

Martina Honsel

**Messung des
betriebswirtschaftlichen Erfolgs
von Nachwuchsleistungszentren
im deutschen Profifußball**

**Entwicklung und Anwendung eines
innovativen Evaluationskonzepts**

Martina Honsel
**Messung des
betriebswirtschaftlichen Erfolgs
von Nachwuchsleistungszentren
im deutschen Profifußball**

Martina Honsel

**Messung des
betriebswirtschaftlichen Erfolgs
von Nachwuchsleistungszentren
im deutschen Profifußball**

**Entwicklung und Anwendung eines
innovativen Evaluationskonzepts**

Dissertation 2023
angenommen von der
Fakultät III
Wirtschaftswissenschaften, Wirtschaftsinformatik und Wirtschaftsrecht
der Universität Siegen

1. Gutachter: Prof. Dr. Volker Stein (Universität Siegen)
2. Gutachter: Prof. Dr. Ralf Lanwehr (Fachhochschule Südwestfalen)

Datum der mündlichen Prüfung: 11. April 2023

Dekan der Fakultät III: Prof. Dr. Marc Hassenzahl

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie;
detaillierte Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

Impressum

Druck und Bindung:
UniPrint, Universität Siegen

Gedruckt auf alterungsbeständigem holz- und säurefreiem Papier.

Siegen 2023: *universi* – Universitätsverlag Siegen
www.uni-siegen.de/universi

ISBN 978-3-96182-151-8

doi.org/10.25819/ubsi/10360

Die Publikation erscheint unter der
Creative Commons Lizenz CC-BY-SA



„Ich habe keine besondere Begabung,
sondern bin nur leidenschaftlich neugierig.“

Albert Einstein (1952)

Vorwort

Was für eine Reise!

Diese Reise beginnt mit dem Besuch des Gymnasiums in den 1970er Jahren als „Arbeiterkind“ und anschließendem Studium der Wirtschafts- und Sozialwissenschaften an der Universität Dortmund. In meiner letzten mündlichen Prüfung zur Diplomkauffrau 1986 in Logistik legte Herr Professor Dr. Dr. Dr. h.c. Reinhardt Jünemann das Saatkorn „Promotion“. Es schlummerte 30 Jahre, während ich meine Karriere in einem global agierenden Unternehmen der Mineralölindustrie in diversen Rollen – auf verschiedenen Hierarchieebenen und Arbeitsgebieten – verfolgte. Es war eine arbeitsintensive und tolle Zeit mit zahlreichen beruflichen Erfahrungen, gesammelt u. a. in Projektmanagement, Führung und Controlling im Umfeld der IT. Ich bin sehr dankbar für diese Zeit und die Menschen, die ich treffen konnte, haben sie mich doch zu der Person gemacht, die ich heute sein darf: Wissenschaftlerin mit „keiner besonderen Begabung, aber leidenschaftlich neugierig“.

Als wissenschaftliche Mitarbeiterin und Lehrbeauftragte für besondere Aufgaben arbeitete ich seit 2017 im Fachbereich Ingenieur- und Wirtschaftswissenschaften im Team von Prof. Dr. Ralf Lanwehr – ihm gilt mein besonderer Dank, gab er mir doch die Chance, das Saatkorn zu entwickeln und die Pflanze entstehen zu lassen. Aufgrund seines motivierenden Führungsstils, der von ihm vertrauensvoll eingeräumten Freiheiten und Freiräumen in Forschung und Lehre sowie seiner charismatischen Art, Umsicht und insbesondere Begeisterungsfähigkeit arbeitete ich jeden Tag mit großer Freude an der Fachhochschule Südwestfalen. Als mein Vorgesetzter fand er stets die richtige Balance zwischen Förderung und Forderung meiner Person und meiner individuellen Leistung und ermöglichte mir so zusammen mit Prof. Dr. Volker Stein als Doktorvater meine kooperative Promotion. Gemeinsam begleiteten mich beide konstruktiv in den Phasen des Dissertationsprojektes. Zudem begleiteten diverse Schulungen des House of Young Talents der Universität Siegen, des Promotionskollegs NRW sowie des Forschungscoachings unter der Leitung von PD Dr. Dr. Guido Strunk meinen gesamten Promotionsprozess als Nachwuchswissenschaftlerin. Dieser ganze Prozess hat mich inspiriert, gefördert und geprägt, und allen gilt mein aufrichtiger Dank. Die Fachhochschule Südwestfalen hat mich nicht nur durch die Finanzierung und auch Ausrichtung von Schulungen und des Forschungscoachings unterstützt, sondern mir persönlich auch zeitliche Ressourcen eingeräumt. Herzlichen Dank dafür! Ohne abendfüllende Diskussionen mit Thorsten Fluck über die Qualität der Nachwuchsförderung im deutschen Fußball, in der ein leidenschaftlicher Fußballtrainer auf mich als eine Controllerseele traf, wäre die Idee zu dieser Arbeit nie entstanden. Herzlichen Dank dafür!

Außergewöhnliche Hilfsbereitschaft im Bereich Data Science und Analytics durfte ich durch meinen ebenfalls promovierenden Kollegen Viktor Wolf (M. A.) und meinen Sohn und gleichzeitig promovierenden Kollegen an der FH Südwestfalen Felix Neubürger (M. Sc.) erfahren. Das hat sehr zum Gelingen dieser Dissertation beigetragen. Ebenso hilfsbereit waren meine Kolleg*innen aus unserem Lehrstuhl-Team „Führung

& Transformation“: immer ein offenes Ohr auch in schweren Zeiten, viele konstruktive und oft auch heitere Diskussionsrunden bis hin zu abendlichen ablenkenden Aktivitäten. Darüber hinaus gaben mir Promotionskolleg*innen der Uni Siegen wertvolles Feedback. Sie haben insbesondere in der Endphase der Promotion einen sehr wertvollen inhaltlichen Austausch ermöglicht und zur Abrundung der Dissertationsschrift beigetragen. Herzlichen Dank dafür! Besonders danken möchte ich auch meinem privaten Umfeld mit all meinen Freunden. Nachdem sich das Erstaunen über mein Ziel gelegt hatte, konnte ich mir ihres Zuspruchs und ihrer Ermutigung stets sicher sein.

Die größte Anerkennung gebührt meiner Familie: meinen Eltern Erwin (†) und Edith Trinat, meinem Mann Peter sowie meinen Söhnen Jonas und Felix Neubürger. Insbesondere mein Mann hat mich über die letzten Jahre immer wieder motiviert, unterstützt, hilfreich und insbesondere geduldig begleitet – das war ohne Frage nicht immer einfach. Meine Söhne zeichneten sich durch liebevolle Unterstützung „dieser verrückten Idee ihrer Mutter“ und der familieninternen Challenge „Wer ist das erste Familienmitglied mit einem Dokortitel?“ aus. Vielen, vielen Dank! Diese Arbeit widme ich Euch!

Was für eine Reise: Arbeiterkind, Uniabsolventin, Trainee, Zwillingismutter, Managerin, „journalistisch schreibende Nachwuchswissenschaftlerin“ bis zur wortsicheren Wissenschaftlerin. Ich bin gespannt, welche Stationen das Leben noch für mich bereithält – neue Aufgaben im Wissenschaftsmanagement am Promotionskolleg für angewandte Forschung NRW liegen aktuell als Nächstes an.

Meschede, November 2022

Martina Honsel

In dieser Dissertationsschrift wird das generische Maskulinum verwendet, da der Fokus der Arbeit in der Untersuchung von Nachwuchsleistungszentren im Profifußball der Männer liegt. Weibliche und anderweitige Geschlechteridentitäten werden lediglich benutzt, soweit es für eine Aussage erforderlich ist.

Teile der Monographie stellen überarbeitete und erweiterte Ausschnitte des Artikels „Die Evaluation von Qualität – Ein Methodenvergleich anhand von Nachwuchsleistungszentren im deutschen Profifußball“ von Lanwehr, Honsel & Wilms (Publikation Januar 2021) dar.

Inhaltsverzeichnis

VORWORT	VI
INHALTSVERZEICHNIS	IX
ABBILDUNGSVERZEICHNIS	XIII
TABELLENVERZEICHNIS	XVI
FORMELVERZEICHNIS	XVII
ABKÜRZUNGSVERZEICHNIS	XVIII
1 EINLEITUNG	1
1.1 Problemstellung	1
1.2 Stand der Forschung (Überblick)	4
1.3 Ziel der Arbeit	6
1.4 Aufbau der Arbeit	8
2 BEGRIFFLICHE GRUNDLAGEN UND FORSCHUNGSSTAND	10
2.1 Nachwuchsausbildung im deutschen Profifußball	10
2.1.1 Nachwuchsleistungszentren.....	10
2.1.2 Abgrenzung Untersuchungsgegenstand	12
2.1.3 Rahmenbedingungen für die Nachwuchsarbeit	14
2.1.3.1 Internationale Rahmenbedingungen für die Nachwuchsarbeit ...	14
2.1.3.1.1 Ausbildungsentschädigung und Solidaritätsbeitrag	14
2.1.3.1.2 Homegrown Player Rule	15
2.1.3.1.3 Finanzielles Fairplay	17
2.1.3.2 Nationale Rahmenbedingungen der Nachwuchsausbildung in Deutschland	21
2.1.3.2.1 Talententwicklung in Deutschland auf individueller Ebene	21
2.1.3.2.2 Organisatorische Ebene – Das Nachwuchsleistungszentren	22
2.1.3.2.3 Kontextuelle Ebene – Das ökonomische Umfeld der Nachwuchsleistungszentren	24
2.2 Evaluationsforschung	28
2.2.1 Evaluation – Begriffliche Klärung	29
2.2.2 Evaluationsdesign	30
2.2.2.1 Evaluationszielsetzung.....	30

2.2.2.2 Evaluationsgegenstand.....	33
2.2.2.3 Evaluationsstakeholder	33
2.2.2.4 Evaluationstyp.....	33
2.2.2.5 Evaluationszeitpunkt.....	34
2.2.2.6 Evaluationskriterien	35
2.2.3 Evaluationsstandards	36
2.2.4 Evaluation von Personalentwicklungsmaßnahmen	39
2.2.4.1 Modelle für die Evaluation von Personalentwicklungsmaßnahmen.....	41
2.2.4.1.1 Das Vier-Ebenen Modell von Kirkpatrick	41
2.2.4.1.2 Brinkerhoff's six-stage model (1987).....	44
2.2.4.1.3 IPO-Modell (Bushnell, 1990)	44
2.2.4.1.4 CIPP-Modell: Stufflebeam (1971).....	45
2.2.4.1.5 CIRO-Modell: Warr, Bird und Rackham (1970).....	45
2.2.4.1.6 Fünf-Level-Modell: Kaufman & Keller (1994).....	45
2.2.4.1.7 ROI-Theorie und -Modell: Phillips (1998).....	45
2.2.4.1.8 VOI: Kellner (2006).....	46
2.2.4.2 Einordnung der Modelle im Hinblick auf das Ziel der Arbeit.....	46
2.2.5 Evaluation der Resultate von Personalentwicklungsmaßnahmen	51
2.2.5.1 Messen des Ergebnisses von Personalentwicklungsmaßnahmen	51
2.2.5.2 Evaluationskriterien für Personalentwicklungsmaßnahmen.....	53
2.2.5.3 ROI für Personalentwicklungsmaßnahmen	57
2.3 Erfolgsfaktorenforschung	58
2.4 Evaluation des Erfolgs im Fußball.....	62
2.4.1 Definition von Erfolg	62
2.4.2 Der sportliche Erfolg von Vereinen	63
2.4.3 Der ökonomische Erfolg von Vereinen.....	64
2.4.4 Erfolg in der Nachwuchsausbildung	67
3 EMPIRISCHE UNTERSUCHUNG – BETRIEBSWIRTSCHAFTLICHER ERFOLG VON NACHWUCHSLEISTUNGSZENTREN	71
3.1 Mentales Modell.....	71

3.2 Planung der Evaluation	74
3.3 Forschungsdesign	76
3.3.1 Wissenschaftstheoretischer Ansatz der Studie	76
3.3.2 Erkenntnisziel und Gegenstand der Studie.....	77
3.3.3 Datengrundlage der empirischen Studie.....	78
3.3.4 Erkenntnisinteresse der Studie.....	79
3.3.5 Anzahl der Untersuchungszeitpunkte und -objekte der Studie	79
3.4 Operationalisierung	80
3.5 Durchführung der Datenerhebung	81
3.5.1 Erhebung der Sekundärdaten.....	83
3.5.1.1 Daten von Transfermarkt.de.....	83
3.5.1.1.1 Spieler- und vereinsbezogene Daten	85
3.5.1.1.2 Datenqualität der spieler- und vereinsbezogenen Daten von Transfermarkt.de	85
3.5.1.1.3 Marktwert.....	89
3.5.1.2 Double Pass Scores	94
3.5.2 Erhebung der Primärdaten	96
3.5.2.1 Rohergebnis Verein.....	96
3.5.2.2 Budget 2019	99
3.5.3 Güte der Datenerhebung.....	101
3.6 Ermittlung der Datenkonstrukte	102
3.6.1 Investment	102
3.6.1.1 Ermittlung Prozentualer Anteil Investment in ein NLZ am Rohergebnis.....	102
3.6.2.2 Verlässlichkeit des Konstrukts Investment in Nachwuchsleistungszentren	104
3.6.2 Nutzen.....	109
3.6.3 Bewertung des Erfolgs von Nachwuchsleistungszentren mittels ROI.....	118
3.6.4 Evaluation der Evaluation der Nachwuchsleistungszentren (Metaevaluation).....	118
3.7 Zwischenfazit Untersuchungsrahmen	119
4 ERGEBNISSE UND DISKUSSION DER EVALUATION	121

