studierenden, welche die Selbsttests öfters durchgeführt haben, die Versuchsgruppe größer positive Veränderungen in den oben genannten Skalen aufweist als die Kontrollgruppe.

Die *Voraussetzungsprüfungen der Normalverteilung* zeigte, dass die Werte in den Skalen Hilflosigkeit (MZP 1), Interesse (MZP 2 und 3) und Selbstwirksamkeit (MZP 3) in etwa normalverteilt sind, während bei allen anderen Skalen bzw. Messzeitpunkten die Voraussetzung der Normalverteilung nicht gegeben ist. Die Prüfung der *Voraussetzung der Varianzenhomogenität* wurde mittels Levene-Tests überprüft und ist bei allen Skalen und Messzeitpunkten gegeben (kleinstes p bei Skala Selbstwirksamkeit (MZP 3): $F(7;160) = 1.89$, $p = .075$).

Die Fragestellungen 2b) bis 2f) wurden mittels zehn univariater Varianzanalysen mit Messwiederholung geprüft. Die Ergebnisse sind in Tabelle 37 zusammengefasst.

Tabellen 37: Ergebnisse der 10 univariaten VA mit Messwiederholung zu den Fragestellungen 2b) bis 2f)

		MZP 1 – MZP 2		MZP 2 – MZP 3	
		HT WS	**Int.Akt.**	**HT SS**	**Int.Akt.**
2b) Implizite Persönlich-keitstheorien (IPT)	*df*	3;158	3;158	3;160	3;160
	F	1.570	0.061	0.234	1.463
	p	**.199**	**.980**	**.873**	**.227**
2c) Selbstwirksamkeits-erwartungen	*df*	3;160	3;160	3;160	3;160
	F	1.931	0.222	0.271	1.781
	p	**.127**	**.881**	**.846**	**.153**
2d) Selbstkonzept Lerner/in	*df*	3;159	3;159	3;159	3;159
	F	0.376	0.536	1.804	0.623
	p	**.771**	**.658**	**.149**	**.601**
2e) Interesse	*df*	3;160	3;160	3;160	3;160
	F	1.041	1.634	1.729	0.607
	p	**.376**	**.184**	**.163**	**.612**
2f) Hilflosigkeit	*df*	3;158	3;158	3;160	3;160
	F	1.663	1.032	0.796	1.579
	p	**.177**	**.380**	**.498**	**.196**

Legende: HT WS = Häufigkeits-Typen des WS; HT SS = Häufigkeits-Typen des SS; Int.Akt. = Interaktion Häufigkeits-Typen und Trackzugehörigkeit

Wie in Tabelle 37 zu erkennen ist, gibt es von MZP 1 auf MZP 2 und von MZP 2 auf MZP 3 in den Skalen “Implizite Persönlichkeitstheorien“ (IPT), “Selbstwirksamkeits-erwartungen“, “Selbstkonzept Lerner/in“, “Interesse“ und “Hilflosigkeit“ keine signifikanten Haupteffekte der ST-Häufigkeits-Typen sowie auch keine signifikanten Interaktionseffekte zwischen den ST-Häufigkeits-Typen und der Trackzugehörigkeit. Dies bedeutet, dass keine Verbesserung in den oben genannten Skalen durch das attributionale Feedback nachgewiesen werden konnte.

2.3.8 Zusätzliche Ergebnisse zu Ziel 2

Spontane Äußerungen von Studierenden zum Selbsttest bzw. dem schriftlichen attributionalen Feedback

An dieser Stelle sollen nun einige Aussagen von Studierenden wiedergegeben werden, welche die Autorin der Diplomarbeit während informeller Gespräche mit Studierenden in Bezug auf die Selbsttests und das schriftliche attributionale Feedback erzählt bekommen hat. Diese enthalten relevante Zusatzinformationen, welche für die Interpretation und Generalisierung der Ergebnisse wichtig erscheinen.

Eine Studierende berichtete, dass sie die Selbsttests relativ regelmäßig und gewissenhaft ausgefüllt habe. Sie habe den jeweiligen Selbsttest noch im Modul oder spätestens ein paar Tage nach Modulende bearbeitet. Dabei habe sie immer alle Fragen ausgefüllt, d.h. bei jedem neuen Versuch habe sie sowohl die schon richtig gelösten als auch die noch nicht richtig gelösten Fragen ausgefüllt und zur Auswertung abgeschickt. Bei falschen Antworten habe sie zunächst nochmals in der Literatur nachgelesen, oder sie habe die Kolleg/inn/en aus ihrer Gruppen befragt, und/oder das Diskussionsforum genutzt (denn z.B. gegen Ende des Studienjahres hin wurden von mehreren Studierenden ziemlich bald nach Freigabe des jeweiligen Selbsttests die richtigen Antworten im Diskussionsforum gepostet). Die Studierende habe jeden einzelnen Selbsttest so oft durchgeführt bis schlussendlich alle Antworten richtig waren, dann habe sie den jeweiligen Selbsttests mit den richtigen Lösungen ausgedruckt und in einer Lernmappe archiviert. Diesen Ausdruck habe sie dann vor der Prüfung nochmals durchgearbeitet. Auf Nachfrage gab die Studentin an, dass ihr das schriftliche attributionale Feedback aufgefallen war, jedoch sei ihr nicht bewusst gewesen, dass aufwendige theoretische Überlegungen hinter dieser schriftlichen Rückmeldung standen.

In einem anderen Gespräch mit einem Studenten erfuhr die Autorin, dass dieser nicht immer alle Fragen ausfüllte, wie es die Intention der Selbsttests war, sondern nur noch jene Fragen, bei welchen ihm die korrekten Antworten noch fehlten. Bei der Auswertung der Selbsttests wurde bei jeder Frage separat angeben, ob diese richtig oder falsch beantwortet wurde.

Wieder ein anderer Student erzählte, dass es für ihn eine Herausforderung war die richtigen Lösungen zu finden, jedoch löste er fehlende Multiple-Choice-Fragen nicht mittels Literaturstudium, sondern mittels systematischen Durchprobierens der verschiedenen Kombinationen der Antwortoptionen.

Zwei weitere Studierende, welche den Track A absolviert hatten, berichteten, dass ihnen zwar die Rückmeldung bzgl. gelöster Aufgaben in Erinnerung geblieben sei, sie konnten sich aber nicht an eine sprachliche Mitteilung erinnern. Während ein anderer Student, ebenfalls aus dem Track A, angab, dass ihm die Rückmeldung bei wenig gelösten Aufgaben eher “nervte“ als motivierte.

Diese unsystematisch erfassten und hier wiedergegebenen Aussagen von Studierenden erheben keinen Anspruch auf Vollständigkeit oder systematische Datenerhebung, sondern sollen lediglich einen Einblick geben, wie unterschiedlich die Selbsttests von den Studierenden genutzt wurden, und dass das schriftliche attributionale Feedback nicht zwingend von allen Studierenden des Track A wahrgenommen wurde.

3 Zusammenfassung und Diskussion

Wie aus der Forschungsliteratur hervorgeht, ist es wichtig, E-Learning-Angebote zu evaluieren (vgl. Kokavecz et al., 1999; Niegemann et al., 2004; Owston, 2000; Tergan, 2004). So wurde auch im Projekt “Vienna E-Lecturing“ (VEL) eine komplexe Evaluationsforschung realisiert. Die vorliegende Diplomarbeit wurde im Rahmen dieser umfassenden Evaluation durchgeführt, wobei der Focus auf der Analyse der Nutzung der Selbsttests und der Evaluation des web-basierten attributionalen Feedbacks lag.

Das Trainingsprogramm “Vienna E-Lecturing“ (VEL) wurde im Arbeitsbereich Bildungspsychologie und Evaluation an der Fakultät für Psychologie in Wien entwickelt. Es handelt sich hierbei um ein Blended-Learning-Konzept, welches im Rahmen der Pflichtlehrveranstaltung “Forschungsmethoden und Evaluation“ (FM&E) im Diplomstudium Psychologie an der Universität Wien umgesetzt wurde. Die vier Lehrziele von VEL sind die Vermittlung von Fachwissen, Lernkompetenz, Kompetenz zum Kollaborativen Lernen und E-Kompetenz. Das didaktische Konzept des Trainingsprogramms VEL beruht auf zwei zentralen Prinzipien, dem Prinzip der Vernetzung und dem Prinzip einer möglichst optimalen Instruktion (vgl. Wagner et al., 2007). VEL beinhaltet neben Präsenzeinheiten zehn Online-Module. Diese Online-Module wurden von den Studierenden auf einer passwortgeschützten Lernplattform bearbeitet, nachdem sie aufeinander folgend zu feststehenden Terminen freigegeben wurden. Ein wesentlicher Bestandteil dieser Online-Module waren die Selbsttests, welche die Studierenden nach Freigabe zu ebenfalls feststehenden Terminen als freiwillige Lernkontrolle jederzeit und so oft sie wollten online nutzen konnten. Die zehn Selbsttests waren in HTML und PHP programmiert und umfassten sechs bis elf Aufgaben im Multiple-Choice-Format. Die Studierenden sollten die richtige/n Antworten aus einer bestimmten Anzahl von Antwort-möglichkeiten auswählen. Nach Eingabe des Codes bzw. der Matrikelnummer und nach einem Mausklick auf den Button “Auswerten“ wurden die Daten der Selbsttest-nutzung (wie z.B. das Datum der Selbsttestdurchführung und die Anzahl der richtig gelösten Aufgaben) der Studierenden in einer MySQL-Datenbank abgespeichert. Parallel erfolgten mittels PHP-Skript die Auswertung des jeweiligen Selbsttests und die anschließende Anzeige des automatisch erstellten Feedbacks. Dieses unmittel-

bare Feedback beinhaltete für beide Varianten des Trainingsprogramms (Track A und Track B) einerseits die in der jeweiligen Selbsttestdurchführung erreichte Punktezahl (*Knowledge of Performance*) und andererseits eine Auflistung, welche Fragen richtig und welche falsch beantwortet wurden (*Knowledge of Result*). Die Studierenden der "Intensiv-Variante" (Track A) erhielten zusätzlich zur oben beschriebenen Rückmeldung ein ergebnisabhängiges individuelles attributionales Feedback, welches ebenso im Arbeitsbereich Bildungspsychologie und Evaluation entwickelt wurde. Die Grundlagen für dieses attributionale Feedback lieferten nach Schober und Spiel (2005) das Reattributionstraining von Ziegler und Schober (1997, 2001) und die Ergebnisse von Dresel (2004). Bei weniger als 40% richtig gelöster Aufgaben wurden Rückmeldungen zu Misserfolg mit internal-variablen Attributionen gegeben, d.h. Attributionen auf die Anstrengung und dem Hinweis auf das Fähigkeitspotential und teilweise auf die Aufgabenschwierigkeit (external-stabile Attribution). Bei 40-69% richtig gelöster Aufgaben wurde ein Feedback zu tendenziellem Erfolg mit wiederum Attributionen auf die Anstrengung und dem Hinweis auf das Fähigkeitspotential sowie zum Teil auf die Aufgabenschwierigkeit gegeben. Bei 70-89% richtig gelöster Aufgaben wurden Rückmeldungen zu Erfolg sowie eine randomisierte Vorgabe von Attributionen auf die Anstrengung (internal-variabel) und auf die Fähigkeit (internal-stabil) vermittelt. Bei mehr als 90% richtig gelöster Aufgaben erfolgte ein Feedback zu großem Erfolg mit wiederum randmomisierter Vorgabe von internalen Attributionen (d.h. auf die Anstrengung und auf die Fähigkeit).

Bei der vorliegenden Forschungsarbeit handelt es sich um eine quasi-experimentelle Untersuchung mit einem Zwei-Gruppen-Prätest-Posttest1-Posttest2-Design. Für die Analyse der Selbsttestnutzung und der Evaluation des web-basierten attributionalen Feedbacks wurden einerseits die Nutzungsdaten der zehn Selbsttests, welche in einer MySQL-Datenbank gespeichert waren, aufbereitet und komprimiert. Andererseits wurden die im "SRL-Fragebogen" erhobenen Daten (z.B. die erst-genannten Erfolgs- bzw. Misserfolgsattributionen) der Studierenden und deren erbrachte Leistung in den Vorlesungsprüfungen zu "Forschungsmethoden und Evaluation I bzw. II" herangezogen.

Für die Beantwortung der Fragestellungen wurden zwei verschiedene Stichproben herangezogen. Die Stichprobe 1, deren Daten für die Darstellung des Verlaufs der Selbsttestnutzung verwendet wurden, umfasst die Studierenden des Studienjahres 2003/04, die an den beiden Lehrveranstaltungen (FM&E I und FM&E II) und an allen drei Messzeitpunkten der Evaluation teilgenommen haben, sowie bei welchen der Code und die Matrikelnummer zusammengeführt werden konnte. Die Stichprobe 2, deren Daten für die Beantwortung des Großteils der Fragestellungen herangezogen wurde, wurde im Arbeitsbereich Bildungspsychologie und Evaluation mittels eines neu entwickelten Matching-Verfahrens gezogen. Die Versuchsgruppe umfasst 84 Studierende der “Intensiv-Variante“ (Track A) des Trainingsprogramms VEL aus den Studienjahren 2003/04 und 2004/05. Die Kontrollgruppe besteht aus 84 Studierenden, die aus dem Pool der Studierenden der “Standard-Variante“ (Track B) des Studienjahres 2003/04 hinsichtlich der Ähnlichkeit zur Versuchsgruppe mit Hilfe der Euclidean Distance Matching (EuM) Prozedur gezogen wurden (vgl. Spiel et al., 2008).

Da sich die Matching-Stichprobe aus zwei Kohorten, d.h. aus den Studienjahren 2003/04 und 2004/05, zusammensetzt, wurde vor der eigentlichen Untersuchung anhand einer Gesamtstichprobe überprüft, ob es Unterschiede in der Durchführungshäufigkeit der Selbsttests zwischen den beiden Studienjahren gibt. Die Ergebnisse zeigen, dass es keine signifikanten Unterschiede zwischen den beiden Studienjahren (2003/04 und 2004/05) gibt, und auch nicht zwischen den beiden Studienjahren innerhalb der beiden Versuchsbedingungen (Track A und Track B).