4.1 Ergebnisse – Deskriptive Aussagen.....	121
4.1.1 Kosten.....	121
4.1.2 Nutzen	124
4.1.2.1 Vereinsqualität	124
4.1.2.2 Individuelle Spielerqualität	128
4.1.2.3 Hervorgebrachte Spielerqualität	132
4.1.2.4 Nutzen – Monetär bewertete Hervorgebrachte Spielerqualität..	135
4.1.3 Return on Investment Nachwuchsleistungszentrum	137
4.2 Ergebnis zur Forschungsfrage.....	141
4.3 Metaevaluation gemäß der Standards der DeGEval.....	141
4.3.1 Nützlichkeit.....	141
4.3.2 Durchführbarkeit.....	143
4.3.3 Fairness	143
4.3.4 Genauigkeit	144
5 ABSCHLUSSDISKUSSION UND AUSBLICK.....	148
5.1 Zusammenfassung	148
5.2 Limitationen	150
5.3 Implikationen für die Evaluation von Personalentwicklungsmaßnahmen .	153
5.4 Implikationen für die Evaluation von Nachwuchsleistungszentren	154
5.5 Implikationen für und Ausblick auf die zukünftige Forschung.....	159
LITERATURVERZEICHNIS	162

Abbildungsverzeichnis

Abbildung 1: Kirkpatrick-Modell plus eigene Erweiterungen. Eigene Darstellung nach D. L. Kirkpatrick, 1967; 1994.	4
Abbildung 2: Aufbau der Arbeit	8
Abbildung 3: Betrachtete Theorien und Forschungsbereiche.....	9
Abbildung 4: Entwicklung rund um die Nachwuchsleistungszentren 2001 bis heute. Eigene Darstellung nach (DFL, 2021b).....	11
Abbildung 5: PASS 4.0 Qualitätsdimensionen. Eigene Darstellung auf Basis der zur Verfügung gestellten Double-Pass- Daten.	13
Abbildung 6: Ausbildungsentschädigungen und Solidaritätsbeiträge weltweit. Eigene Darstellung (FIFA, 2017, 2018, (2019a; FIFA, 2020, 2021).....	15
Abbildung 7: Anteil einheimischer und ausländischer Spieler in der 1. Bundesliga bis Saison 2019/20. Eigene Darstellung (Statista, 2021a).....	19
Abbildung 8: Anteil ausländischer Spieler in den BIG-5-Ligen: Deutschland, England, Spanien, Frankreich und Italien in der Saison 2021/2022. Eigene Darstellung (Statista, 2021b).....	20
Abbildung 9: Einflussfaktoren auf die Entwicklung eines Fußballspielers. Eigene Darstellung nach Seidel (2011).....	22
Abbildung 10: Talentförderung DFB. Eigene Darstellung nach (DFB, 2021c).....	23
Abbildung 11: Investitionen in Nachwuchsarbeit 1. und 2. Bundesliga. Eigene Darstellung (DFL, 2020a).	24
Abbildung 12: Umsatzentwicklung 1. und 2. Bundesliga und ihre Investitionen in den Nachwuchsbereich in Milliarden Euro. Eigene Darstellung (DFL, 2020a).	25
Abbildung 13: Ausbildungsentschädigungen und Solidaritätsbeiträge pro Saison 1. und 2. Fußball-Bundesliga. Eigene Darstellung (DFL, 2016; DFL, 2017; DFL, 2018; DFL, 2019b; DFL, 2020a; DFL, 2021a).....	26
Abbildung 14: Verteilungsschlüssel Medienerlöse	27
Abbildung 15: Programmevaluation Zusammenhänge. Eigene Darstellung nach Stockmann & Meyer (2014).....	31
Abbildung 16: Evaluationsarten in Abhängigkeit von Zielsetzung und Zeitpunkt. eigene Darstellung nach Mittag und Hager (2000) sowie Kauffeld (2016).....	35
Abbildung 17: DeGEval-Standards und ihre Anwendbarkeit auf die zehn Hauptaufgaben für eine Evaluation. Eigene Darstellung nach DeGEval (2017b).....	38

Abbildung 18: Häufigkeit von Evaluationen der Ebenen des Kirkpatrick Modells in Praxis und Wissenschaft im Vergleich. Eigene Darstellung nach van Buren und Erskine (2002) und Arthur et al. (2003).....	47
Abbildung 19: Das Vier-Ebenen-Modell von Kirkpatrick, erweitert um die Ebene ROI durch Phillips mit Angaben zur Häufigkeit der Messung der Erfolgsmaße in Organisationen (Praxis). Eigene Darstellung nach Phillips (2003), Kirkpatrick (1967, 1994), van Buren und Erskine (2002) und Döring und Bortz (2016).	48
Abbildung 20: The New World Kirkpatrick Model. Eigene Darstellung nach J. D. Kirkpatrick & Kirkpatrick W., 2016, S. 11.	50
Abbildung 21: Ansätze zur Evaluation von Personalentwicklungsmaßnahmen. Eigene Darstellung nach J. Müller (2017) und J. Littkemann et al. (2012).	52
Abbildung 22: Human-Capital Bewertung-Entscheidungsmatrix. Eigene Darstellung nach Scholz et al. (2011).	54
Abbildung 23: Methoden der Erfolgsfaktorenforschung. Eigene Darstellung nach Haenecke (2002).....	60
Abbildung 24: Methoden zur Identifikation von Erfolgsfaktoren. Eigene Darstellung nach Haenecke (2002).	61
Abbildung 25: Einnahme-Mix 1. und 2. Bundesliga in der Saison 2018/19. Eigene Darstellung nach DFL (2020a).....	65
Abbildung 26: Übersicht verwendete unabhängige Variablen im Notenmodell. Eigene Darstellung nach Eschweiler und Vieth (2004).....	66
Abbildung 27: Mentales Modell – ROI für Nachwuchsleistungszentren.	73
Abbildung 28: Kriterien zur Planung der Evaluation des betriebswirtschaftlichen Erfolgs von Nachwuchsleistungszentren.....	75
Abbildung 29: Überblick über das Forschungsdesigns der vorliegenden Studie. Eigene Darstellung nach Döring und Bortz (2016).	77
Abbildung 30: Überblick Merkmale des Forschungsdesigns.....	80
Abbildung 31: Überblick Evaluationsprozess und Operationalisierung ROI für Nachwuchsleistungszentren.	81
Abbildung 32: Überblick Datenherkunft und Kategorisierung.....	82
Abbildung 33: Abgrenzung Einflussfaktoren auf den Marktwert und Transferwert. Eigene Darstellung nach Transfermarkt.de (2013).....	93
Abbildung 34: In Investment Nachwuchsleistungszentrum eingehende Finanzdaten.....	96
Abbildung 35: Umsatzentwicklung und Investitionen in den Nachwuchsbereich 1. und 2. Bundesliga in Mrd. Euro und prozentualer Anteil der Investitionen am Umsatz. Eigene Darstellung nach DFL (2020a).....	104

Abbildung 36: Überblick Konstrukt „Nutzen für ROI für NLZ“	109
Abbildung 37: Vereinsqualitäten im Vergleich zur Tabellensituation: 1. Bundesliga, 2. Bundesliga, 3., 4. und 5. Liga in Deutschland, La Liga, 2. und 3. Spanische Liga, Premier League, 2. Liga England über 16 Saisons. Nach Lanwehr et al. (2021).	110
Abbildung 38: Streudiagramm von Individuelle Spielerqualität und Marktwerte der Spieler.	113
Abbildung 39: Zusammenhang Hervorgebrachte Spielerqualität NLZ und Double Pass Scores. Eigene Darstellung.	115
Abbildung 40: Entwicklung Mittelwert und Median Spielerqualität und Marktwert pro Jahr	117
Abbildung 41: Cluster-Kosten NLZ in Mio. Euro.	122
Abbildung 42: Mittelwerte Vereinsqualität per Liga in Deutschland im Zeitverlauf	125
Abbildung 43: Mittelwert Vereinsqualität pro Verein in der 1. Bundesliga im Betrachtungszeitraum (1999-2020).....	126
Abbildung 44: Zusammenhang Mittelwert Vereinsqualität zu Anzahl Jahre zugehörig zur 1. Bundesliga.....	127
Abbildung 45: Mittelwert Vereinsqualität pro Verein in der 2. Bundesliga im Betrachtungszeitraum (1999-2020).....	129
Abbildung 46: Zusammenhang Mittelwert Vereinsqualität zu Anzahl Jahre zugehörig zur 2. Bundesliga.....	130
Abbildung 47: Anzahl Spieler, Median und Mittelwert Individueller Spielerqualitäten im Betrachtungszeitraum.....	131
Abbildung 48: Anzahl Nachwuchsleistungszentren, Median und Mittelwert Hervorgebrachte Spielerqualitäten in Betrachtungszeitraum	132
Abbildung 49: Nutzengenerierung: 49 Nachwuchsleistungszentren 2009-2019	136

Tabellenverzeichnis

Tabelle 1: Überblick über die beschriebenen internationalen Rahmenbedingungen. Eigene Darstellung nach Wallebohr (2019).....	18
Tabelle 2: Übersicht der gängigen output-orientierten Indikatoren und eine Einschätzung ihrer Eignung zur Messung des Erfolgs von Nachwuchsleistungszentren. Eigene Darstellung nach Lanwehr et al. (2021).....	70
Tabelle 3: Abgleich Transfermarkt.de mit polnischen fußballbezogenen Webseiten: Anzahl überprüfter Daten, ermittelter fehlerhafter Daten und deren Anteile in Prozent sowie die Konfidenzober- und Untergrenzen des Fehlerquotienten bei Annahme der Poisson-Verteilung getrennt nach überprüfter Datensatzkategorie. Eigene Darstellung nach Kasparek (2020).	88
Tabelle 4: Übersicht über die Vollständigkeit der für die Studie erstellten Datenbank auf Relationenebene; Eigene Darstellung nach Strauß (2020).	89
Tabelle 5: Deskriptive Statistik Double Pass Scores nach PASS-Zyklen.....	95
Tabelle 6: Übersicht zur Datenerhebung für die vorliegende Studie.....	103
Tabelle 7: Deskriptive Statistik Datensatz Prozentualer Anteil Budget 2019 und Rohergebnis Verein mit NLZ 2018/19	107
Tabelle 8: Deskriptive Statistik PASS-Zyklen und hervorgebrachte Spielerqualität (SQ_NLZ).....	117
Tabelle 9: Marktwertfaktor zur Bewertung der hervorgebrachten Spielerqualität.....	117
Tabelle 10: Kosten Nachwuchsleistungszentren in Mio. EURO pro Liga in 2018/19... ..	122
Tabelle 11: Kosten Nachwuchsleistungszentren sortiert nach Ligazugehörigkeit in 2018/2019.	123
Tabelle 12: Deskriptive Statistik Vereinsqualität per Liga in Deutschland.	124
Tabelle 13: Deskriptive Statistik Individuelle Spielerqualität	131
Tabelle 14: Deskriptive Statistik Hervorgebrachte Spielerqualität.....	133
Tabelle 15: Ranking NLZ nach Summe Hervorgebrachter Spielerqualität über Betrachtungszeitraum, Mittelwert per Jahr und Summen per Pass-Zyklen.	134
Tabelle 16: Nutzen Nachwuchsleistungszentren in Euro.	137
Tabelle 17: Vereine mit Nachwuchsleistungszentrum, Kosten- und Nutzenkalkulation.	139
Tabelle 18: ROI der Vereine mit Nachwuchsleistungszentrum.	140
Tabelle 19: Metaevaluation anhand der Checkliste zur Anwendung der Standards (DeGEval, 2017a).	147

Formelverzeichnis

Formel 1: ROI für Personalentwicklungsmaßnahmen.....	57
Formel 2: Berechnung Vereinsqualität – VQ. Eigene Darstellung nach Lanwehr et al. (2021, S. 31).	111
Formel 3: Berechnung Individuelle Spielerqualität. Eigene Darstellung, Weiterentwicklung nach Lanwehr et al. (2021, S. 31).....	112
Formel 4: Berechnung Hervorgebrachte Spielerqualität NLZ. Eigene Weiterentwicklung nach Lanwehr et al. (2021) 114	
Formel 5: ROI für Nachwuchsleistungszentren (NLZ).....	118

Abkürzungsverzeichnis

CIES	Centre International d'Etude du Sport
COVID-19	Corona Virus Disease 2019
DeGEval	Deutsche Gesellschaft für Evaluation e.V.
DFB	Deutscher Fußball-Bund
DFL	Deutsche Fußball Liga
ebd.	ebenda
ECA	European Club Association
ERS	Evaluation Research Society
FFP	Finanzielles Fairplay
FIFA	Fédération Internationale de Football Association
HCROI	ROI of Human Capital
ICC	Intraklassenkorrelation
IVW	Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e.V.
JCSEE	Joint Committee on Standards for Educational Evaluation
Mio	Millionen
MLS	Major League Soccer
Mrd	Milliarden
PASS	Professional Academy Support System
PIMS	Profit Impact of Marketing Studies
RODI	Return on Development Investment
ROI	Return on Investment
SpielerID	Spieler-Identifikationsnummer
Tsd.	Tausend
U23	Mannschaft mit Spielern unter 23 Jahren
UEFA	Union of European Football Associations
VereinsID	Vereins-Identifikationsnummer

1 Einleitung

1.1 Problemstellung

Bei der Europameisterschaft 2000 schied die deutsche A-Nationalmannschaft der Männer in der Vorrunde aus. Daraufhin etablierte der Deutsche Fußball-Bund (DFB) in Zusammenarbeit mit der Deutschen Fußball Liga e. V. (DFL) in Deutschland ein neues Konzept der Talentförderung. Im Jahr 2001 wurden dazu *Nachwuchsleistungszentren* eingeführt und sie wurden für die Vereine der ersten beiden Bundesligen verpflichtend (DFL, 2021b). Die Nachwuchsleistungszentren werden von den Vereinen mit dem Ziel geführt, die besten Talente durch eine systematische Förderung auf spätere Anforderungen im Lizenzfußball und damit auch für einen eventuellen Einsatz in der Nationalmannschaft vorzubereiten (DFB, 2020; Lanwehr et al., 2021). Zu dem Zeitpunkt der Errichtung der Nachwuchsleistungszentren – in Veröffentlichungen der DFL wird der Begriff Leistungszentrum synonym verwendet – lag der Anteil deutscher Spieler in den drei oberen Ligen des deutschen Profifußballs bei 62 %. Der Anteil der deutschen Spieler im Lizenzfußball hat seitdem um 8 % zugenommen (Transfermarkt.de, 2021a) und liegt damit 2021 bei 70 %. Die DFL interpretiert dies als Erfolg des eingeführten Konzepts (DFL, 2021b).