Das *erste Ziel* dieser Diplomarbeit war die *Analyse der Selbsttestnutzung* und, ob es einen *Zusammenhang zwischen Selbsttestnutzung und der Prüfungsleistung* in “Forschungsmethoden und Evaluation“ gibt.

Für die deskriptive Darstellung der Durchführungshäufigkeit der zehn Selbsttests und des Zeitpunkts der Selbsttestdurchführung wurden die Daten der Stichprobe 1 (N = 226) herangezogen. Die zehn Selbsttests (ST) wurden im Studienjahr 2003/04 von den 226 Studierenden der Stichprobe 1 insgesamt 14 806-mal durchgeführt. Davon wurden die Selbsttests des Wintersemesters (ST 1-5) 8 142-mal (55.0%) und die Selbsttests des Sommersemesters (ST 6-10) 6 664-mal (45.0%) absolviert. Die

Selbsttests wurden einerseits während des jeweiligen Moduls und andererseits ab ca. 4 Wochen vor dem jeweiligen Prüfungstermin genutzt. Wobei diese Trennung bei Selbsttest 5 und 10 nicht möglich ist, da die Zeiträume während des Moduls und vier Wochen vor dem ersten Prüfungstermin deckungsgleich sind. Für alle folgenden Ergebnisse wurden die Daten der Matching-Stichprobe (*N* = 168) herangezogen.

Zunächst wurde überprüft, ob es Unterschiede in der Durchführungshäufigkeit zwischen den einzelnen Selbsttests, den Semestern sowie den beiden Versuchsbedingungen gibt. Obwohl aufgrund bisheriger Ergebnisse aus der Forschungsliteratur (vgl. Doherty, 2006; Liu et al., 2007) und der Evaluation zu VEL (z.B. Schober & Spiel, 2005) Unterschiede erwartet wurden, zeigen die Ergebnisse, dass nur Selbsttest 7 signifikant häufiger durchgeführt worden ist als der Selbsttest 9. Alle anderen Selbsttests unterscheiden sich nicht signifikant in ihrer Durchführungshäufigkeit. Es gibt auch keine statistisch bedeutsamen Unterschiede in der Durchführungshäufigkeit der Selbsttests des Wintersemesters (WS) und der Durchführungshäufigkeit der Selbsttests des Sommersemesters (SS). Weiters gibt es auch keine signifikanten Unterschiede zwischen den Studierenden des Track A und den Studierenden des Track B in der Durchführungshäufigkeit – weder zusammengefasst für das ganze Studienjahr, noch für die beiden Semester (WS und SS), noch für die einzelnen zehn Selbsttests.

Die signifikanten Unterschiede in der Durchführungshäufigkeit zwischen Selbsttest 7 und Selbsttest 9 könnten möglicherweise dadurch bedingt sein, dass Selbsttest 7, im Vergleich zu den anderen Selbsttests, die meisten Aufgaben (11) beinhaltete und im Modul 7 englische Literatur bearbeitet werden musste. Das Modul 7 war auch bei der Bearbeitung insgesamt viel aufwendiger als die anderen Module (vgl. Kapitel 1.4.2 – Evaluationsergebnisse zu VEL). Und obwohl Selbsttest 9 auch relativ viele Aufgaben (10) beinhaltete, bezogen sich diese nicht unmittelbar auf das soeben bearbeitete Modul 9, sondern waren aus den Stoffgebieten der zuvor bearbeiteten Module (1-8) entnommen. Der Selbsttest 9 stellte sozusagen eine Wiederholung der vorhergehenden Module dar.

Der Frage nach unterschiedlichen Typen der Selbsttestnutzung wurde zunächst mit Hilfe der Clusteranalyse nachgegangen. Da diese jedoch keine zufriedenstellenden

Cluster lieferte, wurden mittels theoriegeleitetem Wissen und anhand von Ergebnissen aus der Forschungsliteratur zu Lerngewohnheiten (vgl. Haag & Mischo, 2002; Wagner et al., 2004) vier Typen der Selbsttestnutzung erstellt. Demzufolge waren im Wintersemester 37% der Studierenden der Matching-Stichprobe ST-Nicht-Nutzer/inn/en (d.h. sie haben im Semester und vor den Prüfungen die Selbsttests gar nicht bis selten verwendet) und 30% waren Saisonnutzer/inn/en (d.h. sie führten die Selbsttests während des Semesters gar nicht bis selten durch, nutzten diese jedoch vor den Prüfungen mittel bis sehr oft). Weiters waren im Wintersemester 20% der Studierenden der Matching-Stichprobe Mitlerner/inn/en (d.h. sie verwendeten die Selbsttests während des Semesters mittel bis sehr oft, jedoch vor den Prüfungen gar nicht bis selten) und 13% waren ST-Viel-Nutzer/inn/en (d.h. sie nutzten die Selbsttests im Semester und vor den Prüfungen mittel bis sehr oft). Im Sommersemester waren 46% der Studierenden der Matching-Stichprobe ST-Nicht-Nutzer/inn/en und 42% Saisonnutzer/inn/en, während nur noch 5% Mitlerner/inn/en und 7% ST-Viel-Nutzer/inn/en waren.

Entgegen den Erwartungen aufgrund der Forschungsliteratur (z.B. Liu et al., 2007) gibt es keinen signifikanten Unterschied zwischen den Studierenden des Track A und den Studierenden des Track B in der Art der Selbsttestnutzung, d.h. es gibt keinen statistisch bedeutsamen Unterschied zwischen Track A und Track B hinsichtlich der Zugehörigkeit zu einem der vier ST-Nutzungs-Typen.

Die Überprüfung der Fragestellung, ob es aufgrund unterschiedlicher Selbsttest-nutzung einen Unterschied in der Prüfungsleistung zu FM&E I bzw. II gibt, zeigte, dass es einen signifikanten Unterschied zwischen den ST-Viel-Nutzer/inne/n und den ST-Nicht-Nutzer/inne/n in der Prüfungsleistung zu FM&E II gibt. ST-Viel-Nutzer/innen haben eine signifikant höhere Gesamtleistung in den Prüfungen zu FM&E II als ST-Nicht-Nutzer/innen. Ansonst konnten keine statistisch bedeutsamen Unterschiede zwischen den vier ST-Nutzungs-Typen festgestellt werden, weder in den Prüfungen zu FM&E I *gesamt*, noch in den einzelnen Kategorien *Reproduktion, Verständnis* und *Produktion* der Prüfungen FM&E I und II. So konnte nicht bestätigt werden, wie aufgrund der Ergebnisse aus der Literatur erwartet wurde (vgl. Haag & Mischo, 2002; Wagner et al., 2004), dass Saisonslerner/innen niedrigere Prüfungsleistungen erzielen als andere Typen der Selbsttestnutzung.

Ferner wurden keine signifikanten Interaktionseffekte zwischen den vier ST-Nutzungs-Typen und der Trackzugehörigkeit festgestellt, mit Ausnahme der Kategorie Verständnis bei den Prüfungen FM&E II[27].

Um die signifikanten Unterschiede zwischen den ST-Nicht-Nutzer/innen und ST-Viel-Nutzer/innen in der Prüfungsleistung gesamt in FM&E II bzw. die signifikante Interaktion in der Kategorie Verständnis der Prüfungen zu FM&E II abzusichern, wurden weitere inferenzstatistische Analysen durchgeführt. Hierfür wurden die Studierenden hinsichtlich ihrer ST-Durchführungshäufigkeit im WS bzw. im SS in je vier Gruppen unterteilt (ST-Durchführungs-Häufigkeits-Typen). Für diese Unterteilung wurde die ST-Durchführungshäufigkeit herangezogen, unabhängig davon, wann die Selbsttests genutzt wurden, d.h. der Zeitpunkt der Selbsttestdurchführung fand hierbei keine Berücksichtigung.

Die *Zusatzergebnisse* zeigen, dass es in der Prüfungsleistung *gesamt* zu FM&E II und in den Kategorien *Reproduktion* und *Produktion* statistisch und praktisch bedeutsame Unterschiede zwischen Studierenden, welche die Selbsttests gar nicht bis sehr selten durchgeführt haben, und Studierenden, die die Selbsttests oft bis sehr oft durchgeführt haben, gibt. Unabhängig jedoch von ihrer Trackzugehörigkeit, denn es wurden keine signifikanten Interaktionseffekte festgestellt (auch nicht in der Kategorie Verständnis). Diese Ergebnisse stehen im Einklang mit der Literatur (Brown, 1932; Symonds & Chase, 1929), dass das Feedback *Knowledge of Performance* (KP) lernförderlicher ist als kein Feedback. Jedoch ist diese Interpretation der Ergebnisse nur eingeschränkt gültig, weil es sich bei der vorliegenden Untersuchung nicht um ein Experiment, sondern um ein Quasi-Experiment handelt und somit nicht mitberücksichtigte Störvariablen einen Einfluss auf die Prüfungsleistungen haben könnten. Das vorliegende Ergebnis soll jedoch als Anregung für künftige Untersuchungen dienen. So könnte beispielsweise für die Überprüfung der Wirksamkeit der Selbsttests und der Feedback-Arten *Knowledge of Performance* (KP) und *Knowledge of Result (KR)* ein experimentelles Untersuchungsdesign angelegt werden (vgl. Narciss, 2006).

[27] Dieses Ergebnis wird aber angesichts der sehr ungleichen und teilweise extrem kleinen Zellenbesetzungen in Frage gestellt und nicht interpretiert.

Das *zweite Ziel* der Diplomarbeit war die *Evaluation des web-basierten attributionalen Feedbacks*.

Durch die vorliegende Untersuchung konnten keine Verbesserungen in den Erfolgs- und den Misserfolgsattributionen durch das schriftliche attributionale Feedback nachgewiesen werden, weder im Wintersemester noch im Sommersemester. Weiters konnten auch keine Verbesserungen in den Skalen “Implizite Persönlichkeitstheorien“ (IPT), “Selbstwirksamkeitserwartungen“, “Selbstkonzept Lerner/in“, “Interesse“ und “Hilflosigkeit“ durch das attributionale Feedback festgestellt werden, weder von MZP 1 auf MZP 2 noch von MZP 2 auf MZP 3. Diese Ergebnisse der vorliegenden Untersuchung stehen im Widerspruch zu den Ergebnissen aus der Forschungsliteratur zu Attributionstrainings (vgl. Dresel, 2004; Grünke & Castello, 2008; Heller, 2004; Robertson, 2000; Ziegler & Finsterwald, 2008; Ziegler & Schober, 2001).

Als *Zusatzergebnisse* zum zweiten Ziel der Diplomarbeit wurden einige Aussagen von Studierenden wiedergegeben, welche während informeller Gespräche mit Studierenden in Bezug auf die Selbsttests und das schriftliche attributionale Feedback festgehalten wurden. Diese unsystematisch erfassten Aussagen von Studierenden gewähren einen Einblick, wie unterschiedlich die Selbsttests von den Studierenden genutzt wurden, und dass das web-basierte attributionale Feedback nicht zwingend von allen Studierenden des Track A wahrgenommen wurde. Wenn man diese zusätzlich festgehaltenen, spontanen Äußerungen der Studierenden für die Interpretation heranzieht, kann man als mögliche Ursachen für die fehlende Wirksamkeit des web-basierten attributionalen Feedbacks eine *mangelnde Nutzung des Feedbacks durch die Lernenden* und/oder eine *geringe “Mindfulness“ bei der Feedbackrezeption* festhalten (vgl. Bangert-Drowns et al., 1991; Krause, 2007). So wird für nachfolgende quasi-experimentelle Untersuchungen zu web-basierten attributionalen Feedbacks ein *Manipulation Check* (siehe Bortz & Döring, 2006) empfohlen, um empirisch zu prüfen, ob das attributionale Feedback in der Versuchsgruppe wahrgenommen wird. Denn wo eine Intervention nicht greift, kann sie auch nichts bewirken (in Anlehnung an Bortz & Döring, 2006). Sollte sich anhand des Manipulation Checks zeigen, dass das schriftliche attributionale Feedback im Selbsttest zu wenig präsent ist, wäre eine Möglichkeit zur Verbesserung der

Wahrnehmbarkeit die Präsentation des attributionalen Feedbacks in einem eigenen Pop-Up-Fenster (vgl. Dresel, 2004), anstelle der gegenwärtig realisierten, gleichzeitigen Darbietung mit dem Feedback *Knowledge of Performance* (KP).

Einen weiteren Erklärungsansatz für die nicht hypothesenkonformen Ergebnisse zum attributionalen Feedback bieten die Ergebnisse von Ziegler und Finsterwald (2008), wonach sich erfolgreiche Attributionstrainings unter anderem dadurch auszeichnen, dass eine Vorauswahl der Trainingsteilnehmer/innen getroffen wird. Denn eine undifferenzierte Applizierung des Attributionstrainings schmälert dessen Wirkung. Diese Anforderung konnte beim Trainingsprogramm VEL nicht bewerkstelligt werden, da die Studierenden selbst entscheiden konnten, an welcher Variante sie teilnehmen wollten. Die deskriptiven Ergebnisse der vorliegenden Untersuchung zeigen, dass 80% der Studierenden der Versuchsgruppe (sowie auch 80% der Kontrollgruppe) schon beim ersten Messzeitpunkt “positive“ Erfolgsattributionen angaben. Bezüglich eines Misserfolgs nannten 62% der Studierenden der Versuchsgruppe (bzw. 56% der Kontrollgruppe) “positive“ Attributionen schon beim ersten Messzeitpunkt.

Eine weitere mögliche Ursache für die unerwartet ausgebliebenen Erfolge des web-basierten attributionalen Feedbacks könnte eine fehlende Veridikalität der Attributionsrückmeldungen (vgl. Dresel, 2004; Ziegler & Schober, 2001) sein. Möglicherweise sind die attributionalen Rückmeldungen von den Studierenden als wirklichkeitsfremd und zu unpersönlich aufgefasst worden (Hinweise dafür siehe Schober & Spiel, 2005). Auch dies sollte bei zukünftigen Untersuchungen zu web-basierten attributionalen Rückmeldungen berücksichtigt bzw. erforscht werden.

4 Literaturverzeichnis

Arnold, P. & Thillosen, A. (2002). Aufgabenorientiertes Lernen in telematischen Studienmodulen – Aufgabenformen, Aufgabentypen und Aufgabengestaltung. In G. Zimmer (Hrsg.), *E-Learning: High-Tech or High-Teach? Lernen in Netzen zwischen Aktualität und Potenzialität* (S. 5-18). Bielefeld: Bertelsmann.

ATLAS Gesellschaft für Internationale Kommunikation. (2003). *Computer-Lexikon mit Fachwörterbuch* (7. Aufl.). Unterschleißheim: Microsoft Press.

Atria, M., Reimann, R. & Spiel, C. (2006). Qualitätssicherung durch Evaluation. In C. Steinebach (Hrsg.), *Handbuch Psychologische Beratung* (S. 574-586). Stuttgart: Klett-Cotta.

Baeßler, B., Lücke, S., Koring, B., Kinnebrock, S. & Rössler, P. (2003). E-Learning-Systeme: Theoriegeleitete Konzeption, Qualitätsmanagement, Implementierung. *Zeitschrift für Medienpsychologie, 15* (1), 13-23.

Bangert-Drowns, R.L., Kulik, C.C., Kulik, J.A. & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61* (2), 213-238.

Bloom, B.S. (1956). *Taxonomy of educational objectives. Handbook 1. The cognitive domain.* New York: David McKay.

Bortz, J. & Döring, N. (2002). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (3. Aufl.). Berlin: Springer.

Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. Aufl.). Berlin: Springer.

Breitkopf, L. (1985). Die Hilflosigkeitsskala. *Diagnostica, 31,* 221-233.

Brown, F.J. (1932). Knowledge of results as an incentive in school room practice. *Journal of Educational Psychology, 23,* 532-552.

Bruns, B. & Gajewski, P. (2002). *Multimediales Lernen im Netz. Leitfaden für Entscheider und Planer* (3. Aufl.). Berlin: Springer.

Carell, A. (2006). *Selbststeuerung und Partizipation beim computergestützten kollaborativen Lernen. Eine Analyse im Kontext hochschulischer Lernprozesse.* Münster: Waxmann.

Carver, C.S. & Scheier, M.F. (2000). On the structure of behavioral self-regulation. In M. Boekaerts, P.R. Pintrich & M. Zeidner (Hrsg.), *Handbook of self-regulation* (S. 41-84). San Diego, CA: Academic Press.

Deci, E.L. & Ryan, R.M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik, 39* (2), 223-238.

Deutsche Gesellschaft für Evaluation. (Hrsg.). (2001). *Standards für Evaluation.* Köln: Deutsch Gesellschaft für Evaluation e.V.

Dresel, M. (2004). *Motivationsförderung im Schulischen Kontext*. Göttingen: Hogrefe.

Dresel, M., Schober, B. & Ziegler, A. (2007). Golem und Pygmalion: Scheitert die Chancengleichheit von Mädchen im mathematisch-naturwissenschaftlichen Bereich am geschlechtsstereotypen Denken der Eltern? In P. Ludwig & H. Ludwig (Hrsg.), *Erwartungen in himmelblau und rosarot – Effekte, Determinanten und Konsequenzen von Geschlechterdifferenzen in der Schule* (S. 61-81). Weinheim: Juventa.

Doherty, W. (2006). An analysis of multiple factors affecting retention in web-based community college courses. *Internet and Higher Education, 9,* 245-255.

Dweck, C.S. (1999). *Self-theories: Their role in motivation, personality, and development.* Philadelphia: Taylor & Francis.

Fengler, J. (2004). *Feedback geben: Strategien und Übungen* (3. Aufl.). Weinheim: Beltz.

Finsterwald, M. (1999). *Die Wirksamkeit von Reattributionstrainings zur Verbesserung der Situation begabter Mädchen und Frauen: Ergebnisse einer Metaanalyse.* Unveröffentlichte Magisterarbeit, Ludwig-Maximilians-Universität München.

Finsterwald, M. (2005). Reattributionstrainings: Eine Chance für eine spezifische Förderung von Mädchen im MINT-Bereich? In Österreichisches Zentrum für Begabungsförderung und Begabungsforschung. (Hrsg.), *Die Forscher/innen von morgen. Kongressbericht des 4. Internationalen Begabtenkongresses in Salzburg* (S. 145-152). Innsbruck: Studienverlag.

Friedrich, H.F. & Mandl, H. (1997). Analyse und Förderung selbstgesteuerten Lernens. In F.E. Weinert, (Hrsg.), *Pädagogische Psychologie* (Enzyklopädie der Psychologie, Serie I, Themenbereich D, Praxisgebiete, Bd. 4, S. 237-293). Göttingen: Hogrefe.

Gagné, R.M. (1985). *The conditions of learning and theory of instruction.* New York: Holt, Rinehart & Winston.

Gerstenmaier, J. & Mandl H. (1995). Wissenserwerb unter konstruktivistischer Perspektive. *Zeitschrift für Pädagogik, 41* (6), 867-888.

Gordijn, J. & Nijhof, W.J. (2002). Effects of complex feedback on computer-assisted modular instruction. *Computers & Education, 39,* 183-200.

Grünberger, N. (2005). *Individuelles und kollektives Lernen in Arbeitsgruppen.* Unveröffentlichte Diplomarbeit, Universität Wien.

Grünke, M. & Castello, A. (2004). Attributionstraining. In G.W. Lauth, M. Grünke & J.C. Brunstein (Hrsg.), *Interventionen bei Lernstörungen. Förderung, Training und Therapie in der Praxis* (S. 382-390). Göttingen: Hogrefe.

Haag, L. & Mischo, C. (2002). „Saisonarbeiter" in der Schule – einem Phänomen auf der Spur. *Zeitschrift für Pädagogische Psychologie, 16* (2), 109-115.

Hattie, J. (1999). *Influences on student learning.* Zugriff am 28.07.2008, von http://www.arts.auckland.ac.nz/FileGet.cfm?ID=2d01d774-9d89-4351-a4f6-afee19a2ce2e

Hattie J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77* (1), 81-112.

Heider, F. (1944). Social perception and phenomenal causality. *Psychological Review, 51,* 358-374.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Heller, K.A. (2004). Reattributionstraining (RAT) – ein unterrichtsintegriertes Modell der Begabtenförderung in mathematisch-naturwissenschaftlichen Fächern. In C. Fischer, F.J. Mönks & E. Grindel (Hrsg.), *Curriculum und Didaktik der Begabtenförderung* (S. 304-329). Münster: LIT Verlag.

Howitt, D. & Cramer, D. (2000). *First steps in research and statistics. A practical workbook for psychology students.* London: Routledge.

Ihringer, S. (2000). E-Learning: Modetrend oder neue Potenziale netzgestützter Medien für das Lernen im öffentlichen Dienst. *SpLITter - IT-Nachrichten der Berliner Verwaltung 10* (2), 7-11.

Immler, C. (2006). *Das PC-Handbuch. Windows, Hardware, Software, Internet, Netzwerk, Multimedia.* München: Markt+Technik Verlag.

Jerusalem, M. & Satow, L. (1999). Schulbezogene Selbstwirksamkeitserwartung. In R. Schwarzer & M. Jerusalem (Hrsg.), *Skalen zur Erfassung von Lehrer- und Schülermerkmalen* (S. 15). Berlin: Freie Universität Berlin.

Kelley, H.H. (1967). Attribution theory in social psychology. In D. Levine (Hrsg.), *Nebraska symposium on motivation* (S. 192-238). Lincoln: University of Nebraska Press.

Kerres, M. (2001). Multimediale und telemediale Lernumgebungen. Konzeption und Entwicklung (2. Aufl.), München: Oldenbourg.

Kerres, M. & Jechle, T. (2002). Didaktisch Konzeption des Telelernens. In L.J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia und Internet. Lehrbuch für Studium und Praxis* (S. 267-281). Weinheim: Beltz Psychologie Verlags Union.

Kirkpatrick, D. (1998). *Evaluating training programs: The four levels* (2. Aufl.). San Francisco, CA: Berrett-Koehler.

Klauer, K.J. (1985). Framework for a theory of teaching. *Teaching and Teacher Education, 1,* 5-17.

Kluger, A.N. & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119* (2), 254-284.

Kokavecz, I., Lammers, F. & Holling, H. (1999). Evaluation von computerunterstützten Lern- und Lehrprojekten. In H. Holling & G. Gediga (Hrsg.), *Evaluationsforschung* (S. 59-72). Göttingen: Hogrefe.

Krapp, A. & Weidenmann, B. (2006). *Pädagogische Psychologie* (5. Aufl.). Weinheim: Beltz Verlag.

Krause U.-M. (2007). *Feedback und kooperatives Lernen.* Münster: Waxmann.

Krause, U.-M., Stark, R. & Mandl, H. (2004). Förderung des computerbasierten Wissenserwerbs durch kooperatives Lernen und eine Feedbackmaßnahme. *Zeitschrift für Pädagogische Psychologie, 18* (2), 125-136.

Kulhavy, R.W. & Stock, W.A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review, 1* (4), 279-308.

Kulhavy, R.W., White, M.T., Topp, D.W., Chan, A.L. & Adams, J. (1985). Feedback complexity and corrective efficiency. *Contemporary Educational Psychology, 10,* 285-291.

Kulik, J.A. & Kulik, C.C. (1988). Timing of feedback and verbal learning. *Review of Educational Research, 58* (1), 79-97.

Liu, S., Gomez, J., Khan, B. & Yen, C.-J. (2007). Toward a learner-oriented community college online course dropout framework. *International Journal on E-Learning, 6* (4), 519-542.

Mandl, H. (2004). E-Learning – Trends und zukünftige Entwicklungen. In K. Rebensburg (Hrsg.), *Grundfragen multimedialen Lehrens und Lernens* (S. 17-29). Norderstedt: Books on Demand GmbH.

Meister, D.M. (2004). Online-Lernen und Weiterbildung. In D.M. Meister (Hrsg.), *Online-Lernen du Weiterbildung* (S. 7-26). Wiesbaden: Verlag für Sozialwissenschaften.

Metzger, C., Weinstein, C.E. & Palmer, D.R. (2002). *Wie lerne ich? Lernstrategieinventar für Studentinnen und Studenten (LASSI)* (6. Aufl.). Aarau: Sauerländer.

Möller, J. (2006). Attributionen. In D.H. Rost (Hrsg.). *Handwörterbuch Pädagogische Psychologie* (3. Aufl.), (S. 34-40). Weinheim: Beltz.

Moreno, R. (2004). Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science, 32* (1), 99-113.

Musch, J. (1999). Die Gestaltung von Feedback in computerunterstützen Lernumgebungen: Modelle und Befunde. *Zeitschrift für Pädagogische Psychologie, 13* (3), 148-160.

Narciss, S. (2006). *Informatives tutorielles Feedback. Entwicklungs- und Evaluationsprinzipien auf der Basis instruktionspsychologischer Erkenntnisse.* Münster: Waxmann.

Niegemann, H.M., Hessel, S., Hochscheid-Mauel, D., Aslanski, K., Deimann, M. & Kreuzberger, G. (2004). *Kompendium E-Learning.* Heidelberg: Springer.

Owston, R.D. (2000). Evaluating web-based learning environments: Strategies and insights. *CyberPsychology and Behavior, 3,* 79-87.

Paulus, C. & Strittmatter, P. (2002). Netzbasiertes Lernen in der Hochschule zwischen Anspruch und Wirklichkeit. *Unterrichtswissenschaft, 30* (4), 290-303.

Perrez, M., Huber, G.I. & Geißler, K.A. (2006). Psychologie der pädagogischen Interaktion. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (5. Aufl.), (S. 357-421). Weinheim: Beltz Verlag.

Rasch, D. & Guiard, V. (2004). The robustness of parametric statistical methods. *Psychology Science, 46* (2), 175-208.

Reinmann, G. & Mandl, H. (2006). Unterrichten und Lernumgebungen gestalten. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (5. Aufl.), (S. 613-658). Weinheim: Beltz Verlag.

Reinmann-Rothmeier, G. (2003). *Didaktische Innovation durch Blended Learning. Leitlinien anhand eines Beispiels aus der Hochschule.* Bern: Verlag Hans Huber.

Reinmann-Rothmeier, G. & Mandl, H. (2002). Analyse und Förderung kooperativen Lernens in netzbasierten Umgebungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 34* (1), 44–57.

Robertson, J.S. (2000). Is attribution training a worthwhile classroom intervention for K–12 students with learning difficulties? *Educational Psychology Review, 12* (1), 111-134.

Schiefele U., Krapp, A., Wild, K.-P. & Winteler, A. (1993). Der „Fragebogen zum Studieninteresse“ (FSI). *Diagnostica, 39* (4), 335-351.

Schmitz, B. (2003). Selbstregulation – Sackgasse oder Weg mit Forschungsperspektive? *Zeitschrift für Pädagogische Psychologie, 17* (3/4), 221-232.

Schmitz, B., Perels, F., Bruder, S. & Otto, B. (2003). *Fragebogen Selbstregulation.* Unveröffentlichter Fragebogen, Institut für Psychologie der Technischen Universität Darmstadt.

Schober, B. (2002). *Entwicklung und Evaluation des Münchner Motivationstrainings (MMT).* Regensburg: Roderer.

Schober, B. & Spiel, C. (2005). E-Lecturing: Self-regulated in Higher Education. *Abschlussbericht für das Jubiläumsfondprojekt der Österreichischen Nationalbank Nr. 9956.* Unveröffentlichter Bericht, Arbeitsbereich Bildungspsychologie und Evaluation der Universität Wien.

Schober, B., Reimann, R., Wagner, P., Atria, M. & Spiel, C. (2006). Das VEL-Programm: systematische Förderung von Selbstregulation und Kooperation in der universitären Lehre. In A. Mettinger, P. Oberhuemer & C. Zwiauer (Hrsg.), *eLearning an der Universität Wien. Forschung-Entwicklung-Einführung* (S. 152-169). Münster: Waxmann.

Schober, B., Wagner, P., Reimann, R., Atria, M. & Spiel, C. (2006). Teaching research methods in an internet-based blended-learning setting. Vienna e-lecturing (VEL). *Methodology, 2* (2), 73-82.