Anhand einer im Konzept enthaltenen Strategie zur regelmäßigen Evaluation der vermittelten Kompetenzen, die auf zwei Ebenen durchgeführt wird, wurde durch DFL und DFB auch eine Qualitätsmessung und -sichtung etabliert. Der Aufbau der erforderlichen Funktionen und Strukturen in einem Nachwuchsleistungszentrum wird auf der ersten Ebene über eine formale Lizenzierung sichergestellt. Die zweite Ebene, eine Zertifizierung, quantifiziert, welche Kompetenzen in welcher Qualität im Nachwuchsleistungszentrum aufgebaut wurden (DFB, 2020; Lanwehr et al., 2021). Diese Evaluierung der Kompetenzen erfolgt in mehreren Schritten – Selbstevaluation durch die Vereine, mehrtägige Audits durch Externe vor Ort im Nachwuchsleistungszentrum, Berichterstattung durch die Auditor*innen, Vergabe von Scores und Sternen als Auszeichnung –, womit versucht wird, ein gewisses Maß an Glaubwürdigkeit und Objektivität der Ergebnisse zu erreichen. Diese Kombination der Selbst- und die Fremdevaluation wird auf Seiten der Vereine allerdings als erheblicher Aufwand empfunden, da in den Vereinen für die Zertifizierung eines Nachwuchsleistungszentrums teilweise über 1000 Seiten an Dokumenten produziert werden (Lanwehr et al., 2021; Kroemer, 2015).

Insgesamt vier Evaluationszyklen in 13 Jahren (von 2007 bis 2019) haben dafür gesorgt, dass die innerhalb dieser Phase entwickelten Kompetenzen und Routinen der Nachwuchsleistungszentren den vorgegebenen Kriterien und Idealvorstellungen des DFB bzw. der DFL im Wesentlichen entsprechen. Das ist ein großer Erfolg für das Konzept und die Entwicklung der für die Ausbildung von Talenten für den Lizenzfußball bzw. die Nationalmannschaft benötigten Kompetenzen. Mit ihrer Arbeit tragen die Nachwuchsleistungszentren dazu bei, die beruflichen und persönlichen Qualifikationen der

Talente auszubilden und zu sichern, ihre Interessen zu berücksichtigen und ihre Kompetenzen im Interesse des Vereins bzw. der Fußballverbände auszubauen. Dadurch, dass die Nachwuchsleistungszentren sich bezüglich der erforderlichen Kompetenzen in den untersuchten Qualitätsdimensionen sehr angezogen haben, ist allerdings eine ausgeprägte Differenzierung dieser Zentren zur Bewertung der Güte des Nachwuchsleistungszentrums mit dem zugrundeliegenden Scoring-Modell nicht in ausreichendem Maße möglich.

Die während der COVID-19-Pandemie getroffenen Maßnahmen, wie der Abbruch bzw. die Unterbrechung der Wettbewerbe und sogenannte Geisterspiele, führten zu hohen Einnahmeverlusten bei den Vereinen und zu einer Veränderung des Blickwinkels auf die betriebswirtschaftliche Seite des Erfolgs. In einer ersten Studie zu den Auswirkungen der COVID-19-Pandemie 2020 konstatieren Zülch et al. (2020), dass die finanzielle Situation bei vielen Vereinen der deutschen Bundesliga deshalb bedrohlich ist, weil der Strategieformulierung und -operationalisierung wenig Wert beigemessen wurde. Bei den Finanzen, im Management und insbesondere in der Nachwuchsarbeit sollten die Bundesligavereine daher professioneller werden (Zülch, 2020). Ein professionelles Management und Controlling sollte in den Strategie- und Haushaltsplanungen alle Investitionen – auch die in die Nachwuchsarbeit – hinsichtlich des betriebswirtschaftlichen Erfolgs bewerten. Die hohe Professionalität ihrer Arbeit kann den Nachwuchsleistungszentren mit der bisherigen Zertifizierung bescheinigt werden und sportlicher Erfolg der Nachwuchsmannschaften zeigt sich anhand der Tabellenstände in den Jugendligen.

Personalentwicklung in Unternehmen ist gekennzeichnet durch „die Förderung beruflich relevanter Kenntnisse, Fertigkeiten, Einstellungen“ (Solga et al., 2011, S. 19). Für eine strategierorientierte Personalentwicklung werden Maßnahmen zur Entwicklung von Kompetenzen ergriffen, die der Verwirklichung der strategischen Unternehmensziele dienen (Solga et al., 2011). Die Entwicklung von Talenten für den Lizenzfußball in einem Nachwuchsleistungszentrum ist einer Personalentwicklung gleichzusetzen, sie sollte daher sämtliche Maßnahmen zur systematischen Förderung der beruflichen Handlungskompetenz von Menschen beinhalten, die im Lizenzfußball – im Ausbildungsverein oder in anderen Vereinen – tätig sein werden.

Evaluationen von Personalentwicklungsmaßnahmen sind erforderlich, um Daten zur Beurteilung und Optimierung der Personalentwicklung zu schaffen (Solga, 2011). Mit betriebswirtschaftlichen Kennzahlen soll Transparenz geschaffen werden, die als Entscheidungsgrundlage bei zukünftigen Ressourcenverteilungen und Entwicklung der Unternehmensstrategie dienen soll (Phillips & Schirmer, 2008). Evaluationen von Personalentwicklungsmaßnahmen umfassen sämtliche Aktivitäten zur Qualitäts- und Erfolgskontrolle der durchgeführten Maßnahmen. Sie bewerten nicht nur die inhaltliche Konzeption und die Durchführung der Personalentwicklungsmaßnahmen, sondern vermehrt auch die Wirksamkeit und die Wirtschaftlichkeit des Programms (Solga, 2011). Die von Zülch (2020) angemahte Professionalität bei Finanzen, Management und der Nachwuchsarbeit erfordert eine ebensolche Evaluation der Maßnahmen im Nachwuchsleistungszentrum. Um den Erfolg von Nachwuchsleistungszentren professionell

und wirtschaftlich zu bewerten, ist es erforderlich, ein Evaluationskonzept zu entwerfen, das wissenschaftlichen Qualitätsstandards entspricht und die in der Praxis häufig genannten Grenzen einer Evaluation – wie knappe zeitliche, personelle und monetäre Ressourcen, mangelnde Unterstützung durch das Management, Schwierigkeiten mit einer validen Datenerhebung im betrieblichen Alltag – überwindet (Solga, 2011).

Um den Erfolg von Nachwuchsleistungszentren zu messen, bietet sich eine summative Evaluation an, die die Nachwuchsarbeit nicht prozessbegleitend, sondern abschließend bewertet. Mit der summativen Evaluation wird der Schwerpunkt auf die Evaluation des *Outputs* gelegt, Wirksamkeit und Effizienz können analysiert werden (Solga, 2011). Um die Wirksamkeit einer Personalentwicklungsmaßnahme zu ermitteln, sind geeignete Evaluationskriterien zu identifizieren, die es ermöglichen, das realisierte Ausmaß der Zielerreichung abzulesen (operationale Erfolgsindikatoren, Solga, 2011). Eine der Herausforderungen bei der Auswahl geeigneter operationaler Erfolgsindikatoren besteht darin, ein geeignetes Maß für individuelle Arbeitsleistungen zu identifizieren und gegebenenfalls zu konstruieren sowie die Maße in monetäre Werte zu konvertieren. (Kraiger, 2002; Phillips & Schirmer, 2008). In der Welt des Fußballs liegen massenhaft Daten für den Großteil aller (internationalen) Wettbewerbe vor, die die individuelle (Arbeits-)Leistung repräsentieren oder die Basis für konstruierte Maße bilden können. Die Konvertierung dieser Leistungen zu monetären Werten kann über die ebenfalls vorhandenen Marktwertbewertungen der Individuen bzw. Transferzahlungen erfolgen. Aufgrund dieser vorhandenen Daten bietet sich die Evaluation der Nachwuchsleistungszentren als Anwendungsbeispiel für die Evaluation von Personalentwicklungsmaßnahmen an. Die dafür benötigte Entwicklung einer innovativen Evaluationsmethode und die Identifizierung valider Erfolgsfaktoren für die Bewertung der Nachwuchsleistungszentren werden insbesondere die Evaluations- und Erfolgsfaktorenforschung im Fußball verbessern, da damit das Management der Vereine mit wichtigen betriebswirtschaftlichen Daten für strategische Entscheidungen versorgt werden kann und eine weitere Differenzierung der Nachwuchsleistungszentren – neben den Double Pass Scores und verliehenen Sternen – untereinander ermöglicht wird.

Die Evaluationsforschung betont die Notwendigkeit wissenschaftlicher Qualitätsstandards für eine fundierte valide Evaluation (Hager et al., 2000). Daher werden bei der Entwicklung des innovativen Evaluationskonzepts für Nachwuchsleistungszentren die Ansprüche an Evaluationsprozesse gemäß der Deutschen Gesellschaft für Evaluation (DeGEval, 2017) – Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit – erfüllt. Die entwickelte Evaluationsmethode wird einer Metaevaluation unterzogen.

1.2 Stand der Forschung (Überblick)

Für die vorliegende Untersuchung werden folgende Forschungsbereiche betrachtet: Evaluationsmodelle in der Personalentwicklung, Erfolgsfaktorenforschung mittels Investitionskennzahlen in der Personalentwicklung und die Evaluationsforschung im Profifußballsport. Spiel (2003) berücksichtigt bei ihrer Definition der Evaluationsforschung die Gemeinsamkeiten der in der Fachliteratur aufgeführten Definitionen, nämlich die systematische Bewertung unter Verwendung wissenschaftlicher Methoden: „Evaluations-Forschung untersucht wissenschaftsgestützt unter Berücksichtigung geltender Standards [...] die Effektivität (Ausmaß der Zielerreichung) und Effizienz (Verhältnis von Aufwand und Nutzen) von Gegenständen“ (Spiel, 2003, 883). Effektivität und Effizienz sind die Kriterien, die ergebnisorientierten Evaluationsmodellen in der Personalentwicklung zugrunde liegen (Döring & Bortz, 2016).

Das Kirkpatrick-Modell (siehe Abbildung 1) ist das im Bereich der Evaluation von Personalentwicklungsmaßnahmen am häufigsten angewandte Modell und weist den höchsten Bekanntheitsgrad auf (Alliger et al., 1997; Goldstein & Ford, 2002; Kraiger & Jung, 1997; Salas & Cannon-Bowers, 2001).

Abbildung 1: Kirkpatrick-Modell plus eigene Erweiterungen. Eigene Darstellung nach D. L. Kirkpatrick, 1967; 1994.



Das Vier-Ebenen-Evaluationsmodell (D. L. Kirkpatrick, 1959_1; D. L. Kirkpatrick, 1959_2; D. L. Kirkpatrick, 1960_1; D. L. Kirkpatrick, 1960_2; D. L. Kirkpatrick, 1967; D. L. Kirkpatrick, 1998) bietet allerdings lediglich einige allgemeine Bewertungsrichtlinien statt spezifischer Techniken zur quantitativen Bewertung von Personalentwicklungsmaßnahmen, selbst auf der vierten Ebene. Phillips (2003) hielt es für unabdingbar, quantitativ zu bewerten, ob sich die Maßnahme gelohnt hat, und erweiterte daher das Kirkpatrick-Modell um die fünfte Ebene „Return on Investment (ROI)“. Varianten zum ROI, wie z. B. der Return on Development Investment (RODI) von Avolio et al. (2010) und Cascio et al. (2019), die Effectsize von Meuse et al. (2009) und den weiter gefassten Value of Investment (VOI) von Kellner (2006), ermöglichen umfanglichere betriebswirtschaftliche Analysen.

In den wissenschaftlichen Studien der Evaluationsforschung werden mehrheitlich Reaktions-, Lern- und Verhaltensebene erforscht, die Ergebnisebene ist wie bei der Erfolgsmessung in Organisationen unterrepräsentiert (P. B. Warr et al., 1970; Brinkerhoff, 1987; Kaufman & Keller, 1994). Selbst monetäre Indikatoren basieren häufig auf Selbsteinschätzungen der Stakeholder, d. h. auch die Messung des ROI fußt auf der subjektiven Wahrnehmung eines finanziellen Nutzens (Phillips & Phillips, 1994-2001). Für diese Indikatoren liegen wenige empirischen Daten vor, zudem überwiegend aus dem amerikanischen Raum (McGovern et al., 2001).

Die Wirksamkeit von Personalentwicklungsmaßnahmen wird an allgemeinen organisationalen Leistungskriterien wie Produktivitätskennzahlen oder Bilanzkennzahlen festgemacht (Höft, 2001). Die Evaluation von Personalentwicklungsmaßnahmen wird darüber hinaus bezüglich (Ressourcen-)Aufwand und Ertrag kritisch hinterfragt (Süßmair, 2007).

Die bisher in der Praxis im deutschen Fußball angewandte Evaluationsmethode basiert auf einer input-orientierten Methode, entwickelt im Rahmen von Arbeiten zum Qualitätsmanagement für Fußball-Akademien (Hoecke van et al., 2006; zu input-/outputorientierten Methoden siehe F. G. Becker, 2005 und Mittag & Hager, 2000). Eine *output*-orientierte Evaluation, die das Augenmerk auf Endresultate legt, könnte Entscheidungen im Bereich Personalentwicklungsmaßnahmen effizienter unterstützen, als es die bisherigen Ansätze leisten können (Phillips & Schirmer, 2008).