Schüpbach, E., Guggenbühl, U., Krehl, C., Siegenthaler, H. & Kaufmann-Hayoz, R. (2003). *Didaktischer Leitfaden für E-Learning*. Bern: h.e.p.-Verlag.

Schwarzer, R. (1999). Selbstregulation (REG). In R. Schwarzer & M. Jerusalem (Hrsg.), *Skalen zur Erfassung von Lehrer- und Schülermerkmalen*. Berlin: Freie Universität Berlin.

Slavin, R.E. (1993). Kooperatives Lernen und Leistung: Eine empirisch fundierte Theorie. In G. Huber (Hrsg.), *Neue Perspektiven der Kooperation* (S. 151-170). Hohengehren: Schneider Verlag.

Söllner, R., Flöter, M. & Scheibner, N. (2006) E-Learning in der statistischen Grundausbildung in Psychologie: Das Projekt Statistik in Anwendung (SiA) In G. Krampen & H. Zayer (Hrsg.), *Didaktik und Evaluation in der Psychologie* (S. 193-206). Göttingen: Hogrefe.

Spiel, C. (2005). Program evaluation. In C.B. Fisher & R.M. Lerner (Hrsg.), *Applied developmental science: An encycopedia of research, policies, and programs* (S. 879-883). Thousand Oaks, CA: Sage.

Spiel, C. & Strohmeier, D. (2006). Evaluation und Forschungsmethoden. In U. Kastner-Koller & P. Deimann (Hrsg.), *Psychologie als Wissenschaft* (S. 217-237). Wien: WUV.

Spiel, C., Lapka, D., Gradinger, P., Zodlhofer, E.M., Reimann, R., Schober, B., Wagner, P. & von Eye, A. (2008). A euclidean distance-based matching procedure for non-randomized comparison studies. *European Psychologist, 13* (3), 180-187.

Spiel, C., Reimann, R., Wagner, P., Schober, B., Atria, M., Gradinger, P. & Lapka, D. (2007). Selbstreguliertes Lernen bei Studierenden fördern: Theorie, Praxis und Evaluation einer blended Lehr-Lern-Umgebung. In A. Gastager, T. Hascher & H. Schwetz (Hrsg.), *Pädagogisches Handeln: Balancing zwischen Theorie und Praxis. Beiträge zur Wirksamkeitsforschung in pädagogisch-psychologischem Kontext* (S. 175-185). Landau: VEP.

Spiel, C., Strohmeier, D., Faradji, S., Schober, B., Gradinger, P., Zens, B. Aichinger, A. & Reimann, R. (2004). Selbstreguliertes Lernen durch Vienna E-Lecturing (VEL). Konzept, Umsetzung und Evaluation. In W. Fröhlich & W. Jütte (Hrsg.), *Qualitätsentwicklung in der postgradualen Weiterbildung. Internationale Entwicklungen und Perspektiven* (S. 377-388). Münster: Waxmann.

Stiensmeier-Pelster, J. & Heckhausen, H. (2006). Kausalattribution von Verhalten und Leistung. In J. Heckhausen & H. Heckhausen (Hrsg.), *Motivation und Handeln* (3. Aufl.), (S. 355-392). Heidelberg: Springer.

Symonds, P.M. & Chase, D.H. (1929). Practice vs. motivation. *Journal of Educational Psychology, 20,* 19-35.

Tergan, S.O. (2004). Realistische Qualitätsevaluation von E-Learning. In D.M. Meister, S.O. Tergan & P. Zentel (Hrsg.), *Evaluation von E-Learning. Zielrichtungen, methodologische Aspekte, Zukunftsperspektiven* (S. 131-154). Münster: Waxmann.

Vogt, C. (2004). *Informatik. Eine Einführung in Theorie und Praxis.* Heidelberg: Spektrum Akademischer Verlag.

von Eye, A. (2002). *Configural Frequency Analysis – Methods, models, and applications.* Mahwah, NJ: Lawrence Erlbaum Associates.

Wagner, P., Schober, B. & Spiel, C. (2004). Learning types according to time spent in working at home for school, self-concept and other psychological variables. In H.W. Marsh, J. Baumert, G.E. Richards & U. Trautwein (Hrsg.), *Self-Concept, Motivation and Identity. Where to from here? Proceedings of the Third International Biennal SELF research Conference* (S. 930-937). University of Western Sydney: Self Research Center.

Wagner, P., Schober, B., Reimann, R., Atria, M. & Spiel, C. (2007). Vienna-E-Lecturing: Trainingskonzept zum Selbstregulierten Lernen im Studium. In M. Landmann & B. Schmitz (Hrsg.), *Selbstregulation erfolgreich fördern. Praxisnahe Trainingsprogramme für effektives Lernen* (S. 290-311). Stuttgart: Kohlhammer.

Weiner, B. (1994). *Motivationspsychologie* (3. Aufl.). Weinheim: Beltz.

Weinert, F.E. (1982). Selbstgesteuertes Lernen als Voraussetzung, Methode und Ziel des Unterrichts*. Unterrichtswissenschaft, 10* (2), 99-110.

Wild, K.-P. (2000). *Lernstrategien im Studium. Strukturen und Bedingungen.* Münster: Waxmann.

Winkler, P. (2006). *Computer Lexikon 2006.* München: Markt+Technik Verlag.

Wosnitza, M. (2000). *Motiviertes selbstgesteuertes Lernen im Studium.* Landau Verlag Empirische Pädagogik.

Wottawa, H. & Thierau, H. (1998). *Lehrbuch Evaluation* (2. Aufl.). Bern: Huber.

Zentraler Informatikdienst der Universität Wien. (2008a). *Blackboard Vista.* Zugriff am 25.03.2008, von http://www.univie.ac.at/ZID/elearning-vista/

Zentraler Informatikdienst der Universität Wien. (2008b). *Informationsfolder für Studierende.* Zugriff am 25.03.2008, von http://www.univie.ac.at/ZID/elearning/data/ Folder_eLearning_Stud.pdf

Ziegler, A. & Finsterwald, M. (2008). Attributionstrainings. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogische Psychologie* (S. 416-427). Göttingen: Hogrefe.

Ziegler, A. & Schober, B. (1996). Geschlechtsspezifische Unterschiede in mathematikbezogenen Kognitionen zum Zeitpunkt des Gymnasialeintritts. In K.-P. Treumann, G. Neubauer, R. Möller & J. Abel (Hrsg.), *Methoden und Anwendungen empirischer pädagogischer Forschung* (S. 145-153). Münster: Waxmann.

Ziegler, A. & Schober, B. (1997). *Reattributionstrainings*. Regensburg: Roderer.

Ziegler, A. & Schober, B. (2001). *Theoretische Grundlagen und praktische Anwendung von Reattributionstrainings*. Regensburg: Roderer.

Zimmer, G. (2002). E-Learning führt zu einer anderen Kultur des Lehrens und Lernens. Folgen für didaktische Gestaltung. In G. Zimmer (Hrsg.), *E-Learning: High-Tech or High-Teach? Lernen in Netzen zwischen Aktualität und Potenzialität* (S. 5-18). Bielefeld: Bertelsmann.

Zimmerman, B.J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81* (3), 329-339.

Zimmerman, B.J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P.R. Pintrich & M. Zeidner (Hrsg.), *Handbook of self-regulation* (S. 13-39). San Diego, CA: Academic Press.

Zimmerman, B.J. (2001). Theories of self-regulated learning and academic achievement: An overview and analysis. In B.J. Zimmerman & D.H. Schunk (Hrsg.), *Self-regulated learning and academic achievement: Theoretical perspectives* (2. Aufl.), (S. 1-38). Mahwah, NJ: Lawrence Erlbaum Associates.

5 Anhang

5.1 Anhang – Terminübersicht für WS 2003/04 und SS 2004

Tabelle 38: Übersicht der Termine für die Online-Module und “Face-to-face“-Einheiten des Wintersemester 2003/04 und des Sommersemester 2004

Online-Einheiten Track A + B	“Face-to-face“-Einheiten
	15.10.03 Kick-off Track A + B
	20.10.03 Proseminar Track A
	22.10.03 Proseminar Track A
	27.10.03 Proseminar Track A
28.10. - 11.11.03 Modul 1	
	29.10.03 Meet-the-Expert I Track A + B
	10.11.03 Proseminar Track A
11.11. - 25.11.03 Modul 2	
25.11. - 09.12.03 Modul 3	
	26.11.03 Meet-the-Expert II Track A + B
09.12. - 07.01.04 Modul 4	
	15.12.03 Proseminar Track A
07.01. - 21.01.04 Modul 5	
	21.01.04 Meet-the-Expert III Track A + B
	28.01.04 1. Prüfungstermin WS 03/04
	03.03.04 2. Prüfungstermin WS 03/04
	23.06.04 3. Prüfungstermin WS 03/04
	15.03.04 Proseminar Track A
16.03. - 02.04.04 Modul 6	
	31.03.04 Meet-the-Expert IV Track A + B
02.04. - 23.04.04 Modul 7	
	19.04.04 Proseminar Track A
27.04. - 14.05.04 Modul 8	
	05.05.04 Meet-the-Expert V Track A + B
	10.05.04 Proseminar Track A
18.05. - 04.06.04 Modul 9	
08.06. - 25.06.04 Modul 10	
	16.06.04 Meet-the-Expert VI Track A + B
	30.06.04 1. Prüfungstermin SS 04
	11.10.04 2. Prüfungstermin SS 04
	24.01.05 3. Prüfungstermin SS 04

5.2 Anhang – Attributionale Rückmeldungen der zehn Selbsttests

Modul 1 – 8 Aufgaben

Tabelle 39: Ergebnisabhängige attributionale Rückmeldungen zum Selbsttest 1

8 bis 8 **100%**	– Man merkt, Sie haben sich intensiv mit dem Thema auseinandergesetzt. – Sie beherrschen das offenbar ohne Schwierigkeiten. – Super, Sie waren bei Ihren Antworten sehr genau. – Fein, Sie kennen sich offenbar schon recht gut aus. – Sie haben sehr gründlich und mit viel Verständnis gearbeitet. – Das war sehr gut, Sie wissen wirklich viel zu Modul 1. – Die Aufgaben waren teilweise ganz schön knifflig, aber Sie wissen wirklich eine Menge!
6 bis 7 **75 – 88%**	– Wenn Sie über die Aufgaben noch etwas genauer nachdenken, werden Sie sicher alle lösen können. – Im Wesentlichen haben Sie die wichtigsten Inhalte von Modul 1 verstanden. Wenn Sie noch etwas genauer sind, wird es sicher perfekt. – Sie hatten schon viele richtige Lösungen, schauen Sie sich die wenigen Unklarheiten noch einmal speziell an! – Denken Sie genau nach, dann schaffen Sie sicher auch noch die letzten Aufgaben! – Respekt. Sie können sicher die restlichen Aufgaben auch noch lösen.
4 bis 5 **50 – 63%**	– Wenn Sie den Aufgaben mehr Aufmerksamkeit widmen, werden Sie sicher noch mehr lösen können. – Denken Sie etwas genauer über die Sache nach und lesen Sie sich die entsprechenden Stellen im Text noch mal durch! – Lesen Sie gezielt nach und probieren Sie es noch mal, Sie schaffen dann sicher noch mehr richtige Lösungen! – Die Aufgaben sind wirklich nicht so leicht. Nutzen Sie Ihre Fehler, um gezielt Wissenslücken zu schließen! – Sie wissen schon einiges, beschäftigen Sie sich aber noch mal genau mit den relevanten Stellen in der Literatur.
0 bis 3 **0 – 38%**	– Sie sollten sich Modul 1 nochmals genau ansehen. – Sie haben sich mit dem Basisvokabular der empirischen Forschung offenbar noch nicht genug beschäftigt. – Sie sollten die Literatur noch mal genau durchgehen, dann werden Sie die Aufgaben besser lösen können. – Sie schaffen sicher mehr, wenn Sie der Sache mehr Aufmerksamkeit widmen! – Die Aufgaben haben schon ihre Tücken, lesen Sie in der Literatur zu Modul 1 nach und versuchen Sie es noch mal! – Antworten Sie aufmerksamer, dann werden Sie sicher mehr Aufgaben lösen können.

Modul 2 – 10 Aufgaben

Tabelle 40: Attributionale Rückmeldungen zu Selbsttest 2

9 bis 10 **90 – 100%**	– Super, Sie haben die Fragen sehr konzentriert bearbeitet! – Ausgezeichnet, Ihre Bemühungen, Modul 2 zu verstehen, haben sich gelohnt. – Sie haben Modul 2 offenbar sehr gründlich und mit viel Verständnis bearbeitet. – Die Abgrenzung von Evaluations- und Grundlagenforschung haben Sie voll im Griff! – Gratulation, Sie beherrschen die wichtigsten Inhalte von Modul 2. – Das war sehr gut, Sie wissen wirklich viel zu Modul 2. – Die Fragen waren nicht einfach, Sie haben sie aber trotzdem hervorragend gelöst.
7 bis 8 **70 – 80%**	– Sie kennen sich bei der Abgrenzung von Evaluations- und Grundlagenforschung schon recht gut aus. Lesen Sie noch mal gezielt nach, beim nächsten Mal beantworten Sie sicher alle richtig! – Nicht schlecht! Sie wissen schon eine ganze Menge, wenn Sie noch etwas genauer sind, wird es sicher perfekt. – Das war schon sehr gut, ein paar Stellen aus Modul 2 sollten Sie aber noch mal durchgehen. – Denken Sie genau nach, dann schaffen Sie sicher auch noch die letzten Aufgaben!
4 bis 6 **40 – 60%**	– Das war schon knapp die Hälfte. Lesen Sie die Fragen genauer und denken Sie gründlich nach, dann schaffen Sie noch mehr. – Sie könnten sicher noch mehr Aufgaben lösen, versuchen Sie es noch Mal! – Sie haben schon einiges über die Evaluations- und Grundlagenforschung gelernt. Nutzen Sie Ihre Fehler, um gezielt Wissenslücken zu schließen! – Diese Fragen sind nicht ganz einfach, aber Sie können sicher noch mehr schaffen. – Lesen Sie gezielt zu den falschen Antworten nochmals nach und Sie werden beim nächsten Mal mehr als die Hälfte lösen können!
0 bis 3 **0 – 30%**	– Sie sollten sich mit Modul 2 noch genauer beschäftigen. – Die Inhalte von Modul 2 sind nicht ganz banal, Sie sollten es nochmals konzentriert durcharbeiten. – Geben Sie sich etwas mehr Mühe, Sie schaffen sicher mehr, wenn Sie aufmerksamer lesen und genauer nachdenken. – War auch nicht so einfach! Lassen Sie sich Zeit, denken Sie gründlich nach und versuchen Sie es noch mal! – Arbeiten Sie genauer und Sie werden mehr schaffen!