Studien zu Monitoring, Controlling und Finanzierung von Fußballvereinen zeigen, dass eine Betrachtung der geschäftlichen Aktivitäten mittels betriebswirtschaftlicher Kennzahlen im Fußball von vermehrtem Interesse ist (Bosscher et al., 2006; 2009; Güllich et al., 2005). Wallebohr und Daumann (2019) beurteilten in ihrer Studie den Erfolg in der Nachwuchsausbildung Schweizerischer Fußballvereine anhand der Erlöse, die mit dem Verkauf von Nachwuchsspielern erzielt wurden, der produzierten Marktwerte, die im Rahmen der vorliegenden Untersuchung erhoben wurden, sowie der Anzahl der in den Vereinen ausgebildeten BIG-5-Spieler (Big-5-Spieler sind Spieler, die in einer der Ligen des professionellen Fußballs der Länder Spanien, England, Deutschland, Italien und Frankreich spielen). Die gewählten Indikatoren konnten aufgrund der dort aufgezeigten Limitationen lediglich eine Tendenz zwischen den Vereinen aufzeigen. Als Limitationen

wurden u.a. aufgeführt, dass die Herkunft des Spielers keine eindeutige Aussage darüber ermöglicht, in welchem Land der Spieler ausgebildet wurde, die Einsatzzeiten der Spieler im Verein einer Big-5-Liga nicht berücksichtigt worden sind, ein Time-lag-Problem zur Bewertung der Spieler anhand der Marktwerte besteht, und die Daten auf Basis der bestehenden Literatur, der von den Vereinen zur Verfügung gestellten Dokumente und den geführten Interviews subjektiv sind, die Einflussfaktoren sich auf die Nachwuchsausbildung der Schweiz konzentrieren sowie nur eine Zeitpunktbetrachtung durchgeführt wurde. Es wurde im Fazit der Studie empfohlen, die Studien auf andere Länder auszuweiten, die Anzahl der hervorgebrachten Spieler über einen längeren Zeitraum rückwirkend einzubeziehen sowie die Anzahl der Spieler, die den Sprung von den Clubs zum Profispieler geschafft haben, tiefergehend zu analysieren.

1.3 Ziel der Arbeit

Ziel der Nachwuchsleistungszentren ist es, die besten Talente durch eine systematische Förderung auf den Profifußball vorzubereiten (DFB, 2020; Lanwehr et al., 2021). Zur Überprüfung des Erfolgs der Bemühungen um den fußballerischen Nachwuchs und des Qualitätsmanagements in den Nachwuchsleistungszentren ist eine Evaluation der Aktivitäten und vorhandenen Kompetenzen erforderlich. Die finanzielle Situation vieler Vereine der deutschen Bundesliga ist insbesondere seit den Einschränkungen des Spielbetriebs während der COVID-19-Pandemie durchaus bedrohlich (Zülch, 2020), sodass Strategien auch bei der Investition in die Nachwuchsausbildung hinterfragt werden müssen. Evaluationen der Arbeit in den Nachwuchsleistungszentren mittels betriebswirtschaftlicher Kennzahlen sollen Transparenz schaffen und als fundierte faktenbasierte Entscheidungsgrundlage bei zukünftigen Ressourcenverteilungen und der Entwicklung der Unternehmensstrategie dienen (Phillips & Schirmer, 2008). Aufgrund der bedrohlichen finanziellen Lage werden aufwändige ressourcenverzehrende Evaluationsaktivitäten von den Vereinen aber nicht durchgeführt, aber belastbare valide Bewertungen der Nachwuchsarbeit erwartet. Dieses Dilemma führt zur Forschungsfrage:

Wie kann der Erfolg von Nachwuchsleistungszentren im deutschen Profifußball mittels Investitionskennzahl und angemessenem Erhebungsaufwand valide evaluiert werden?

Evaluationen der Effektivität von Nachwuchsleistungszentren im deutschen Profifußball sind bisher input-orientiert, eher auf das Vorhandensein spezifischer Qualitätsdimensionen, wie z.B. Infrastruktur & Ausstattung, Strategie & Finanzen und Organisation & Verfahren, ausgelegt (Kroemer, 2015; Lanwehr et al., 2021). Es gibt für die Nachwuchsleistungszentren keine ökonomisch einsetzbaren Verfahren, die ein angemessenes Verhältnis von Aufwand und Nutzen der Evaluation aufweisen. Darüber hinaus sind Evaluationen, die überwiegend auf Selbsteinschätzungen und subjektiven Einschätzungen von Auditoren beruhen, wie sie sehr häufig in der Praxis der Messung des Erfolgs

von Personalentwicklungsmaßnahmen durchgeführt werden, von fehlender Objektivität geprägt. Die bisherigen Erfolgsmessungen in der Personalentwicklung werden eher kurzfristig direkt nach der Personalentwicklungsmaßnahme durchgeführt (die ersten drei Ebenen des Kirkpatrick-Modells), obwohl der Einfluss einer Personalentwicklungsmaßnahme mehrere Monate oder Jahre braucht, um sich zu entfalten und eine Betrachtung im Zeitverlauf. In der betriebswirtschaftlichen Erfolgsfaktorenforschung, in der insbesondere geeignete Indikatoren auf organisatorischer Ebene gesamthaft gemessen werden, fehlen geeignete Leistungsindikatoren, die die Wirksamkeit einer Personalentwicklungsmaßnahme individuell nachweisen kann.

Unter Berücksichtigung der aus den Forschungsarbeiten, der drei betrachteten Bereiche – Evaluationsmodelle in der Personalentwicklung, Investitionskennzahlen in der Personalentwicklung und die Messung des Erfolgs von Nachwuchsarbeit im deutschen Profifußball – hervorgegangenen Empfehlungen und beschriebenen Limitationen, ist das umfassende Ziel der Arbeit daher die Entwicklung eines Evaluationskonzepts, das

- outcome-orientiert statt input-orientiert ist,
- die Resultate im Zeitverlauf betrachtet statt lediglich zu einem Zeitpunkt direkt nach der Maßnahme,
- öffentlich zugängliche objektive Daten nutzt statt Selbsteinschätzungen der Teilnehmer,
- Individual-Leistungs-Indikatoren der Teilnehmer berücksichtigt statt allgemeine gesamtorganisatorische Erfolgskennzahlen,
- mit geringem Aufwand leicht anzuwenden ist und
- mittels geeigneter betriebswirtschaftlicher Kennzahlen statt nicht-monetärer Indikatoren evaluiert.

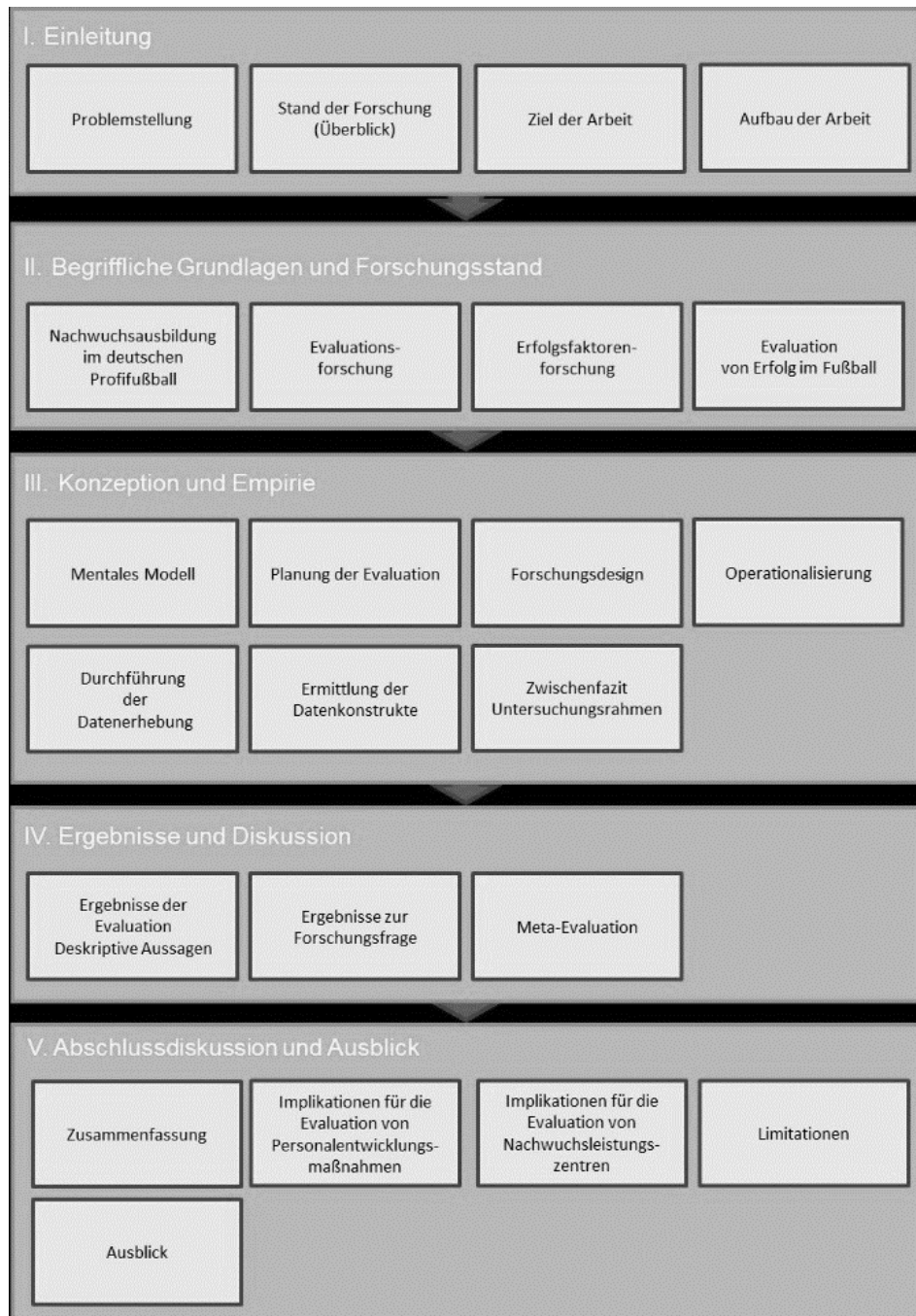
Die Evaluation folgt dabei den Standards für Evaluation, formuliert durch die Deutsche Gesellschaft für Evaluation (DeGEval, 2017c).

Mit der Analyse liefert die Studie zudem eine empirische Grundlage, die zur Theoriendebatte der Evaluation des monetären Erfolgs von Personalentwicklungsmaßnahmen beitragen kann.

1.4 Aufbau der Arbeit

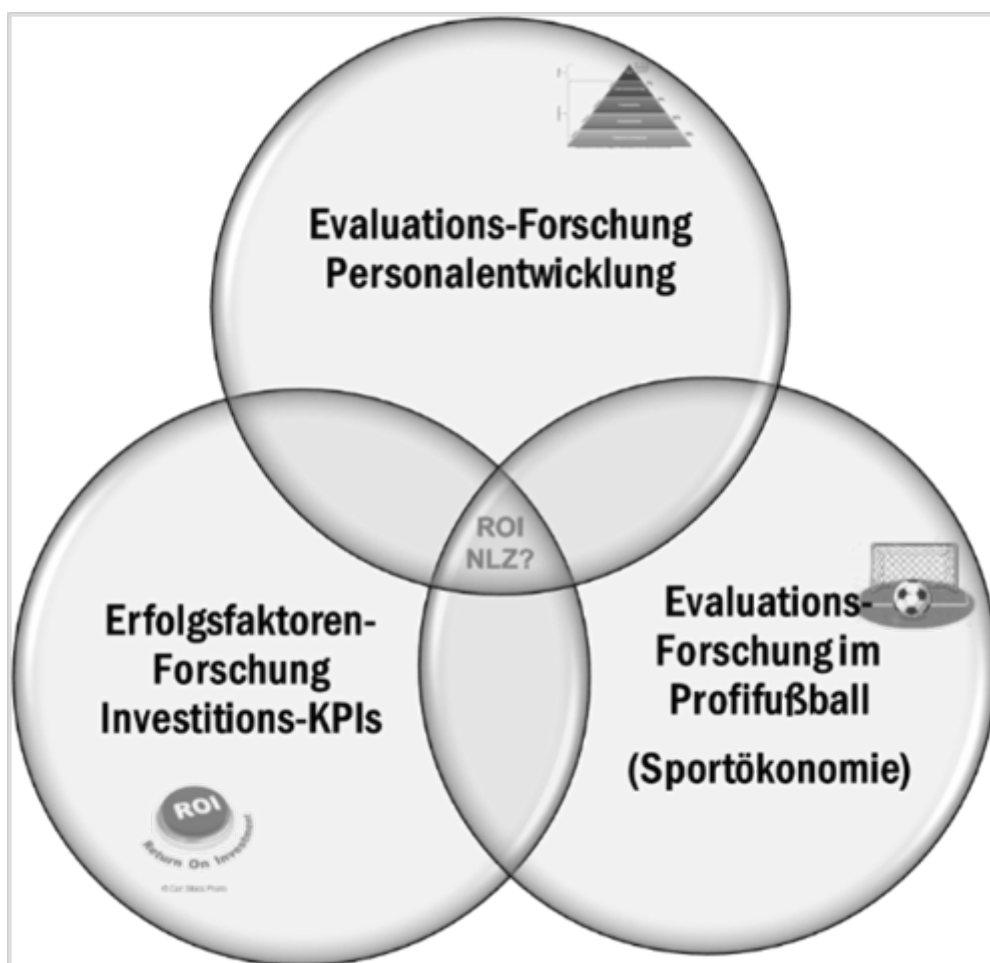
Die Arbeit ist in fünf Teile und vierundzwanzig Kapitel gegliedert (vgl. Abbildung 2). Hinzu kommen zu Beginn bzw. Ende der Arbeit das Vorwort, die Verzeichnisse (Inhalt, Abbildungen, Tabellen und Abkürzungen) sowie die eidesstattliche Erklärung.