Modul 3 – 6 Aufgaben

Tabelle 41: Ergebnisabhängige attributionale Rückmeldungen zum Selbsttest 3

6 bis 6 **100%**	– Sie haben die Untersuchungsarten in Evaluationsstudien wirklich gründlich durchgearbeitet. Kompliment! – Sehr gute Leistung - ein bisschen Zeit in genaues Durcharbeiten zu investieren lohnt sich! – Wenn Sie genau arbeiten, können Sie wirklich sehr gute Antworten geben, super! – Sie haben die Untersuchungsarten in Evaluationsstudien sehr gut verstanden! – Die Untersuchungsarten haben Sie im Griff. Machen Sie weiter so! – Sie haben zum Thema Untersuchungsarten wirklich eine Menge gelernt!
5 bis 5 **83%**	– Wenn Sie über die Fragen noch etwas genauer nachdenken, werden Sie sicher auch noch die Letzte lösen können. – Man merkt, dass Sie sich mit dem Thema Untersuchungsarten auseinandergesetzt haben! – Das war schon sehr gut, gehen Sie das Thema hinsichtlich der letzten falschen Antwort nochmals genau durch. – Die Problematik der Untersuchungsarten haben Sie schon fast durchschaut!
3 bis 4 **50 – 67%**	– Sie können sicher noch mehr Fragen richtig beantworten, versuchen Sie es noch einmal! – Lesen Sie gezielt nach und konzentrieren Sie sich noch mehr auf die genaue Formulierung der Frage, dann werden Sie sicher noch mehr lösen können. – Die Fragen sind nicht leicht, aber mit ein bisschen mehr Nachdenken sicher alle für Sie machbar. – Sie haben schon etwa die Hälfte der Fragen richtig, die genaue Differenzierung der Untersuchungsarten ist aber auch schwierig. – Das war schon ganz o.k. Diskutieren Sie doch die unklaren Fragen mal mit Ihren KollegInnen!
0 bis 2 **0 – 33%**	– Ein bisschen genauer sein müssen Sie schon! Mit mehr Einsatz werden Sie mehr richtige Antworten geben können. – Das Thema Untersuchungsarten muss man verstehen und sich praktisch vorstellen. Lesen Sie nach und versuchen Sie es nochmals! – So ganz einfach sind diese Aufgaben nicht. Denken Sie genauer nach!

Modul 4 – 6 Aufgaben

Tabelle 42: Ergebnisabhängige attributionale Rückmeldungen zum Selbsttest 4

6 bis 6 **100%**	– Die Fragen sind ganz schön knifflig, aber Sie wissen wirklich viel über Versuchspläne und Hypothesenprüfung! – Das war hervorragend! Bearbeiten Sie alle Inhalte aus Modul 4 mit der gleichen Sorgfalt wie diese und Sie haben die inhaltlichen Ziele dieses Moduls bravourös erreicht. – Das war sehr gut, Sie haben ein wirklich differenziertes Wissen zu Versuchsplänen und Hypothesenprüfung. – Sehr gute Leistung - Sie haben über Versuchspläne und Hypothesenprüfung viel gelernt! – Bravo! Man merkt, Sie haben sich intensiv mit dem Thema auseinandergesetzt.
5 bis 5 **83%**	– Sie haben schon fast alle Fragen richtig beantwortet, schauen Sie sich die wenigen Unklarheiten noch einmal speziell an! – Das war ziemlich gut! Sie wissen schon einiges und wenn Sie noch genauer sind, wird es sicher perfekt. – Es wird deutlich, dass Sie schon einiges können. Lesen Sie zu der einen falschen Aufgabe nochmals gezielt nach. – Diese Aufgaben bereiten vielen Studierenden Schwierigkeiten. Sie haben aber offenbar nur noch kleinere Probleme, die Sie sicher auch noch lösen können!
3 bis 4 **50 – 67%**	– Die Aufgaben sind wirklich nicht so leicht. Nutzen Sie Ihre Fehler um gezielt die Lücken zu schließen! – Das war schon nicht schlecht, aber denken Sie noch etwas genauer nach, bevor Sie antworten. Was haben Sie noch nicht verstanden? – Die Fragen haben schon ihre Tücken. Probieren Sie es noch einmal mit mehr Aufmerksamkeit. – Sie können Ihre Fähigkeiten sicher noch um einiges steigern! Diskutieren Sie Versuchspläne und Hypothesenprüfung z.B. detailliert mit Ihren KollegInnen. – Haben Sie sich wirklich konzentriert? Falls ja, lesen Sie im Text zu Modul 4 gezielt nach. Falls nein, versuchen Sie es einfach noch mal, wenn Sie mehr Energie haben. – Sie sollten Modul 4 nicht unterschätzen und um Versuchspläne und Hypothesenprüfung wirklich zu verstehen, braucht es Ihre volle Aufmerksamkeit!
0 bis 2 **0 – 33%**	– Die Fragen sind nicht leicht, arbeiten Sie also den Stoff genauer durch, dann lösen Sie sicher mehr. – Die Fragen haben schon ihre Tücken. Probieren Sie es noch einmal mit mehr Aufmerksamkeit. – Sie können Ihre Fähigkeiten sicher noch um einiges steigern! Diskutieren Sie Versuchspläne und Hypothesenprüfung z.B. detailliert mit Ihren KollegInnen. – Haben Sie sich wirklich konzentriert? Falls ja, lesen Sie im Text zu Modul 4 gezielt nach. Falls nein, versuchen Sie es einfach noch mal, wenn Sie mehr Energie haben. – Sie sollten Modul 4 nicht unterschätzen und um Versuchspläne und Hypothesenprüfung wirklich zu verstehen, braucht es Ihre volle Aufmerksamkeit!

Modul 5 – 10 Aufgaben

Tabelle 43: Ergebnisabhängige attributionale Rückmeldungen zum Selbsttest 5

9 bis 10 **90 – 100%**	– Super, Sie haben sehr gründlich und mit viel Verständnis gearbeitet! – Sehr gut - Ihre Bemühungen, die Literatur zu den Untersuchungsdesigns zu verstehen, haben sich gelohnt! – Das war nicht leicht, aber Sie haben schon viel gelernt und es bravourös gemeistert! – Die Aufgaben haben es teilweise ganz schön in sich, aber Sie kann man nicht so leicht in die Irre führen! – Bravo, beim Thema Untersuchungsdesigns kennen Sie sich jetzt aus. – Hervorragend! Sie wissen wirklich viel über Untersuchungsdesigns in Evaluationskontexten!
7 bis 8 **70 – 80%**	– Sie hatten schon sehr viele Aufgaben gut im Griff. Wenn Sie über die falsch beantworteten nochmals ganz genau nachdenken, schaffen Sie sicher auch diese! – Sie wissen schon viel über Untersuchungsdesigns, lesen Sie die paar letzten Unklarheiten noch gezielt nach. – Versuchen Sie es noch mal und sehen Sie genau hin! Sie schaffen die letzten Aufgaben bestimmt auch noch! – Diesmal sind die Aufgaben für viele Studierende in der Tat nicht leicht, aber SIE sind auf dem besten Weg! – Gehen Sie die 2-3 relevanten Stellen in Modul 5 nochmals durch, dann haben Sie's!
4 bis 6 **40 – 60%**	– In etwa die Hälfte der Fragen ist schon richtig. Sie können sicher noch mehr! – Diesmal ist der Selbsttest wahrlich nicht ohne - aber Sie sind schon ganz gut dabei. Nachlesen und noch mal versuchen! – Den Grundstock zu den Untersuchungsdesigns haben Sie, nutzen Sie Ihre Fehler um Ihr Wissen zu vervollständigen! – Seien Sie etwas genauer, Sie wissen schon einiges! – Beschäftigen Sie sich nochmals gezielt mit den unklaren Stellen und Sie schaffen beim nächsten Durchgang mehr! – Nicht schlecht, aber noch genauer hinsehen und nachdenken!
0 bis 3 **0 – 30%**	– Sie sollten sich Modul 5 nochmals ganz genau ansehen! – Modul 5 ist das Finale des ersten Semesters, das sollten Sie nicht unterschätzen! – Erst lesen und nachdenken, dann ankreuzen! – Damit sollten SIE sich nicht zufrieden geben - auf ein Neues! – Sie sollten mit den Untersuchungsdesigns doch noch etwas mehr Zeit verbringen! – Ok, die Aufgaben sind nicht leicht und erfordern genaues Nachdenken. Lassen Sie sich Zeit und lesen Sie noch mal nach!

Modul 6 – 7 Aufgaben

Tabelle 44: Ergebnisabhängige attributionale Rückmeldungen zum Selbsttest 6

7 bis 7 **100%**	– Sie haben die Datenerhebungsmethoden schon ziemlich gut durchschaut! – Sie haben sehr differenziert gelernt. Machen Sie weiter so und Sie haben die Themen der Vorlesung am Ende voll im Griff! – Bravo! Das war nicht leicht, aber Sie haben sich sehr gut vorbereitet. – Ihr Wissen zu den Datenerhebungsmethoden ist unglaublich exakt! Gratuliere. – Bravo! Man merkt, Sie haben sich intensiv mit dem Thema auseinandergesetzt. – Die Fragen waren schwierig, Sie aber haben alle richtig. Gekonnt bearbeitet, super!
5 bis 6 **71 – 86%**	– Nicht schlecht! Sie wissen schon eine ganze Menge, wenn sie noch genauer sind, schaffen Sie sicher noch mehr Richtige! – Das war schon ziemlich gut! überdenken Sie Ihre wenigen Fehler nochmals und Sie lernen noch was dazu. – Es wird deutlich, dass Sie schon einiges können. Lesen Sie zu der/den falschen Aufgabe(n) nochmals gezielt nach. – Diese Aufgaben bereiten vielen Studierenden Schwierigkeiten. Sie haben aber offenbar nur noch kleinere Probleme, die Sie sicher bald lösen werden!
3 bis 4 **43 – 57%**	– Die Aufgaben sind wirklich nicht so leicht. Die andere Hälfte der Aufgaben werden Sie sicher auch noch verstehen. Diskutieren Sie mit Ihren KollegInnen! – Das war nicht schlecht, aber denken Sie noch etwas genauer nach, bevor Sie antworten. – Achten Sie auf die genaue Formulierung der Fragen und der Antworten, dann werden Sie sicher noch mehr richtig beantworten. – Das war immerhin schon knapp die Hälfte. Diskutieren Sie doch die Inhalte der unklaren Fragen mal übers Forum!
0 bis 2 **0 – 29%**	– Sie können Ihre Fähigkeiten sicher noch um einiges steigern! Diskutieren Sie am Besten mit ihren KollegInnen über die Erhebungsmethoden. – Lesen Sie genauer, das wird sich sicher auszahlen!</ – Haben Sie sich wirklich konzentriert? Falls ja, lesen Sie im Text zu Modul 6 gezielt nach. Falls nein, versuchen Sie es einfach noch mal, wenn Sie mehr Energie haben. – Die Fragen sind nicht leicht, man muss sich schon ein bisschen anstrengen, um den Stoff genau zu verstehen! – Hier wird wirklich differenziertes Basiswissen für PsychologInnen abgefragt! Es zu verstehen ist wichtig, das braucht aber Ihre volle Aufmerksamkeit! Also noch einmal ran an den Stoff!

Modul 7 – 11 Aufgaben

Tabelle 45: Ergebnisabhängige attributionale Rückmeldungen zum Selbsttest 7

11 bis 11 **100%**	– Man merkt, Sie haben nicht einfach drauf los geklickt, sondern gut überlegt und Ihr Wissen angewendet. Bravo! – Super, Sie haben die Fragen sehr konzentriert bearbeitet! – Ausgezeichnet, Ihre Bemühungen, Modul 7 zu verstehen, haben sich gelohnt. – Die Fragen waren teilweise nicht leicht, aber Sie wissen offenbar schon sehr gut, worauf es ankommt. – Beim Thema Datenauswertung in Evaluationskontexten kann man Ihnen jetzt nicht mehr viel vormachen! – Auch der etwas mühsamere englische Text war für Sie keine Verständnishürde. Seien Sie sich bewusst, dass Sie nun schon eine Menge zu diesem Thema können.
8 bis 10 **73 – 91%**	– Respekt. Wenn Sie noch etwas genauer sind, können Sie sicher die restlichen Aufgaben auch noch lösen. – Man kann Sie beim Thema Datenauswertung nicht mehr so leicht in die Irre führen. Lesen Sie aber die letzten unklaren Stellen noch mal nach! – Sie haben schon eine Menge über Datenauswertung gelernt; nutzen Sie die wenigen Fehler, um Ihr Wissen gezielt zu vervollständigen. – Es fehlt nicht mehr viel und Sie lösen die Aufgaben perfekt. Denken Sie in der nächsten Runde genau nach! – Sie haben zentrale Inhalte von Modul 7 schon sehr gut verstanden, arbeiten Sie noch etwas konzentrierter und Sie haben sie im Griff!
5 bis 7 **45 – 64%**	– Sie wissen schon einiges, beschäftigen Sie sich aber noch mal genau mit den relevanten Stellen in der Literatur. – Lesen Sie gezielt nach und probieren Sie es noch mal, Sie schaffen dann sicher noch mehr richtige Lösungen! – Das war immerhin schon knapp die Hälfte. Und mit mehr Konzentration beim Lesen und Antworten ist für Sie sicher noch mehr drin! – Die Fragen sind keineswegs alle so leicht. Man muss sehr genau lesen und antworten. Versuchen Sie es noch mal! – Der Stoff von Modul 7 ist Ihnen keineswegs mehr fremd, aber Sie sollten gezielt die noch bestehenden Wissenslücken zu schließen!
0 bis 4 **0 – 36%**	– Die Datenauswertung im Evaluationskontext ist alles andere als banal. Sehen Sie sich Modul 7 nochmals genau an. – Sie schaffen sicher mehr, wenn Sie der Sache mehr Aufmerksamkeit widmen! – Versuchen Sie es noch mal mit mehr Konzentration! Wenn das noch nicht den Durchbruch bringt, beschäftigen Sie sich einfach noch mal mit dem Text! – Etwas mehr Mühe müssen Sie sich schon geben beim Antworten wie beim vorhergehenden Lesen! Man sollte dieses Thema nicht unterschätzen – Lesen Sie in der Literatur zu Modul 7 noch mal nach und versuchen es dann noch mal!

Modul 8 – 7 Aufgaben

Tabelle 46: Ergebnisabhängige attributionale Rückmeldungen zum Selbsttest 8

7 bis 7 **100%**	– Bravo! Ihre Bemühungen sich mit der Ergebnisdarstellung auseinander zu setzen haben sich 100%ig gelohnt! – Die harte Arbeit hat sich ausgezahlt! Für's wissenschaftliche Beschreiben von Ergebnissen sind Sie jetzt bestens gerüstet. – Methodische Details in einem Artikel sind für Sie kein Problem. Nutzen Sie ihr Können beim kritischen Lesen weiterer Artikel! – Trotz der Schwierigkeit des englischen Artikels haben Sie die wesentlichen Punkte erfasst. – Wunderbar, Sie haben wichtige Facetten bei der Ergebnisdarstellung verstanden!
6 bis 6 **86%**	– Das war schon fast perfekt! Wenn Sie nochmals genau nachdenken und/oder nochmals nachlesen, werden Sie die letzte Frage auch noch richtig beantworten. – Klar ist: Sie haben sich schon sehr viel Wissen angeeignet! Beim Diskutieren wird vielleicht die letzte Unklarheit auch noch ausgeräumt. – Methodische Details in einem Artikel scheinen Ihnen zu liegen. Noch ein bisschen mehr Konzentration und Sie können alle Fragen richtig beantworten! – Die Fragen zu Modul 8 bereiten vielen Studierenden Schwierigkeiten. Sie brauchen nur noch eine richtige Antwort, dann haben Sie alle angesprochenen Aspekte verstanden! – Die letzte richtige Antwort werden Sie auch noch schaffen! Sie wissen jetzt schon wirklich sehr viel!
4 bis 5 **57 – 71%**	– Das war immerhin schon gut die Hälfte! Lesen und Diskutieren Sie doch nochmals die unklaren Punkte. – Die Fragen sind keineswegs alle so leicht. Man muss sehr genau lesen und antworten. Versuchen Sie es noch mal! – Schon ganz o.k., aber manche Fragen verlangen schon einiges Verständnis! Denken Sie genau noch und probieren Sie's dann wieder. – Schade, nicht alle Antworten waren korrekt! Die Details in einem wissenschaftlichen Artikel bedürfen aber auch wirklich genauer Überlegungen. – Sie können sicher noch mehr! Beschäftigen Sie sich konkret mit den unklaren Stellen und beim nächsten Durchgang schaffen Sie sicher mehr!
0 bis 3 **0 – 43%**	– Die wissenschaftliche Ergebnisdarstellung ist alles andere als banal. Sehen Sie sich Modul 8 nochmals genau an. – Sie können Ihre Fähigkeiten noch steigern! Diskutieren Sie mit kompetenten KollegInnen, dann werden Sie auch die komplexen Facetten eines Artikels verstehen lernen. – Modul 8 geht in die Bewertung Ihrer Leistung ein. Mit diesem Stoff sollten Sie sich genauer auseinandersetzen! – Investieren Sie noch mehr Zeit in die Aspekte des Basisartikels! Bei mehr Einsatz werden Sie diese grundlegend wichtigen Inhalte sicher verstehen. – Arbeiten Sie genauer und Sie werden mehr schaffen!

Modul 9 – 10 Aufgaben

Tabelle 47: Ergebnisabhängige attributionale Rückmeldungen zum Selbsttest 9

9 bis 10 **90 – 100%**	– Kompliment! Sie haben den vollen Überblick! – Die harte Arbeit hat sich für Sie ausgezahlt! Sie zeigen grundlegendes Verständnis für Forschungsmethoden und Evaluation. – Hut ab! Trotz der Vielfalt der Themen zeigen Sie wirklich ausgezeichnetes Können. – Arbeiten Sie genauso gewissenhaft bei der Abschlussprüfung und es kann Ihnen nichts mehr passieren! – Wenn Sie sich weiterhin so erfolgreich einsetzen, könnten Sie eines Tages ein(e) ausgezeichnete(r) Evaluator(in) werden.
7 bis 8 **70 – 80%**	– Sie wissen schon sehr viel über Forschungsmethoden und Evaluation. Lesen Sie die Fragen noch ein bisschen genauer, dann schaffen Sie auch die letzten Antworten richtig. – Klar ist: Sie haben sich schon sehr viel Wissen angeeignet! Beim Diskutieren wird vielleicht die letzte Unklarheit auch noch ausgeräumt. – Die Wiederholung der bisherigen Stoffgebiete ist gar nicht so einfach. Aber es fehlt nicht mehr viel, probieren Sie's noch mal! – Nur weiter so, die letzten Unklarheiten werden Sie auch noch beseitigen. – Das war schon sehr gut. Diskutieren Sie am besten die unklaren Details noch über die Plattform.
4 bis 6 **40 – 60%**	– Ca. die Hälfte der Fragen ist schon richtig beantwortet! Setzen Sie sich nochmals mit den Unklarheiten detailliert auseinander! – Die Fragen sind wirklich nicht so leicht, aber es geht ja um wichtiges Basiswissen. Seien Sie noch genauer! – Nicht schlecht, aber schauen Sie noch genauer hin und denken Sie gut nach, bevor Sie antworten! – Sie können sicher noch mehr! Gehen Sie die relevanten Stoffteile mit Ihren KollegInnen nochmals durch. – Mit diesen zentralen Fragen des bisherigen Stoffes sollten Sie sich doch noch genauer auseinandersetzen. Auf ein Neues!
0 bis 3 **0 – 30%**	– Sie sollten ihre Wissensgrundlagen auf jeden Fall noch ausbauen! Nehmen Sie sich mehr Zeit, um Unklares noch genauer zu durchdenken. – Sie schaffen sicher mehr, wenn Sie der Sache mehr Aufmerksamkeit widmen! – In Forschungsmethoden und Evaluation kommt es sehr auf Ihr Verständnis an, vor allem bei diesem Grundlagenwissen! Wiederholen Sie die Stoffgebiete, die Ihnen noch Schwierigkeiten bereiten! – Die Aufgaben sind nicht leicht, sondern erfordern genaues Lesen und Nachdenken. Lassen Sie sich Zeit und lesen Sie nochmals nach. – Wissenslücken sind nichts Schlechtes, solange man sich bemüht sie zu schließen!

Modul 10 – 10 Aufgaben

Tabelle 48: Ergebnisabhängige attributionale Rückmeldungen zum Selbsttest 10

9 bis 10 **90 – 100%**	– Ausgezeichnet, Sie haben zur Entwicklung der Evaluationsforschung schon sehr viel gelernt! – Sie haben den Durchblick! Kompliment! – Bearbeiten Sie alle Themenbereiche so wie die Entwicklung der Evaluationsforschung und Sie werden die Prüfung problemlos bestehen! – Bravo! Man merkt, Sie haben sich intensiv mit dem Thema auseinandergesetzt. – Sie haben Modul 10 offenbar sehr gründlich und mit viel Verständnis bearbeitet.
7 bis 8 **70 – 80%**	– Man sieht, dass Sie schon einiges gelernt haben! Lesen Sie nochmals nach, dann schaffen Sie sicher auch noch die restlichen Fragen! – Noch genauer lesen und nachdenken! Bei Ihnen sind "10 Richtige" drin! – Das war schon sehr gut, ein paar Stellen aus Modul 10 sollten Sie aber gezielt noch mal durchgehen. – Sie sind auf dem besten Wege! Klären Sie noch die letzten Unklarheiten zur Entwicklung der Evaluationsforschung! – Ihr konzentriertes Arbeiten hat sich gelohnt, Sie wissen schon einiges über die Entwicklung der Evaluationsforschung. Gehen Sie es noch mal an und kümmern Sie sich gezielt um die Aspekte, die noch unklar sind!
4 bis 6 **40 – 60%**	– Wenn Sie den Aufgaben mehr Aufmerksamkeit widmen, werden Sie sicher noch mehr lösen können. – Sie haben schon einiges über die Entwicklung der Evaluationsforschung gelernt. Nutzen Sie Ihre Fehler, um gezielt Wissenslücken zu schließen! – Das war schon knapp die Hälfte. Lesen Sie die Fragen genauer und denken Sie gründlich nach, dann schaffen Sie noch mehr. – Die Fragen sind nicht ganz einfach, aber Sie können sicher noch mehr lösen. – Diskutieren Sie doch die unklaren Aspekte mal in Ihrer Gruppe und versuchen Sie es dann noch mal! – Denken Sie gerade über Ihre Fehler noch mal genau nach und Sie werden beim nächsten Mal mehr als die Hälfte lösen können!
0 bis 3 **0 – 30%**	– Geben Sie sich etwas mehr Mühe, Sie können sicher mehr schaffen! – Sie sollten sich mit der Entwicklung der Evaluationsforschung doch noch mal genauer auseinandersetzen. – Auch wenn es diesmal nicht mit komplizierter Statistik zu tun hat, ist Modul 10 keineswegs banal. Man sollte das Thema nicht unterschätzen. – Viele Fehler sind hier erst mal kein Problem, nutzen Sie sie, um Ihre Lücken zu schließen!

5.3 Anhang – SRL-Fragebogen des MZP 2 im Studienjahr 2003/04

Wie arbeite ich?

Ihr CODE:
(falls Code vergessen, bitte Matrikelnr.)

Ihr Geschlecht: ☐ weiblich ☐ männlich

Bitte lesen Sie den folgenden Text im Kasten durch und bearbeiten Sie die nachstehende Aufgabe, so gut es Ihnen möglich ist.

Uli soll in 8 Tagen ein Referat zum Themenbereich „Forschungsmethoden" halten. Das genaue Thema kann Uli frei wählen.

Heute war Uli in der Bibliothek, um Literatur für das Referat zu suchen, dabei ist Uli nochmals klar geworden, dass „Forschungsmethoden" ein sehr weiter Begriff ist.

Nun befürchtet Uli, den Überblick zu verlieren und dass die Zeit für eine gute Vorbereitung zu kurz werden könnte.

Bitte stellen Sie sich vor, Sie wären in Ulis Situation.
Wie würden Sie vorgehen, um dieses Referat möglichst gut vorzubereiten?

Bitte machen Sie Ihre Angaben so konkret wie möglich. Zur Erleichterung ist ein Raster unterteilt in Tagen vorgegeben.

Was genau tun Sie an den 7 verbleibenden Tagen?

Tag 1: _____

Tag 2: _____

Tag 3: _____

Tag 4: _____

Tag 5: _____

Tag 6: _____

Tag 7: _____

Tag 8: 10:00 Uhr: Referat muss präsentiert werden!

Wie lerne ich?

Wir bitten Sie, die folgenden Fragen genau durchzulesen und jeweils die Antwort anzukreuzen, die am besten für Sie zutrifft. Der Fragebogen betrifft **Ihr Lernen und Arbeiten im Bereich der psychologischen Forschungsmethoden** (hierzu zählen alle LVen zu Methodenlehre, Statistik, Auswertungsmethoden, etc.). Antworten Sie also bitte immer mit Blick auf *diesen* Bereich. Es gibt hierbei keine richtigen oder falschen Antworten – es geht einzig um Ihre ganz persönliche Einschätzung.

Achten Sie bitte darauf, dass Sie keine Fragen auslassen.