Abbildung 2: Aufbau der Arbeit



Im ersten Teil der Arbeit wird die Problemstellung erläutert, ein Überblick zum Stand der Forschung gegeben und das Ziel der Arbeit vorgestellt. Dieser Teil beinhaltet neben der Abgrenzung des Themas auch die Forschungsfrage, mit der sich die Arbeit auseinandersetzt. Darauf aufbauend zeigt der zweite Teil der Arbeit die wichtigsten Grundlagen im Zusammenhang mit dem Untersuchungsgegenstand *Nachwuchszentrum* im deutschen Profifußball auf und stellt die erklärungsrelevanten Rahmenbedingungen und Theorien sowie den aktuellen Stand der Forschung aus der Evaluations-, Erfolgsfaktorenforschung und der Evaluation des Erfolgs im Fußball vor (siehe dazu Abbildung 3).

Abbildung 3: Betrachtete Theorien und Forschungsbereiche



Der dritte Teil der Dissertationsschrift umfasst den konzeptionellen und empirischen Teil, das konkrete Vorgehen, die Beschreibung und Entwicklung der Datenkonstrukte und deren Validierung. Im vierten Teil werden die Ergebnisse und die gewonnenen Resultate vorgestellt und diskutiert. Zuletzt werden im fünften Teil der Arbeit die Implikationen für die Evaluation von Entwicklungsmaßnahmen diskutiert und Vorschläge für die an die Arbeit anknüpfende Forschung aufgeführt.

2 Begriffliche Grundlagen und Forschungsstand

Der Forschungsgegenstand *Nachwuchsleistungszentrum* ist multidimensional zu betrachten. Zum einen liegt das an der Vielzahl der Beteiligten und damit der diversen Perspektiven auf die Nachwuchsleistungszentren, zum anderen kann der Erfolg eines Nachwuchsleistungszentrums auf verschiedenen Ebenen betrachtet werden. Um den Untersuchungsgegenstand darzustellen, erfolgt in Kapitel 2.1 zunächst eine Präzisierung der *Nachwuchsausbildung im deutschen Profifußball* sowie eine Abgrenzung zum internationalen Markt. Kapitel 2.2 enthält eine begriffliche Klärung von *Evaluation* und eine Vorstellung des Forschungsstandes hinsichtlich existierender Evaluationsansätze mit dem Schwerpunkt auf Personalentwicklungsmaßnahmen. Im Anschluss daran wird in Kapitel 2.3. die Erfolgsfaktorenforschung im Hinblick auf Investitions-Kennzahlen beschrieben und in Kapitel 2.4 auf die Evaluation des Erfolgs im Fußball eingegangen. Auf Basis der Betrachtung der drei Forschungsfelder erfolgt in Kapitel 2.5 die Ableitung der Forschungsfrage.

2.1 Nachwuchsausbildung im deutschen Profifußball

Das Untersuchungsobjekt dieser Arbeit sind die Nachwuchsleistungszentren im deutschen Profifußball. Ihr Ziel ist es „im Rahmen der übergeordneten Ziele ‚Weltspitze in Bundesliga und Nationalmannschaften‘ sowie ‚Optimale Ausschöpfung des Talentpools‘, die Qualität der Talentförderung im Lizenzbereich und oberen Amateurbereich zu optimieren“ (DFB, 2013, S. 1). Die Vielzahl der an der Ausbildung Beteiligten – Vereine, (inter-)nationale Fußballverbände, Spieler, Berater, Trainer, Eltern, Psychologen, Physiotherapeuten, Pädagogen, Kaufleute, Mediziner, Vereinsmanagement – ermöglichen eine Vielfalt an Perspektiven, mit denen der Untersuchungsgegenstand betrachtet werden kann. Um das Untersuchungsobjekt darzustellen, werden zunächst die Organisation und die Aktivitäten von Nachwuchsleistungszentren beschrieben und eine Abgrenzung der für diese Arbeit gewählten Perspektive vorgenommen. Im Anschluss daran werden die Rahmenbedingungen der Nachwuchsausbildung im Fußball erläutert. Zum Abschluss des Kapitels erfolgt eine Zusammenfassung zur Nachwuchsausbildung im deutschen Profifußball.

2.1.1 Nachwuchsleistungszentren

Die Aufgabe eines Nachwuchsleistungszentrums besteht darin, junge talentierte Spieler zu erkennen und zu fördern, um sie an den Lizenzfußball heranzuführen. Kern des Unternehmens ist es, durch eine systematische, intensive und innovative Förderung das Potential der Talente in einem hochprofessionellen Umfeld auf spätere Anforderungen im Lizenzfußballbereich vorzubereiten (DFB, 2014a). Per Lizenzierungsstatut haben die am „Spielbetrieb der 1. und 2. Bundesliga teilnehmenden Vereine und Kapitalgesellschaften als Fördereinrichtung ein Leistungszentrum zu führen“ (DFL, 2021c). Der

Aufbau und die Unterhaltung der Nachwuchsleistungszentren wird durch DFL und DFB mit der Einführung eines Qualitätsmanagementsystems unterstützt. Die Durchführung systematischer Evaluationen der erforderlichen Kompetenzen und Ressourcen auf zwei Ebenen soll zur Sicherung und Steigerung einer nachhaltigen Ausbildungsqualität in den Leistungszentren führen. Im Rahmen einer Lizenzierung wird das Vorhandensein der zugehörigen Funktionen, wie z.B. Physiotherapie, schulische und psychologische Unterstützung für die Talente, Ausbildung der Trainer, geprüft.

Welche Kompetenzen in welcher Qualität im Nachwuchsleistungszentrum aufgebaut wurden, wird über eine Zertifizierung quantifiziert (DFB, 2014b). Die ursprünglichen Vorgaben sind im Laufe der vergangenen Jahre in nahezu allen Bereichen verfeinert und damit stetig weiterentwickelt worden (DFB, 2021c). Seit dem 01.01.2021 gehen die grundsätzlichen Anforderungen, die die Vereine für den Betrieb von Leistungszentren erfüllen müssen, aus der neuen Lizenzierungsordnung und den neuen Richtlinien für den Aufbau und Unterhaltung von Nachwuchsleistungszentren der Teilnehmer der Lizenzligen hervor (Lanwehr et al., 2021; DFL, 2020c). Die Entwicklung der Leistungszentren seit 2001 zeigt Abbildung 4.

Die DFL hat in den Richtlinien die Anforderungen der Nachwuchsleistungszentren für die Lizenzvereine wie folgt definiert (die folgenden Ausführungen basieren auf (DFL, 2021c): Die Nachwuchsleistungszentren werden in verschiedene Güteklassen eingeteilt. Ein Verein der 1. Bundesliga muss dabei mindestens die Kategorie II erreichen, während die Vereine der 2. Bundesliga die Kategorie III erreichen müssen. Es sind zahlreiche Voraussetzungen für die verschiedenen Kategorien zu erfüllen. Die Leistungszentren sind in Grundlagen-, Aufbau- und Leistungsbereich aufgeteilt. Im Grundlagenbereich (U8-U11) muss mindestens eine Mannschaft gestellt werden, der Aufbaubereich (U12-U15) besteht aus einer Mannschaft pro Jahrgang und der Leistungsbereich (U16-U23) umfasst drei bis fünf Mannschaften, wobei die U23 optional ist. Somit ergeben sich mindestens acht bis maximal zehn

Abbildung 4: Entwicklung rund um die Nachwuchsleistungszentren 2001 bis heute. Eigene Darstellung nach (DFL, 2021b).



Mannschaften für die Vereine. Auch die Anzahl der Spieler pro Mannschaft wird von der DFL vorgegeben. Bis zur U15 zählt ein Kader 20 Spieler, anschließend je nach Aufbau des Nachwuchsleistungszentrums bis zu 24 Spieler. Zur Sicherung des Nachwuchses für die Nationalmannschaft sind auf den Spielerlisten des Leistungsbereichs immer zwölf Spieler deutscher Nationalität zu melden. Diese Regelungen sollen sicherstellen, dass der Anteil der deutschen Spieler im Lizenzfußball steigt und das Durchschnittsalter von Bundesligaprofis sinkt. 2019 wurden eine Restrukturierung des Systems *Nachwuchsleistungszentrum* zum Abschluss gebracht und ein neuer Ansatz zur Qualitätssicherung implementiert (DFL, 2021c). Die bisherige Zertifizierung der Nachwuchsleistungszentren wurde abgeschafft, stattdessen ist ein ganzheitlicher Ansatz zur stetigen Förderung der besonderen Talente, der sog. Qualitätsmanagement-Kreis, implementiert worden (DFL, 2021c).

Das Qualitätsmanagement umfasst acht Teilbereiche: Lizenzierung, Regelmäßige Analysen, Unterstützung/Soziales, Wissensmanagement, Entwicklung und Innovation, Anreizsysteme, Kommunikation und Steuerungsgruppen (DFB, 2014b; DFL, 2021c). Die Lizenzierung sichert die Rahmenbedingungen für das Arbeiten auf Leistungssportniveau in den Leistungszentren und bildet die Grundlage für Anreizsysteme. Sie gliedert sich in die Bereiche Management, Fußballausbildung und Unterstützung. Die Vereine arbeiten im Bereich der individuellen Analyse eng mit dem DFB und der DFL zusammen und können so konkrete Bereiche analysieren und Ansätze zu Weiterentwicklung der Nachwuchsleistungszentren entwerfen (DFL, 2021c). Neben dem Vereinsleben werden auch die Möglichkeiten zur Erbringung guter schulischer Leistungen der jungen Talente durch das Qualitätsmanagement des DFB, basierend auf den Eliteschulen des Fußballs oder einer direkten Kooperation zwischen Schulen und Clubs, bewertet. Eine Sensibilisierung der Spieler und Mitarbeiter der Leistungszentren zu Themen wie Rassismus oder Doping und auch die Unterbringung der Spieler sind ebenfalls Teil des Qualitätsmanagements. Fortbildungen und Serviceportale für die Mitarbeiter der Leistungszentren bilden den abschließenden Bereich des Qualitätsmanagements. Dazu finden regelmäßig Fortbildungsveranstaltungen der DFL für verschiedene Bereiche der Nachwuchsleistungszentren statt und Best-Practice Beispiele sollen den Vereinen helfen, sich weiterzuentwickeln.

2.1.2 Abgrenzung Untersuchungsgegenstand

Gegenstand der vorliegenden Untersuchung sind die Nachwuchsleistungszentren im deutschen Profifußball. Evaluiert wird die Qualität ihrer Nachwuchsausbildung gemessen anhand eines Indikators, der den Erfolg dieser Bemühungen misst. Dies geschieht auf zwei Ebenen: *individuell* auf der Ebene des einzelnen Fußballspielers und *organisatorisch* auf der Ebene eines Vereins mit Nachwuchsleistungszentrum. Das System des Profifußballs und der Nachwuchsförderung in Deutschland ist im Kontext mit internationalen und nationalen Fußballverbänden (z. B. Fédération Internationale de Football Association (FIFA), Union of European Football Associations (UEFA), DFB, DFL)

und deren Rahmenbedingungen zu betrachten, da der DFB, als Mitglied des Weltverbands (FIFA) und des europäischen Fußballverbands UEFA, den Regelungen dieser Verbände folgt.

Individuelle Ebene

Alle Faktoren, die auf der Ebene des einzelnen Fußballspielers zur Bewertung seiner Qualität genutzt werden, bilden die *individuelle Ebene*. Unter der Annahme, dass die Beschaffenheit der Ausbildung in den Nachwuchsleistungszentren in erheblichem Maße zur Bildung von Fähigkeiten und Erfahrungen und damit zur Qualität eines Spielers beiträgt, ist ein wesentlicher Faktor die *Verweildauer des Spielers im Nachwuchsleistungszentrum*. Eine Betrachtung anderer Einflussfaktoren, wie Talent, spezielle physische oder psychische Fähigkeiten, ausbildende Trainer, Trainingsmethoden o. ä., erfolgt in dieser Studie nicht. Einen Überblick über den aktuellen Stand der Erfolgsfaktorenforschung auf der individuellen Ebene zeigen Bosscher et al. (2006), Brand, S., Emrich, E., Güllich, A., Prohl, R., Turbanski, S. (2006) Durand-Bush und Salmela (2001) und der Deutscher Sportbund (2006). Darüber hinaus stellen (Emrich et al., 2005) eine Liste mit Studien deutscher Talentforscher zur Verfügung.

Organisatorische Ebene

In Analogie zu dem bis Ende 2019 zur Bewertung der Qualität angewandten Systems – basierend auf den Arbeiten von Knop et al. (2004), Bosscher et al. (2006) und Hoecke van et al. (2006) – wird auf der gesamtorganisatorischen Ebene die Nachwuchsausbildung im Verein betrachtet. Dabei wird aber in dieser Arbeit im Gegensatz zum *PASS 4.0*-Ansatz, mit dem die Ausgestaltung von fünf Qualitätsdimensionen und zugehörigen Sub-Qualitätsdimensionen (siehe dazu Abbildung 5) eines Nachwuchsleistungszentrums bewertet wird, das Nachwuchsleistungszentrum als *Black-Box* betrachtet. Der Erfolg des Nachwuchsleistungszentrums wird abhängig gemacht von der Summe des Erfolgs der einzelnen Fußballspieler, indem die Summe aller hervorgebrachten Spielerqualitäten, gewichtet mit der Verweildauer im Nachwuchsleistungszentrum für die Organisation gebildet wird.

Abbildung 5: PASS 4.0 Qualitätsdimensionen. Eigene Darstellung auf Basis der zur Verfügung gestellten Double-Pass- Daten.