Vielen Dank!

	nie	selten	manch-mal	oft	sehr oft	fast immer
Neue Begriffe, Definitionen usw. lerne ich, indem ich mir entsprechende Beispiele und Situationen vorstelle.						
Sobald ich merke, dass ich mein Lernpensum nicht schaffen werde, teile ich mir den Stoff neu ein.						
Auch wenn ich bestimmte Inhalte langweilig finde, arbeite ich sie durch.						
Wenn ich bestimmte Arbeiten erledige, so setze ich mir einen zeitlichen Rahmen und halte mich auch daran.						
Wenn ich einmal Hilfe brauche, hole ich sie mir von meinen KollegInnen.						
Wenn ich für eine Prüfung lerne, überlege ich jeweils nach dem Lernen, ob ich richtig vorgegangen bin.						
Wenn ich einmal etwas nicht verstehe, dann mache ich es mir auf einem anderen Weg klar.						
Ich arbeite so lange, bis ich mir sicher bin, einen schwierigen Stoff verstanden zu haben.						
Ich schreibe mir kurze Zusammenfassungen des zu lernenden Stoffes auf.						
Ich versuche mir den Lernstoff durch Wiederholungen zu merken.						
Ich unterbreche mein Lernen von Zeit zu Zeit, um über mein bisheriges Vorgehen nachzudenken.						
Nach dem Lernen überlege ich mir, ob mein Vorgehen sinnvoll war.						
Wenn ich mit dem Lernen beginnen sollte, dann schiebe ich es hinaus.						
Während des Lernens versuche ich immer wieder herauszufinden, was ich noch nicht verstehe.						
Um den Lernstoff zu strukturieren, mache ich mir Übersichten, Tabellen oder Skizzen.						
Ich lese mir den Lernstoff mehrmals hintereinander durch.						
Ich arbeite hart, um erfolgreich zu sein, selbst wenn ich die betreffende Lehrveranstaltung nicht mag.						
Ich stelle mir selbst Fragen, um sicher zu sein, dass ich den Stoff verstanden habe.						

	stimme gar nicht zu	stimme nicht zu	stimme eher nicht zu	stimme eher zu	stimme zu	stimme völlig zu
Im Bereich Forschungsmethoden geht es mir vor allem darum, eine Idee davon zu bekommen, wie empirische Forschung funktioniert.						
Die Beschäftigung mit Forschungsmethoden gehört nicht gerade zu meinen Lieblingstätigkeiten.						
Es ist nicht von Vornherein festgelegt, wie viel ich im Bereich Forschungsmethoden verstehen kann; ich kann schrittweise dazulernen und mich immer weiter verbessern.						
Misserfolge zeigen mir auf, was ich anders machen könnte.						
Im Bereich Forschungsmethoden geht es mir vor allem darum, dass die Anderen nicht glauben, ich sei zu dumm dafür.						
Mir liegt viel daran, im Bereich Forschungsmethoden viel zu wissen.						
Ich weiß, dass ich den Anforderungen des Faches gewachsen bin.						
Eigentlich ist mir der Bereich Forschungsmethoden nicht wichtig.						
Im Bereich Forschungsmethoden geht es mir vor allem darum, eine gute Note zu bekommen.						
Ich kann mein Können im Bereich Forschungsmethoden steigern.						
Misserfolge sind für mich Anlass, darüber nachzudenken, was ich das nächste Mal anders machen sollte.						
Wenn sich beim Lernen etwas als sehr schwierig herausstellt, deprimiert mich das und ich habe kaum mehr Lust weiterzumachen.						
Ich kann auch schwierige Fragen lösen, wenn ich mich anstrenge.						
Im Bereich Forschungsmethoden geht es mir vor allem darum, bei der Prüfung nicht durchzufallen.						

	nie	selten	manch-mal	oft	sehr oft	fast immer
Wenn Stoffinhalte schwierig sind, dann setze ich mich mit ihnen nicht weiter auseinander.						
Ich versuche so oft wie möglich zusammen mit anderen KollegInnen zu lernen.						
Den Stoff, den ich lerne, versuche ich mit eigenen Worten auszudrücken.						
Wenn mir ein bestimmter Inhalt unklar ist, gehe ich ihn noch einmal genau durch.						
Sobald ich merke, dass ich ein bestimmtes Ziel nicht erreichen werde, ändere ich mein Vorgehen.						
Ich lerne wichtige Stichwörter auswendig, um mich besser an die wesentlichen Lerninhalte erinnern zu können.						
Nachdem ich gelernt habe, überprüfe ich, ob ich das, was ich mir vorgenommen habe, auch geschafft habe.						
Während des Lernens versuche ich, Zusammenhänge zwischen verschiedenen Aspekten, Ideen und Themen herzustellen.						
Beim Lernen halte ich mich an einen Zeit- und Arbeitsplan.						
Ich überlege mir während der Arbeit immer wieder, ob mein bisheriges						

	nie	selten	manch-mal	oft	sehr oft	fast immer
Wenn beim Lernen was nicht klappt, versuche ich die Schwierigkeiten festzustellen und den Stoff gezielt noch einmal durchzugehen.						
Ich lerne Merksätze und Definitionen auswendig.						
Nur bei wirklich interessanten Inhalten mache ich mir die Mühe, sie zu verstehen.						
Ich lerne Teile des Stoffes möglichst auswendig.						
Um Ordnung in meine Gedanken zu bekommen, mache ich mir Stichworte.						
Wenn ich ein schwieriges Problem lösen soll, passe ich mein Vorgehen der Schwierigkeit an.						
Wenn ich eine Aufgabe bearbeite, überprüfe ich immer wieder, ob ich gezielt vorgehe.						
Ich gebe nicht auf, auch wenn der Stoff sehr schwierig ist.						
Ich versuche Beziehungen zu finden zwischen dem, was ich gerade lerne und dem, was ich bereits weiß.						
Während der Aufgabenbearbeitung überlege ich mir, ob ich meine Planung einhalten kann.						

Wenn ich im Bereich Forschungsmethoden **einen Misserfolg** erlebt habe, dann lag das ...

meistens an _____

manchmal auch an _____

mitunter auch an _____

	stimme gar nicht zu	stimme nicht zu	stimme eher nicht zu	stimme eher zu	stimme zu	stimme völlig zu
Im Bereich Forschungsmethoden geht es mir vor allem darum, so viel zu lernen wie möglich.						
Den Bereich Forschungsmethoden finde ich wichtig.						
Wenn ich beim Lernen besorgt bin, die Anforderung nicht zu schaffen, kann ich diesen Gedanken nur schwer von mir wegschieben.						
Ich kann zwar neue Inhalte lernen, aber richtig verstehen werde ich die Forschungsmethoden nie, dazu fehlt mir die Begabung.						
Im Bereich Forschungsmethoden geht es mir vor allem darum, meine Fähigkeiten zu steigern.						
Trotz meiner Bemühungen komme ich in Forschungsmethoden oft nicht mit und empfinde ein Gefühl der Ohnmacht.						
Mein Können im Bereich Forschungsmethoden hängt vor allem davon ab, wie viel ich lerne.						
Im Bereich Forschungsmethoden geht es mir vor allem darum, keine schlechte Note zu bekommen.						
Es fällt mir leicht, neue Stoffinhalte zu verstehen.						
Über forschungsmethodische Inhalte und Themen nachzudenken und zu reden macht mir Spaß.						
Wenn beim Lernen Versagensängste auftauchen, kann ich mich schnell beruhigen und in meiner Arbeit gut weitermachen.						

	stimme gar nicht zu	stimme nicht zu	stimme eher nicht zu	stimme eher zu	stimme zu	stimme völlig zu
Ich bin für Forschungsmethoden nicht sehr begabt – daran kann ich auch nichts ändern.						
Ich fühle mich in Forschungsmethoden oft überfordert.						
Im Bereich Forschungsmethoden geht es mir vor allem darum, dass die Anderen merken, dass ich gut bin.						

Wenn ich im Bereich Forschungsmethoden **einen Erfolg** erzielt habe, dann lag das ...

meistens an _____

manchmal auch an _____

mitunter auch an _____

	stimme gar nicht zu	stimme nicht zu	stimme eher nicht zu	stimme eher zu	stimme zu	stimme völlig zu
Im Bereich Forschungsmethoden geht es mir vor allem darum, viele neue Dinge zu lernen.						
Alle anderen Bereiche des Studiums sind mir lieber als die Forschungsmethoden.						
Wenn ich einen Misserfolg erlebe, dann versuche ich meine Fehler zu erkennen, um zu wissen, was ich das nächste Mal anders machen muss.						
Im Bereich Forschungsmethoden geht es mir vor allem darum, die Prüfung zu bestehen.						
Auch wenn Inhalte auf den ersten Blick sehr schwierig erscheinen, so weiß ich doch, dass ich alles verstehen und lernen kann, wenn ich mich ausreichend damit auseinandersetze.						
Auch wenn ich viel lerne werde ich im Bereich Forschungsmethoden nie wirklich gut sein.						

Denken Sie bitte an das vergangene Semester zurück und an Ihre **Auseinandersetzung mit der Literatur zu 'Forschungsmethoden und Evaluation'**. Welches sind Ihrer Meinung nach die wichtigsten Schritte für einen erfolgreichen Umgang mit solchen methodischen Lehrbuchtexten?

Denken Sie bitte an Ihren Arbeitsaufwand und Ihr Lernen für die Lehrveranstaltung 'Forschungsmethoden und Evaluation' im letzten Semester.

A) Wie viel Zeit haben Sie durchschnittlich **pro Modul** investiert?	während des **Semesters**:	bei der **Prüfungsvorbereitung**:
Zeit für Lesen der Literatur:	Stunden	Stunden
Gesamtzeit: *(Lesen der Lit. + alle anderen Aktivitäten!)*	Stunden	Stunden

B) Wie **häufig** haben Sie …	während des **Semesters**:					bei der **Prüfungsvorbereitung**:				
den Selbsttest jedes Moduls bearbeitet?	○ nie	○ selten	○ manch-mal	○ oft	○ sehr oft	○ nie	○ selten	○ manch-mal	○ oft	○ sehr oft
die Plattform zur Kommunikation genutzt?	○ nie	○ selten	○ manch-mal	○ oft	○ sehr oft	○ nie	○ selten	○ manch-mal	○ oft	○ sehr oft

C) Mein Lernfortschritt in Bezug auf die LV-Inhalte ist …	sehr gering	gering	mittel-mäßig	hoch	sehr hoch
D) Meine allgemeine Zufriedenheit mit der LV ist …	sehr gering	gering	mittel-mäßig	hoch	sehr hoch

Zum Abschluss noch einige wenige ***allgemeine*** Fragen,
die ***über*** das Lernen und Arbeiten im Bereich Forschungsmethoden ***hinausgehen***.

Stellen Sie sich folgende Situation vor: Studierende erhalten bei Rückgabe einer Prüfungsarbeit zusätzlich zur Zeugnisnote eine verbale Rückmeldung. Uli bekommt folgende Rückmeldung: „Sie haben sich offenbar nicht ausreichend auf die Prüfung vorbereitet!" Was bedeutet das für Motivation und Selbstwert von Uli in diesem Fach?

Bitte nennen Sie Strategien zur Verbesserung des eigenen Zeitmanagements!

Ich habe Vertrauen in meine Fähigkeiten, mein Lernen generell effektiv zu gestalten.	trifft gar nicht zu	trifft nicht zu	trifft eher nicht zu	trifft eher zu	trifft zu	trifft völlig zu

Mein Ziel im Studium ist …	gerade positiv abzuschließen	durchschnittliche Leistungen zu erbringen	gute Leistungen zu erbringen	viele sehr gute Leistungen zu erbringen

Meine allgemeine Zufriedenheit mit dem Psychologiestudium insgesamt ist …	sehr gering	gering	mittel-mäßig	hoch	sehr hoch

Ich weiß, dass ich mein Lernen allgemein gut steuern kann.	trifft gar nicht zu	trifft nicht zu	trifft eher nicht zu	trifft eher zu	trifft zu	trifft völlig zu

Meine Kompetenz im Umgang mit „Neuen Medien" (Internet, Email, Foren, etc.) schätze ich folgendermaßen ein:	sehr gut	gut	ausreichend	nicht ausreichend

Ich finde, mein Vorgehen beim Lernen ist insgesamt sehr erfolgreich.	trifft gar nicht zu	trifft nicht zu	trifft eher nicht zu	trifft eher zu	trifft zu	trifft völlig zu

Vielen Dank für das Beantworten der Fragen!

5.4 Anhang – Reliabilität der Skalen des SRL-Fragebogens

Tabelle 49: Skalen des Fragebogens zur Selbstregulation inkl. Itemanzahl und Reliabilität “Cronbachs Alpha“ nach Schober & Spiel (2005)

Ebene	Skala	Itemanzahl	Reliabilität
Metakognition	Bewertung/Reflexion	6	.83
	Regulation/Adaptation	5	.74
	Ressourcenbezogene Strategien I: Zeit- / Arbeitsmanagement	3	.81
	Ressourcenbezogene Strategien II: Soziale Ressourcen	2	.64
Kognition	Lernstrategien I: Organisieren / Strukturieren	3	.78
	Lernstrategien II: Elaboration	4	.76
	Lernstrategien III: Memorieren	4	.62
Motivation	Ausdauer	6	.67
	Interesse (Wichtigkeit & Freude)	4	.87
	Zielorientierung I: Lernziel	4	.82
	Zielorientierung II: Leistungsziel	4	.81
	Selbstwirksamkeit	4	.74
	Implizite Persönlichkeitstheorie	4	.70
	Selbstkonzept Lerner	3	.84
Emotion	Umgang mit Misserfolg	3	.69
	Emotionsregulation	3	.77
	Hilflosigkeit	3	.84

5.5 Anhang – Kategorieschema der Attributionen

Tabelle 50: Kategorieschema für die Zuordnung der Erfolgs- und Misserfolgsattributionen des freien Antwortformats

Code	Kategorie (Zuordnung)	Antwortbeispiele
1	Emotionale Zustände **intern-variabel**	Tagesverfassung positiver Einstellung / guter Tag momentane Befindlichkeit Ich hatte einen guten/schlechten Tag Ich war aufgeregt Ich hatte Angst Nervosität Unlust/ Stress
2	Externale Hilfestellung **extern-variabel**	Mir hat jemand geholfen Ich hatte Nachhilfe Art des Vortrages guten/schlechten Lehrenden/Professoren fehlender Erklärung d. Lehrenden zu wenig/stundenlangen Diskussionen mit Kollegen Falschinformationen von Kollegen Bearbeiten in der Gruppe
3	Durchführungs-charakteristika 1 **extern-variabel**	Die Aufgabenschwierigkeit in der konkreten Situation Die (konkreten) Aufgaben waren leicht/schwer Fragen waren leicht/schwer Gruppendynamik Gruppenproblematik Prüfer, Beurteiler, Beurteilung Prüfung, Prüfungsbogen unklare Prüfungsfragen Ich hatte viel Zeit bei der Aufgabenbearbeitung ausreichender Lernzeit wenig Zeit Ich hatte zuwenig Zeit zum Lernen Ich hatte zuwenig Zeit für die Vorbereitung Terminkollision, Zeitknappheit
4	Durchführungs-charakteristika 2 **intern-variabel**	Ich habe die ganze Zeit ausgenutzt zu spät mit dem Lernen begonnen Zeit gut genützt Unterschätzung des Stoffumfanges zuviel auswendig gelernt statt auf Verständnis Lernstrategie falsch lernen Zeitmanagement-Probleme Schlechte Zeiteinteilung den Stoff gut durchdacht durchgegangen nicht effektiv gelernt gut gelernt gute Vorbereitung zu wenig Zusammenhänge hergestellt

5	Anstrengung **intern-variabel**	Ich habe mich angestrengt/bemüht Ich habe viel gelernt brav gelernt (und verstanden) Ich habe während der Vorlesung aufgepasst, mitgearbeitet Ich habe die Vorlesung besucht Interesse Motivation nicht wollen sich entmutigen lassen schnelles Aufgeben Spaß an der Sache Spaß am analytischen Denken viel Zeit genommen zu wenig Zeit genommen zu wenig (genau) lernen lernen ausreichendem Lernen, Üben durcharbeiten des Stoffes/Bsp Auseinandersetzung (mit dem Stoff) Lernaufwand
6	Konzentration **intern-variabel**	Ich habe (k)einen Leichtsinnsfehler gemacht Ich habe (k)einen Rechenfehler gemacht Ich war konzentriert Ich habe mich nicht ablenken lassen Ich habe die Aufgaben genau gelesen Es war ein Missverständnis Ich war müde
7	Begabung / Fähigkeit **intern-stabil**	Ich habe den Stoff verstanden Ich bin gut in Statistik Das Fach liegt mir Der Stoff ist <u>für mich</u> leicht/schwer Die Anforderungen sind <u>für mich</u> hoch/niedrig Meiner Lernfähigkeit zu wenig Wissen Nicht-Verstehen, Nicht-Verständnis Zusammenhänge nicht verstanden
8	Zufall **external-variabel**	Ich hatte Glück/Pech Es ist das gekommen, was ich gelernt habe Passende Prüfungsfragen
9	Externale stabile Faktoren **external-stabil**	Statistik ist ein leichtes Fach Die Anforderungen sind hoch/niedrig Der Stoff ist leicht/schwer Inhalte leicht/schwer Skript gut/schlecht Zeitpunkt der Prüfung den Inhalten Stoff zu trocken, zu viele Zahlen
10	Attributionsunsicherheit	Ich weiß nicht woran es lag
11	nicht zuordenbar	mir meiner Einstellung

5.6 Anhang – Übersicht der Freischaltung 2004/05

Tabelle 51: Übersicht der Freischaltung der Module und Selbsttests mit Wochenangabe für das WS 2004/05

43	44	45	46	47	48	49	50	51	52	1	2	3	4	5
M 1														
	S 1													
		M 2												
			S 2											
				M 3										
					S 3									
						M 4		Weihnachtsferien						
							S 4							
											M 5			
												S 5		Prüf.