2.1.3 Rahmenbedingungen für die Nachwuchsarbeit

In diesem Kapitel werden die Rahmenbedingungen, die bei der Betrachtung der Nachwuchsausbildung im deutschen Profifußball zu berücksichtigen sind, grundlegend erläutert. Dazu gehören die internationalen Rahmenbedingungen, insbesondere die Regulierungen der FIFA und der UEFA, die kontextrelevanten Informationen zur Nachwuchsausbildung, ein Überblick über die an der Nachwuchsarbeit Beteiligten sowie für die Thematik relevante Kennzahlen zur Talentförderung im Profifußball in Deutschland.

2.1.3.1 Internationale Rahmenbedingungen für die Nachwuchsarbeit

Késenne (2011) empfiehlt auf internationaler Ebene sowie auf der Ebene der nationalen Ligen und Verbände, Regulierungen einzuführen, die Vereinen (finanzielle) Anreize bieten, vermehrt in den eigenen Nachwuchs zu investieren. FIFA und UEFA sowie auch DFB und DFL in Deutschland sind dieser Forderung bereits nachgekommen. Die internationalen Verbände haben Regulierungen eingeführt, die es den Vereinen ermöglichen, die Nachwuchsarbeit auch unter ökonomischer Perspektive – Kosten-Nutzen-Betrachtung – zu verstehen. Die *Ausbildungsentschädigung*, der *Solidaritätsbeitrag* und die *Home Grown Player Rule* sind hier als die wichtigsten zu nennen (Wallebohr & Daumann, 2019). Innerhalb der Regelungen zum *Finanziellen Fairplay* (Financial Fair Play) der UEFA werden Investitionen in die Nachwuchsarbeit gesondert betrachtet (UEFA, 2014).

2.1.3.1.1 Ausbildungsentschädigung und Solidaritätsbeitrag

Die FIFA hat seit 2005 mit der Einführung der *Ausbildungsentschädigung* und des *Solidaritätsbeitrags* finanzielle Anreize für Vereine geschaffen, in die Ausbildung von Spielern zu investieren (FIFA, 2019b). Investitionen in die Nachwuchsarbeit werden honoriert, indem den ausbildenden Vereinen eine Entschädigung für die Kosten der Ausbildung des Spielers gezahlt wird. Sie stellen eine Einnahmequelle für die Vereine dar (zum Einnahme-Mix von Vereinen siehe Kapitel 2.4.3). Eine Ausbildungsentschädigung ist als „eine festgeschriebene Ablösesumme für einen Spieler unter 23 Jahren, der seinen ersten Profivertrag nicht in dem Verein unterzeichnet, der ihn ausgebildet hat, bezeichnet“ (FIFA, 2019b). Die Ausbildungsentschädigung geht auch bei jedem nachfolgenden Transfer bis zum Ende der Spielzeit, in der der Spieler 23 Jahre alt wird, an den Verein, in dem der Spieler ausgebildet wurde. Die Berechnung der Höhe der Ausbildungsentschädigung für ehemalige Vereine basiert auf dem finanziellen Aufwand, den der aufnehmende Verein gehabt hätte, wenn der Spieler dort ausgebildet worden wäre. Er richtet sich nach der Spielzeit, in der der Spieler 12 Jahre alt geworden ist, und endet mit der Spielzeit, in der der Spieler 21 Jahre alt wird (zu den Berechnungsdetails und Regulierungen siehe: FIFA, 2019b, S. 69-70).

Der *Solidaritätsbeitrag* wurde mit dem gleichen Ziel wie die Ausbildungsentschädigung implementiert, wird aber auch für Transfers fällig, die nach der Ausbildung des Spielers

erfolgen (FIFA, 2019b). „Wird ein Spieler vor Ablauf seines Vertrags transferiert, erhalten alle Vereine, die zu seinem Training und seiner Ausbildung beigetragen haben, einen Teil der Entschädigung, die an seinen ehemaligen Verein entrichtet wird (Solidaritätsbeitrag)“ (FIFA, 2019b, S. 26). Die Ermittlung der Höhe des Solidaritätsbetrags erfolgt auf Basis des Verhältnisses zur Spielzeit, die der Spieler zwischen seinem 12. und 23. Geburtstag bei den jeweiligen Vereinen verbracht hat (zu den Berechnungsdetails und Regulierungen siehe: FIFA, 2019b, S. 72-73).

Die Ausbildungsentschädigung ist rechtlich umstritten. Das Landgericht Oldenburg urteilte in 2005, dass diese gegen den Artikel 12 des Grundgesetzes – Gewährleistung der Berufsfreiheit – verstößt. Mit der Ausbildungsentschädigung könnten sich potentielle Arbeitgeber

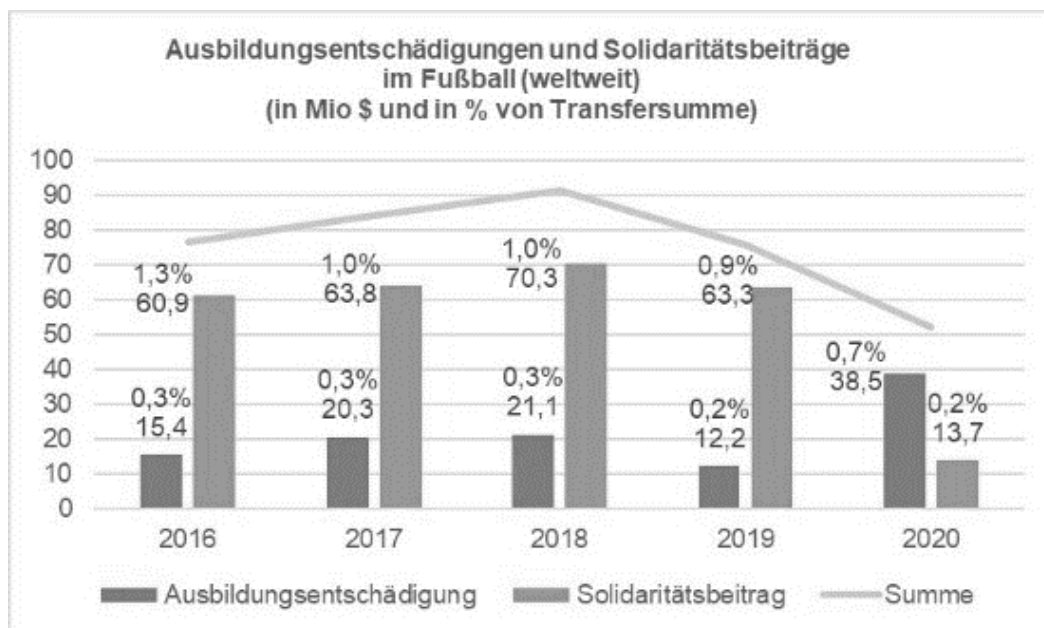
des Fußballspielers abschrecken lassen (Müller-Römer, 2006). Der Europäische Gerichtshof dagegen erklärte die Ausbildungsentschädigung für rechtmäßig. „Vor diesem Hintergrund könnten ausbildende Vereine davon abgehalten werden, in die Ausbildung junger Spieler zu investieren, wenn sie keinen Ersatz der dafür aufgewendeten Beträge erhalten könnten, falls ein Spieler nach Abschluss seiner Ausbildung einen Vertrag als Berufsspieler mit einem anderen Verein abschließt“ (EUGH, 2010, S. 44).

Zwischen 2016 und 2020 sind Entschädigungen für die Ausbildung von Spielern von 15,4 Mio. US-Dollar in 2016 über das zeitweilige Hoch in 2018 von 21,1 Mio. US-Dollar auf 13,7 Mio. US-Dollar gesunken. Die Solidaritätsbeiträge machen ein Vielfaches der Ausbildungsentschädigungen aus, sie sind im gleichen Zeitraum von 60,9 Mio. US-Dollar, über 70,3 Mio. US-Dollar in 2018 auf 38,5 Mio. US-Dollar in 2020 gefallen. Die Höhe der Beträge stellen einen prozentualen Anteil der Transfersumme dar, sodass damit die Differenzen erklärt werden können. Bis auf das Jahr 2020, dessen stark gesunkene Zahlen mit den Folgen der COVID-19-Pandemie erklärt werden können, liegt der prozentuale Anteil an der Transfersumme bei Ausbildungsentschädigungen um die 0,3 %, bei Solidaritätsbeiträgen um die 1 % (FIFA, 2021). Einen Einblick in die Höhe der Ausbildungsentschädigungen und Solidaritätsbeiträge weltweit (Bereich der FIFA) gibt die Abbildung 6.

2.1.3.1.2 Homegrown Player Rule

Ziel der Homegrown Player Rule ist es, die „lokale Ausbildung junger Spieler zu fördern und die Offenheit und Fairness der europäischen Wettbewerbe zu erhöhen“ (UEFA, 2005). Seit der Saison 2008/09 müssen Vereine in der UEFA Champions League und UEFA Europa League mindestens acht einheimische Spieler in einem auf 25 Spieler begrenzten Kader haben. Einheimische Spieler sind Spieler, die „unabhängig von ihrer Nationalität mindestens drei Jahre lang zwischen dem Alter von 15 und 21 Jahren von ihrem Verein oder einem anderen Verein desselben Nationalverbands ausgebildet wurden“ (UEFA, 2005). Bis zu 50 % der lokal ausgebildeten Spieler müssen aus dem Verein selbst stammen, die anderen entweder aus dem lokalen Verein oder aus anderen Vereinen desselben nationalen Verbandes kommen (UEFA, 2005).

Abbildung 6: Ausbildungsentschädigungen und Solidaritätsbeiträge weltweit. Eigene Darstellung (FIFA, 2017, 2018, 2019a; FIFA, 2020, 2021).



In Deutschland wurde zur Saison 2006/07 die Regel eingeführt, dass in jedem Verein der 1. und 2. Bundesliga mindestens zwölf deutsche Lizenzspieler unter Vertrag stehen (DFL, 2020b). Verabschiedet wurde darüber hinaus ebenfalls die Einführung der von der UEFA praktizierten Homegrown Player Rule zur Förderung des Nachwuchses. „2006/07 mussten pro Club mindestens vier lokal – von einem deutschen Verein– ausgebildete Spieler unter Vertrag stehen, 2007/08 mindestens sechs und seit 2008/09 mindestens acht (4+4). Ein von Vereinsseite ausgebildeter Spieler muss in drei Spielzeiten/Jahren im Alter zwischen 15 und 21 Jahren für den Verein oder die Kapitalgesellschaft spielberechtigt gewesen sein, ein vom Verband ausgebildeter Spieler in drei Spielzeiten/Jahren im Alter zwischen 15 und 21 Jahren für einen Verein oder eine Kapitalgesellschaft im Bereich des DFB“ (DFL, 2020b). Ähnliche Regeln sind auch in mehreren nationalen Ligen in Europa in Kraft (Smokvina, 2013). Die Homegrown Player Rule der UEFA hat allerdings bis dato lediglich einen geringen Einfluss auf die Ausbildung von jungen Spielern in europäischen Vereinen. Die Anzahl der einheimischen Spieler in den Lizenzmannschaften hat nur leicht zugenommen (Dalziel et al., 2013, Smokvina, 2013). Veränderungen hinsichtlich der Investitionen oder der Qualität der Nachwuchsleistungszentren konnten im Rahmen der Studie nicht identifiziert werden (Wallebohr & Daumann, 2019).

2.1.3.1.3 Finanzielles Fairplay

Ziel des im September 2009 verabschiedeten Konzepts für *Finanzielles Fairplay* (FFP) ist, das Wohlergehen des Fußballs in Zeiten von ernsthaften, sich verschärfenden finanziellen Problemen des europäischen Vereinsfußballs zu fördern. Regelmäßige Aktualisierungen in den Jahren 2012, 2015, 2018 basieren weiterhin auf diesen originären grundlegenden Prinzipien und Zielen (UEFA, 2014). Die wirtschaftliche und finanzielle Leistungsfähigkeit der Vereine soll verbessert sowie ihre Transparenz und Glaubwürdigkeit erhöht werden (Heermann, 2022). Darüber hinaus soll sichergestellt werden, dass die Vereine ihren Verbindlichkeiten – gegenüber ihren Angestellten, den Sozialversicherungen, Steuerbehörden und anderen Gläubigern – nachkommen können. Die Regelung soll ebenfalls Vereine dazu bringen, bei gleichzeitiger langfristiger Förderung des Fußballs wirtschaftlich zu arbeiten (UEFA, 2014). Ausgaben durch Finanzierungen in Stadien, Jugendakademien oder gemeinnützige Projekte zur Förderung junger Fußballspieler sind von der Break-even-Berechnung ausgenommen (siehe Artikel 58-64, UEFA, 2018).

Drei Regulierungen internationaler Verbände (wie FIFA und UEFA) bieten damit den Vereinen direkte oder indirekte finanzielle Anreize, in die Nachwuchsförderung zu investieren: Ausbildungsentschädigung, Solidaritätsbeitrag und Homegrown Player Rule. Ein Überblick über die beschriebenen internationalen Rahmenbedingungen kann Tabelle 1 entnommen werden.

Die Homegrown Player Rule vermeidet einen unmittelbaren Bezug zu der Staatsangehörigkeit (Parrish & Miettinen, 2008). Damit unterscheidet sie sich von der im sog. Bosman Verfahren für rechtswidrig in der EU erklärten 6+5-Regel. Diese über die Homegrown Player Rule hinausgehende vorgesehene Regulierung (6+5-Regel), initiiert durch FIFA und die UEFA, sollte zusätzlich dazu beitragen, die Anzahl der einheimischen Spieler in den Clubs erhöhen. Vorgaben des EUGHs standen der Einführung dieser Regel entgegen (Battis et al., 2010; INEA, 2008).

Mit den Auswirkungen des Bosman Urteils auf den europäischen Profifußball haben sich zahlreiche Studien bereits beschäftigt (Cachay & Riedl, 2002; Schroeder, 2016; Geyer & Dilger, 2009), eine tiefere Untersuchung der Auswirkungen auf die Nachwuchsarbeit erfolgt im Rahmen der vorliegenden Studie nicht.

Die Homegrown Player Rule könnte eine indirekte Diskriminierung darstellen, durch die der Status eines Profifußballspielers in der Mannschaft beeinträchtigt werden könnte.

Tabelle 1: Überblick über die beschriebenen internationalen Rahmenbedingungen. Eigene Darstellung nach Wallebohr (2019).

Rahmenbedingung	Verband	Regelung
Ausbildungsentschädigung	FIFA	Bei Unterzeichnung des ersten Vertrags als Profi durch den Spieler und bei jedem Transfer bis zum Ende der Spielzeit, in der der Spieler 23 Jahre alt wird, erhalten Vereine, die einen Spieler ausgebildet haben, eine Ausbildungsentschädigung.
Solidaritätsbeitrag	FIFA	Bei einem Spielertransfer vor Ablauf des Vertrags erhalten alle Vereine, die zum Training und der Ausbildung des Spielers beigetragen haben, einen Teil der Entschädigung.
Homegrown Player Rule	UEFA	In einer Mannschaft mit 25 Spielern müssen mind. acht einheimische Spieler vertreten sein. D. h., es darf ein maximaler Anteil ausländischer Spieler von 68 % in der Mannschaft bestehen; mindestens 32 % müssen einheimische Spieler sein.
Finanzielles Fairplay	UEFA	Für einen für den UEFA-Wettbewerb qualifizierten Verein ist nachzuweisen, dass im Verlauf einer Saison keine überfälligen Verbindlichkeiten gegenüber Dritten bestehen und kostendeckend gewirtschaftet wird.

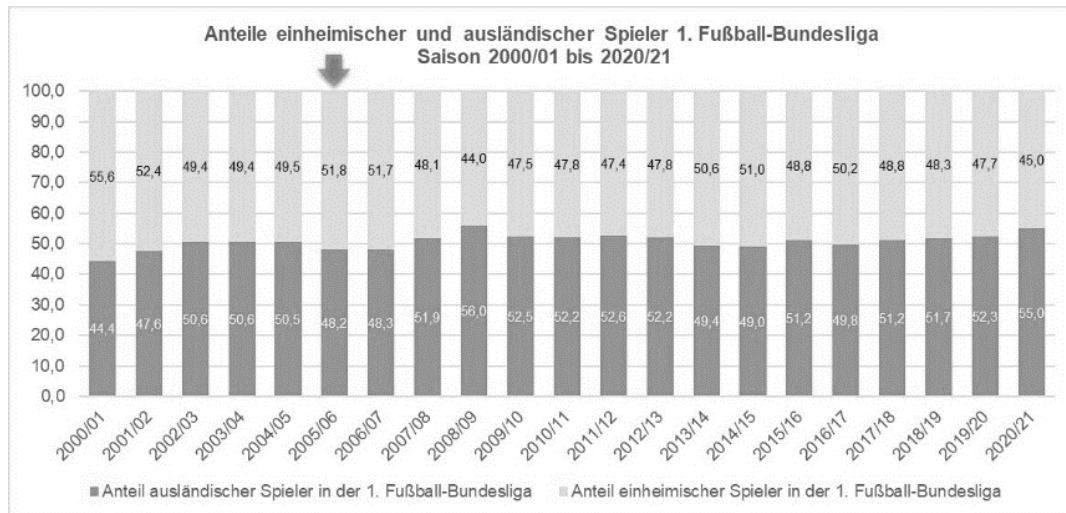
Eine Beeinträchtigung liegt dann vor, wenn statt der Staatsangehörigkeit andere Unterscheidungsmerkmale herangezogen werden, die mit einer ausländischen Herkunft verknüpft sein können (Kliesch, 2017). Mit großer Wahrscheinlichkeit werden lokal ausgebildete Spieler *Einheimische* sein, da sie sich eher einem Verein aus dem eigenen Land anschließen (Dalziel et al., 2013, S. 89).

Die Auswirkungen der Regulierungen der FIFA und UEFA auf den Profifußball in Deutschland und anderen Ländern können anhand der fußballspezifischen Kennzahl *Anteil Einsatz ausländischer bzw. selbst ausgebildeter Spieler* – wenn auch nur näherungsweise – für die ersten Ligen gemessen werden. Kontinuierliche Untersuchungen der Forschungsgruppe „Football Observatory“ des internationalen Zentrums für Sportstudien (CIES) zu den Entwicklungen im Profifußball in den größten Ligen Europas berichten u. a. diese Kennzahl (vgl. dazu diverse Reports - CIES Football Observatory (football-observatory.com)).

Über den Betrachtungszeitraum der Saisons 2000/01 bis 2020/21 sind im Mittel 50,8 % der Spieler in der 1. Fußball-Bundesliga ausländischer Herkunft, seit der Einführung der

Regulierungen der FIFA und UEFA zur Saison 2005 (Ausbildungsentschädigung, Solidaritätsbeitrag) sogar im Mittel 51,2 %. Der Anteil der Legionäre (Nicht-Einheimischen) ist seit der Saison 2005/06 permanent über die damalige 48,2 % Marke gestiegen (Statista, 2021a). Den Verlauf des prozentualen Anteils ausländischer und einheimischer Spieler in der 1. Bundesliga in den Saisons von 2000/01 bis 2020/21 zeigt Abbildung 7.

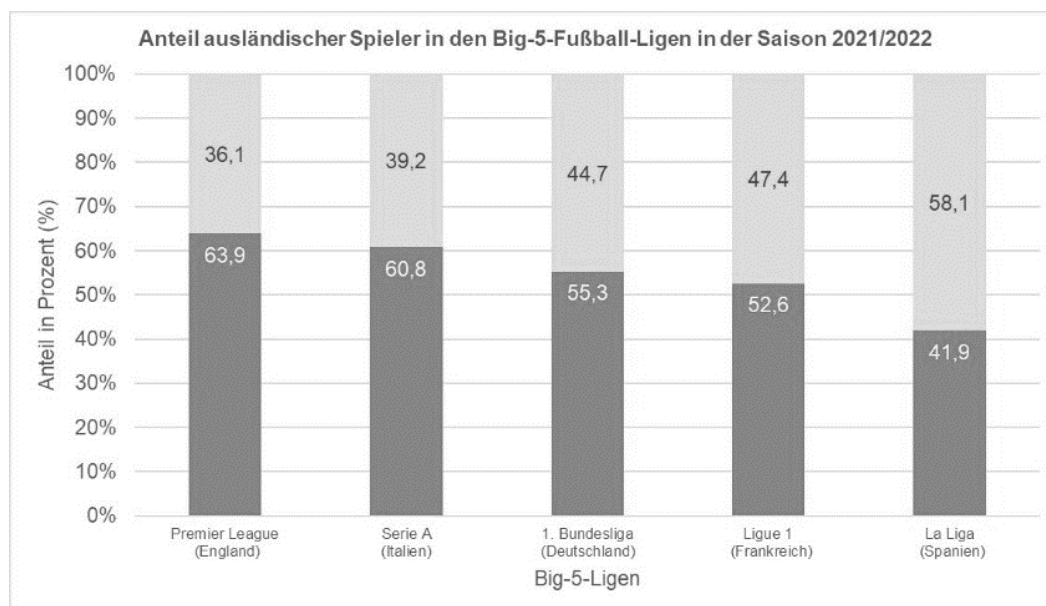
Abbildung 7: Anteil einheimischer und ausländischer Spieler in der 1. Bundesliga bis Saison 2019/20. Eigene Darstellung (Statista, 2021a).



Im September 2021 (Saison 21/22) sind 55 % der gesamten Spieler der 1. Fußball-Bundesliga ausländischer Herkunft, somit 45 % im eigenen Verein oder nationalen Verband ausgebildete Spieler Teil der Mannschaft. Der Mindestanteil an *Einheimischen* – gemäß Homegrown Player Rule acht Spieler einer Mannschaft von 25 Spielern, also 32,0 % – wird damit um 12,3 Prozentpunkte überschritten. Der höhere Anteil ausländischer Spieler hat einen rückläufigen Anteil der im Verein oder im deutschen Verband ausgebildeten Spieler (Einheimische) zur Folge. Die Bundesliga liegt im Vergleich mit anderen Top-Ligen Europas gemessen am Anteil der Nicht-Einheimischen (Legionäre) im Mittelfeld. Den höchsten Anteil ausländischer Spieler weist die englische Premier League mit 63,9 % aus (maximal erlaubt sind lt. Homegrown Player Rule 68 %), den geringsten Anteil von 41,9 % die spanische La Liga (siehe dazu Abbildung 8, Statista, 2021b).

Mit der Einführung der Regulierungen durch FIFA und UEFA (Gültigkeit ab den Saisons 2005/06 für Ausbildungsentschädigung und Solidaritätsbeitrag sowie 2006 für die Homegrown Player Rule) ist der *Anteil einheimischer Spieler in der 1. Fußball-Bundesliga* von 51,8 % (2005/06) auf 44,7 % (2021/22) um 7,1 Prozentpunkte gesunken. Zwar wird in der 1. Bundesliga und in den übrigen Big-5-Ligen die Mindestanforderung von 32 % einheimischen Spielern in einer Mannschaft übererfüllt, die Erreichung des originären Ziels der Regulierungen jedoch, eine hohe Anzahl an hochqualifizierten Profifußballspielern – auch für die Nationalmannschaften – hervorzubringen, ist anhand dieses Indikators nicht erkennbar.

Abbildung 8: Anteil ausländischer Spieler in den BIG-5-Ligen: Deutschland, England, Spanien, Frankreich und Italien in der Saison 2021/2022. Eigene Darstellung (Statista, 2021b).



Die steigende Entwicklung des *Anteils ausländischer Spieler* und der damit verbundene konstante Rückgang der Zahl der vom Verein ausgebildeten Spieler (*Einheimische*) innerhalb einer Mannschaft kann mit einer höheren Mobilität, auch früh in der Karriere eines Fußballspielers, erklärt werden (CIES, 2017a). Eine Analyse des CIES zur Entwicklung der Demografie des europäischen Fußballs zeigt, dass von 2009 bis 2017 der Anteil der im Verein ausgebildeten Spieler (club-trained) in den untersuchten Mannschaften (n=466, UEFA members association) in 31 Ligen von 23,2 % auf 18,5 % gesunken ist. 2017 war das achte Jahr mit einem Rückgang in Folge. Auffällig ist, dass in den besten Ligen der Big-5-Länder recht niedrige Anteile der im Verein ausgebildeten Spieler ausgewiesen werden. Spanien liegt mit 22,8 % im Mittelfeld (Platz 13) des Rankings aller 31 betrachteten Ligen (1. Platz = höchster Anteil an im Verein ausgebildeten Spielern), während Frankreich (18,3 %, Platz 19), Deutschland (14,1 %, Platz 24), England (10,3 %, Platz 27) und Italien (8,9 %, Platz 29) im letzten Drittel zu finden sind. Der europäische Fußballarbeitsmarkt wird damit zunehmend internationalisiert, die Mobilität der Spieler nimmt in Europa kaum Rücksicht auf nationale Grenzen (CIES, 2017a).

Das Analyseergebnis der CIES Studie (2017a) verdeutlicht neben der zunehmenden Mobilität von Fußballspielern auch die Unwirksamkeit der Maßnahmen, die zur Förderung der Beschäftigung von Spielern aus dem eigenen Land ergriffen wurden (wie die Homegrown Player Rule; CIES, 2017a). Wohlhabende Ligen und Vereine sind in der Lage, im Rahmen der Regulierungen der FIFA, UEFA und nationalen Verbände trotz der Ausgaben für Ausbildungsentschädigung, und Solidaritätsbeitrag die besten Spieler unabhängig von ihrer Herkunft unter Vertrag zu nehmen. Die Regulierung der Nachwuchsausbildung in Deutschland sollte sich demnach nicht nur auf den heimischen Markt in Deutschland fokussieren, sondern sich den Herausforderungen des europäischen Fußballarbeitsmarkts stellen.

2.1.3.2 Nationale Rahmenbedingungen der Nachwuchsausbildung in Deutschland

Die deutsche Fußballlandschaft wird durch feste Strukturen und Ordnungen reguliert. Der deutsche Fußballbund (DFB) stellt die oberste Instanz des deutschen Fußballs dar. Der Dachverband hat 7,1 Mio. Mitglieder, organisiert in 24 Tsd. Vereinen (DFB, 2021a). Aufgabenbereiche des DFB sind die flächendeckende Verwaltung der 21 Landesverbände, organisiert in fünf Regionalverbänden (DFB, 2018). Des Weiteren ist der DFB für das nationale Regelwerk zuständig und veranstaltet den saisonalen Wettbewerb Deutscher Pokal (Röttger, 2013).

Im Jahre 2000 wurde die Eigenständigkeit aller 36 Profimannschaften mittels Gründung der deutschen Fußball Liga GmbH (DFL) durch den DFB genehmigt (Laupus, 2019). Die Durchführung des Spielbetriebs und die Lizenzierung der 36 Vereine der 1. und 2. Bundesliga gehören zu den wichtigsten Aufgaben der DFL. Darüber hinaus erfolgt durch die Gesellschaft die Vermarktung sowie die Weiterentwicklung der Marke *Bundesliga*. Die Verantwortung für die Rechtevermarktung der Spiele der DFL für Fernseh-, Radio- und Internetübertragungen liegt bei der in 2008 gegründeten 100-prozentigen Tochterfirma *DFL Sports Enterprises* der DFL (DFB, 2013).