Tabelle 52: Übersicht der Freischaltung der Module und Selbsttests mit Wochenangabe für das SS 2005

10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
M 6		Osterf.														
	S 6															
					M 7											
						S 7										
							M 8									
								S 8								
									M 9							
										S 9						
												M10				
													S 10			Pr.

5.7 Anhang – Korrelationsmatrizen ad Fragestellung 1h

Tabelle 53: Korrelationsmatrix der Leistungen in den Prüfungen FM&E I für die drei Kategorien Reproduktion, Verständnis und Produktion

Korrelationen (Kendall-Tau-b) für das WS		**Reproduktion**	**Verständnis**	**Produktion**
Reproduktion	*r*	1	0.150	0.246
	Sig. (2-seitig)	.	.005	.000
	N	167	167	167
Verständnis	*r*	0.150	1	0.130
	Sig. (2-seitig)	.005	.	.014
	N	167	167	167
Produktion	*r*	0.246	0.130	1
	Sig. (2-seitig)	.000	.014	.
	N	167	167	167

Tabelle 54: Korrelationsmatrix der Leistungen in den Prüfungen FM&E II für die drei Kategorien Reproduktion, Verständnis und Produktion

Korrelationen (Kendall-Tau-b) für das SS		**Reproduktion**	**Verständnis**	**Produktion**
Reproduktion	*r*	0.647	1	0.154
	Sig. (2-seitig)	.000	.	.004
	N	164	164	164
Verständnis	*r*	0.467	0.154	1
	Sig. (2-seitig)	.000	.004	.
	N	164	164	164
Produktion	*r*	0.527	0.321	0.225
	Sig. (2-seitig)	.000	.000	.003
	N	90	90	90

5.8 Anhang – Häufigkeiten der Attributionsnennungen

Tabelle 55: Häufigkeiten und Prozent der erstgenannten Erfolgsattributionen

	MZP 1	**MZP 2**	**MZP 3**
	Häufigkeit	**Häufigkeit**	**Häufigkeit**
emotionale Zustände	0 (0.0%)	0 (0.0%)	0 (0.0%)
externale Hilfestellung	12 (7.1%)	4 (2.4%)	4 (2.4%)
Durchführungscharakteristika 1	4 (2.4%)	11 (6.5%)	7 (4.2%)
Durchführungscharakteristika 2	27 (16.1%)	21 (12.5%)	28 (16.7%)
Anstrengung	77 (45.8%)	89 (53.0%)	80 (47.6%)
Konzentration	2 (1.2%)	4 (2.4%)	2 (1.2%)
Begabung, Fähigkeit	25 (14.9%)	21 (12.5%)	21 (12.5%)
Zufall	4 (2.4%)	3 (1.8%)	4 (2.4%)
Externale stabile Faktoren	0 (0.0%)	1 (0.6%)	2 (1.2%)
Attributionsunsicherheit	0 (0.0%)	0 (0.0%)	0 (0.0%)
nicht zuordenbar	12 (7.1%)	14 (8.3%)	17 (10.1%)
Fehlende Werte	5 (3.0%)	0 (0.0%)	3 (1.8%)
Gesamt	168 (100.0%)	168 (100.0%)	168 (100.0%)

Tabelle 56: Häufigkeiten und Prozent der erstgenannten Misserfolgsattributionen

	MZP 1	MZP 2	MZP 3
	Häufigkeit	Häufigkeit	Häufigkeit
emotionale Zustände	4 (2.4%)	2 (1.2%)	1 (0.6%)
externale Hilfestellung	5 (3.0%)	6 (3.6%)	1 (0.6%)
Durchführungscharakteristika 1	19 (11.3%)	20 (11.9%)	36 (21.4%)
Durchführungscharakteristika 2	21 (12.5%)	21 (12.5%)	18 (10.7%)
Anstrengung	69 (41.1%)	67 (39.9%)	76 (45.2%)
Konzentration	5 (3.0%)	9 (5.4%)	3 (1.8%)
Begabung, Fähigkeit	20 (11.9%)	19 (11.3%)	10 (6.0%)
Zufall	5 (3.0%)	2 (1.2%)	3 (1.8%)
Externale stabile Faktoren	7 (4.2%)	6 (3.6%)	3 (1.8%)
Attributionsunsicherheit	0 (0.0%)	0 (0.0%)	0 (0.0%)
nicht zuordenbar	9 (5.4%)	13 (7.7%)	12 (7.1%)
Fehlende Werte	4 (2.4%)	3 (1.8%)	5 (3.0%)
Gesamt	168 (100.0%)	168 (100.0%)	168 (100.0%)

Tabelle 57: Häufigkeiten der Erfolgsattributionen (jeweils die erste Nennung) aller 3 MZP aufgeteilt nach Trackzugehörigkeit (Track A / Track B)

	MZP 1		MZP 2		MZP 3	
Track	A	B	A	B	A	B
internal-stabil	10	15	11	10	13	8
external-stabil	0	0	1	0	1	1
internal-variabel	59	47	60	54	58	52
external-variabel	8	12	10	8	6	9
nicht zuordenbar	5	7	2	12	6	11
Fehlend	2	3	0	0	0	3
Gesamt	84	84	84	84	84	84

Tabelle 58: Häufigkeiten der Misserfolgsattributionen (jeweils die erste Nennung) aller 3 MZP aufgeteilt nach Trackzugehörigkeit (Track A / Track B)

	MZP 1		MZP 2		MZP 3	
Track	A	B	A	B	A	B
internal-stabil	10	10	9	10	6	4
external-stabil	4	3	1	5	1	2
internal-variabel	52	47	53	46	52	46
external-variabel	13	16	14	14	19	21
nicht zuordenbar	3	6	5	8	4	8
Fehlend	2	2	2	1	2	3
Gesamt	84	84	84	84	84	84

Abstract (deutsch)

Diese Diplomarbeit entstand im Kontext des internetgestützten Lehr-Lern-Konzepts "Vienna E-Lecturing" (VEL) und verfolgte mittels quasi-experimentellem Zwei-Gruppen-Prätest-Posttest1-Posttest2-Design die zwei Ziele: (1) Die Analyse der Selbsttestnutzung im Rahmen der Lehrveranstaltung "Forschungsmethoden und Evaluation" (FM&E) und die Überprüfung des Zusammenhangs der Selbsttest-nutzung mit den erbrachten Leistungen in den Lehrveranstaltungsprüfungen, sowie (2) die summative Evaluation des neu entwickelten web-basierten attributionalen Feedbacks. Für die Beantwortung der Fragestellungen wurden zwei Stichproben herangezogen. Stichprobe 1 (N = 226) umfasste nur Studierende aus dem Studienjahr 2003/04, während Stichprobe 2 (N = 168), eine hinsichtlich verschiedener Aspekte "gematchte" Stichprobe, Studierende aus den Studienjahre 2003/04 und 2004/05 beinhaltete (diese beiden Kohorten unterscheiden sich nicht signifikant in der Durchführungshäufigkeit der Selbsttests). Insgesamt wurden die Selbsttests (ST) 14 806-mal von den Studierenden der Stichprobe 1 durchgeführt. In der Matching-Stichprobe (Stichprobe 2) gibt es keine signifikanten Unterschiede in der ST-Durchführungshäufigkeit zwischen verschiedenen Lehrveranstaltungs-varianten, verschiedenen Semestern oder (bis auf eine singuläre Differenz) den einzelnen ST. Studierende, die die Selbsttests oft bis sehr oft (> 31-mal) durchgeführt haben, erreichten eine signifikant bessere Gesamtprüfungsleistung in FM&E II und in reproduktiven und produktiven Prüfungsanforderungen als jene Studierenden, die die Selbsttests gar nicht bis sehr selten (0-10mal) durchgeführt haben. Die summative Evaluation des web-basierten attributionalen Feedbacks zeigte keine signifikanten Verbesserungen in Erfolgs- und Misserfolgsattributionen, sowie in den Merkmalen "Implizite Persönlichkeitstheorie", "Selbstwirksamkeitserwartungen", "Selbstkonzept Lerner/in", "Interesse" und "Hilflosigkeit". Jedoch gibt es Hinweise darauf, dass diese Ergebnisse durch eine geringe "Mindfulness" bei der Feedbackrezeption durch die Lernenden bedingt sind.

Abstract (englisch)

This thesis emerged in the context of the internet-supported learning teaching concept “Vienna E-Lecturing“ (VEL) and pursued two aims via quasi-experimental two-group-pretest-posttest1-posttest2-design. The first aim was to analyse the use of selftests within the course “research methods and evaluation” and the examination of the connection between using the selftests and the achievement at the exam of the course. The second aim was the summative evaluation of the new developed web-based attributional feedback. For this research two samples were used. The first sample (*N* = 226) included only students from the academic year 2003/04. The second sample (*N* = 168) was a matched sample, which included students from the academic years 2003/04 and 2004/05 (These two cohorts did not significantly differ in the frequency of which the selftests were used). The results of the first aim shows that the selftests were used a total of 14 806 times in the first sample. In the matched sample there are no significant differences between the two experimental groups (track A and track B), the two semesters (winter semester and summer semester) or (except one singular difference) the ten selftests. Students, which used the selftests often or very often (> 31 times), achieved significantly better results at the exam of the course “research methods and evaluation II”, both in the total score and in the subcategories reproduction and production, than students who did not or rarely used the selftests (0 - 10 times). The summative evaluation of the attributional feedback showed that there are no significant improvements in the attributions of achievement and of failure, nor in the characteristics ”implicit theories of personality”, ”self-efficacy expectancy”, ”self-concept as a learner”, ”interest” and ”helplessness” through the attributional feedback. However there are indications that these unexpected non-significant results are due to a low mindfulness from the students at the reception of the feedback.

Hiermit bestätige ich, Michaela Stefanie Fasching, dass ich die vorliegende Arbeit in allen relevanten Teilen selbstständig durchgeführt habe.

Ich habe mich bemüht, sämtliche Inhaber/innen der Bildrechte ausfindig zu machen und ihre Zustimmung zur Verwendung der Bilder in dieser Arbeit eingeholt. Sollte dennoch eine Urheberrechtsverletzung bekannt werden, ersuche ich um Meldung bei mir.

Michaela S. Fasching

LEBENSLAUF

Name:	Michaela Stefanie Fasching
Geburtsdatum:	30.07.1975
Wohnort:	1020 Wien, Rueppgasse 10/9
Mobiltelefon:	0650 / 54 52 50 3
Email:	michaela.fasching@univie.ac.at

Ausbildung / Studium:

1981 – 1989	Volks- und Hauptschule in Pöchlarn
1989 – 1990	Einjährige Haushaltungsschule in Ybbs
1990 – 1993	Ausbildung zur Einzelhandelskauffrau bei Firma Gustav Führlinger in Melk; Lehrabschlussprüfung am 16.06.1993 mit Erfolg bestanden
1997 – 2001	Handelsakademie für Berufstätige in St. Pölten; Reifeprüfung am 15.06.2001 bestanden
2001 – 2008	Diplomstudium Psychologie an der Universität Wien

Berufliche Tätigkeiten:

08/1993 – 06/1994	kaufmännische Angestellte im ehemaligen Lehrbetrieb Firma Gustav Führlinger in Melk
06/1994 – 12/1994	kaufmännische Angestellte bei Hausmann KG in St. Pölten
01/1995 – 08/1997	kaufmännische Angestellte bei LIBRO GmbH. in St. Pölten
10/1997 – 11/1997	Telefonistin bei der Firma INOVAS in St. Pölten
01/1998 – 01/2001	bei Mövenpick-Marché in Kemmelbach, anfangs Kassiererin, dann Büroangestellte und schließlich Assistentin der Buchhaltung
06/2001 – 07/2002	geringfügig beschäftigte Büroangestellte bei Mövenpick-Marché in Wien
06/2005 – 08/2005	Praktikantin am Institut für Wirtschaftspsychologie, Bildungspsychologie und Evaluation der Fakultät Psychologie der Universität Wien
10/2005 – 02/2009	Studienassistentin am Institut für Wirtschaftspsychologie, Bildungspsychologie und Evaluation der Fakultät Psychologie der Universität Wien
06/2007 – 02/2008	Projektmitarbeiterin am Projekt “Evaluation der Eignungstests für das Medizinstudium in Österreich“

Wien, am 11.11.2008