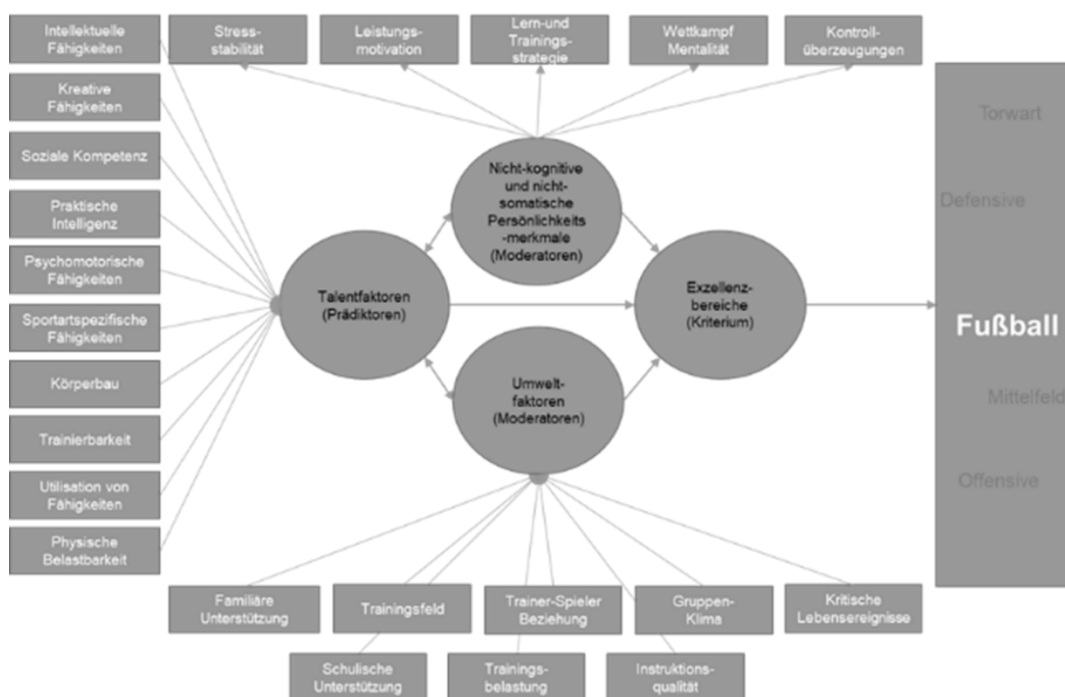
2.1.3.2.1 Talententwicklung in Deutschland auf individueller Ebene

Die *individuelle Ebene* bezieht sich auf Faktoren, die bei der Nachwuchsarbeit auf der Ebene des einzelnen Spielers zur Vorbereitung auf den Lizenzfußball relevant sind. Eine tiefere Analyse dieser Faktoren auf der Ebene eines einzelnen Spielers findet im Rahmen dieser Studie nicht statt. Der Fokus der vorliegenden Arbeit liegt nicht auf dem individuellen Erfolg eines Talents, sondern auf dem ~~Output~~-Outcome des Kollektivs *Nachwuchsleistungszentrum* als Summe aller hervorgebrachten Spielerqualitäten. Zur besseren Einordnung der Erfolgsfaktoren auf kollektiver Ebene wird dennoch ein kurzer Überblick über die wichtigsten Grundlagen der Talentforschung auf individueller Ebene gegeben.

Die Talentforschung zeigt, dass mehrere Faktoren interagieren, um zu Erfolg in der Nachwuchsarbeit zu kommen. Im Folgenden wird auf die wichtigsten Theorien und identifizierte Erfolgsfaktoren für die individuelle Ebene in der Ausbildung von Nachwuchsspielern hingewiesen. Die Forschung im Bereich individueller Talentförderung beschäftigt sich mit der „Suche, Sichtung, Identifikation, Rekrutierung, Entwicklung bis hin zur Förderung von Talenten“ (Wallebohr, 2019, S. 44). Im Rahmen der Nachwuchsförderung haben die zugrundeliegenden Prozesse für die Identifikation, Selektion und Entwicklung von Talenten in den letzten Jahren an Bedeutung zugenommen (Murr, 2019). Die bisherige Praxis, Spieler in Talentförderungsprogrammen durch die Trainerexperten anhand subjektiver Bewertungen zu identifizieren und zu selektieren, wird zunehmend durch objektive Diagnostiken zu physiologischen, anthropometrischen und psychologischen Talentprädiktoren im Fußball ersetzt (Murr, 2019).

Für eine detailliertere Sicht auf die Talentforschung sei auf die entsprechende Literatur verwiesen: Durand-Bush und Salmela (2001), Hohmann et al. (2002), Joch (2001) und Regnier et al. (1993). Eine Liste vorhandener Studien deutscher Talentforscher wurde von Emrich et al. (2005) herausgegeben. Talententwicklung ist laut Gulbin et al., 2010, 2013) abhängig von verschiedenen Trainingsformen, laut Farrow (2010) von der Bedeutung der kognitiven Wahrnehmungsfähigkeit oder auch von unkonventionellen Übungsumgebungen (Araújo et al., 2010). Baker et al. (2010), Weir et al. (2010), Lidor et al. (2010), Ali (2011) und Faude et al. (2010) haben Überblicksarbeiten mit konkreten methodischen Empfehlungen zum Einsatz sowie den Stärken und Schwächen von Tests insbesondere zur Leistungsdiagnostik im Fußball bereitgestellt (Höner, 2012). Die eine Forschungsperspektive analysiert Entwicklungsstufen/-abschnitte der Talententwicklung aus sportlicher Sicht (Bloom, 1985; Ericsson & Lehmann, 1996; Reilly et al., 2000; Hohmann & Seidel, 2003), eine andere die verschiedenen Entwicklungsabschnitte aus Gesamtsicht auf den Sportler, auch im Hinblick auf die nicht sportlichen Aspekte einer Sportlerkarriere (Abbott & Collins, 2004; Côté, 1999; Côté et al., 2007; Durand-Bush & Salmela, 2002; Wylleman et al., 2004). Einen Überblick über Faktoren, die Einfluss auf die Talententwicklung eines Fußballspielers haben können, zeigt die Abbildung 9 mit dem *Modell des sportlichen Talents* nach Seidel (2011).

Abbildung 9: Einflussfaktoren auf die Entwicklung eines Fußballspielers. Eigene Darstellung nach Seidel (2011).

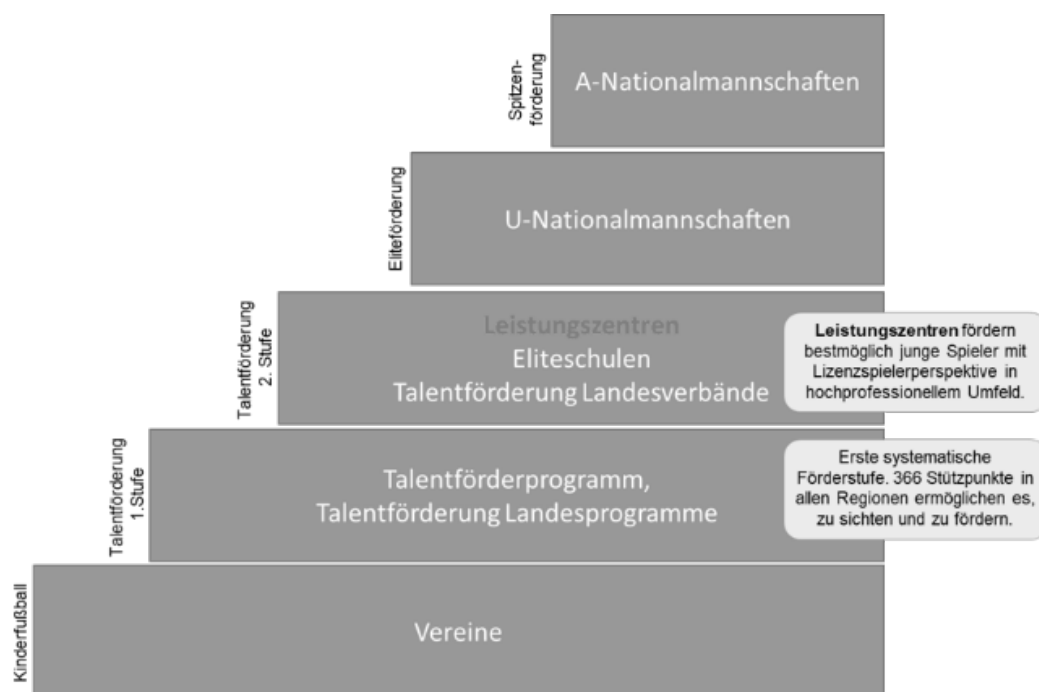


2.1.3.2.2 Organisatorische Ebene – Das Nachwuchsleistungszentren

Auf der organisatorischen Ebene gibt es derzeit 56 Nachwuchsleistungszentren im Bereich des Deutschen Fußball-Bundes (Stand Januar 2021, DFB, 2014b). Eingebettet sind diese in die Förderstufen des DFB-Talentförderprogramms, das in der Saison 2002/03

aufgelegt wurde (siehe Abbildung 10). Die erste Stufe der Talentförderung besteht aus derzeit 366 Stützpunkten; diese bilden die Brücke zwischen der Nachwuchsarbeit in den Vereinen und der Ausbildung von Top-Talenten in den Leistungszentren (zweite Stufe der Talentförderung), in den Eliteschulen des Fußballs und des Sports sowie in den Landesverbänden (DFB, 2014b). Durch die Regulierungen internationaler (FIFA, UEFA) und nationaler Verbände und Organisationen (DFB, DFL) werden gezielt Anreize geschaffen, verstärkt in ein eigenes Nachwuchsleistungszentrum zu investieren. Die FIFA-Regularien fordern die Zahlung einer Entschädigung durch den aufnehmenden Verein an den Ausbildungsverein, wenn ein U23-Spieler einen Profivertrag erhält, und bei internationalen Transfers eine 5-%-Beteiligung an den Transfersummen für an der Ausbildung beteiligte Vereine. Die UEFA wiederum honoriert die Qualität der Arbeit in den Nachwuchsleistungszentren für Clubs der 2. Bundesliga. Die DFL hingegen honoriert die Einsatzminuten eines lokal ausgebildeten Spielers in der U23 zugunsten des Vereins sowie die Ausbildung erstmaliger U23-Lizenzspieler. Der DFB wiederum honoriert sowohl an der Ausbildung beteiligte Amateurvereine für den erstmaligen Einsatz eines Junioren-Nationalspielers als auch die Nachwuchsleistungszentren der Liga 3 und der Regionalliga für die Qualität der Arbeit. Darüber hinaus sehen es die gemeinsamen Regularien des DFB und der DFL vor, dass bei einem Transfer von einem Nachwuchsleistungszentrum zum nächsten Entschädigungszahlungen durch den aufnehmenden Verein an den abgebenden Verein erfolgen (Pabst, o.J.).

Abbildung 10: Talentförderung DFB. Eigene Darstellung nach (DFB, 2021c).



2.1.3.2.3 Kontextuelle Ebene – Das ökonomische Umfeld der Nachwuchsleistungszentren
Nachwuchsleistungszentren in Deutschland sind mit der Einführung der Talentförderprogramme professionalisiert worden. Im Folgenden wird das ökonomische Umfeld anhand der Ausgaben (Kosten) für Nachwuchsarbeit und der potentiellen Einnahmequellen, die direkt der Nachwuchsarbeit zugeschrieben werden können, erläutert.

Ausgaben

Die Ausgaben werden anhand des *durchschnittlichen Investitionsaufkommen* in den Nachwuchsbereich aufgezeigt. Dieses erhöhte sich seit der Saison 2003/04 stetig von 1,33 Mio. Euro auf 5,41 Mio. Euro pro Nachwuchsleistungszentrum der Vereine der 1. und 2. Bundesliga (18 Vereine pro Liga). In den letzten fünf Jahren des Betrachtungszeitraums (2015/16 bis 2019/2020) betrug die Wachstumsrate der Investitionen in die Nachwuchsleistungszentren der 1. und 2. Bundesliga 22,87 %. Die Aufwendungen für Nachwuchsleistungszentren der Vereine der Liga 3 oder Regionalligen sind in diesen Daten nicht berücksichtigt. Seit der Saison 2003/04 investierten die Vereine der 1. und 2. Bundesliga in ihren Nachwuchsbereich insgesamt 1,98 Mrd. Euro (DFL, 2020; DFL, 2021) bei Gesamtumsatzerlösen von 47,09 Mrd. Euro (siehe Abbildung 11).

Abbildung 11: Investitionen in Nachwuchsarbeit 1. und 2. Bundesliga. Eigene Darstellung (DFL, 2020a).



Gemessen am Anteil der Investitionen in den Nachwuchsbereich an den Umsatzerlösen der Vereine zeichnet sich keine Steigerung der Investitionen nach Einführung der FIFA- und UEFA-Regularien zur Förderung des Fußballnachwuchses in 2005 (Ausbildungsentschädigung, Solidaritätsbeitrag) und 2006 (Homegrown Player Rule) ab. Bis auf die Saison 2019/20, in der bedingt durch die COVID-19-Pandemie Umsatzeinbußen von 5,7 % im Vergleich zur Vorsaison hingenommen werden mussten (DFL, 2021), bewegt sich der Anteil der Investitionen konstant um die 4,0 % des Umsatzes (in der Saison 2019/20 liegt der Anteil bei 4,33 % des Umsatzes). Das kontinuierliche Wachstum des Fußballgeschäfts in Europa und auch in Deutschland wird vor allem getrieben durch

steigende Einnahmen aufgrund des Verkauf von TV- und Medienrechten, vermehrtem Sponsoring und absolut steigenden Transfererlösen. Die Entwicklung des Umsatzes der 1. und 2. Bundesliga und die Investitionen in den Nachwuchsbereich je Saison bis 2019/20 zeigt Abbildung 12.

Abbildung 12: Umsatzentwicklung 1. und 2. Bundesliga und ihre Investitionen in den Nachwuchsbereich in Milliarden Euro. Eigene Darstellung (DFL, 2020a).



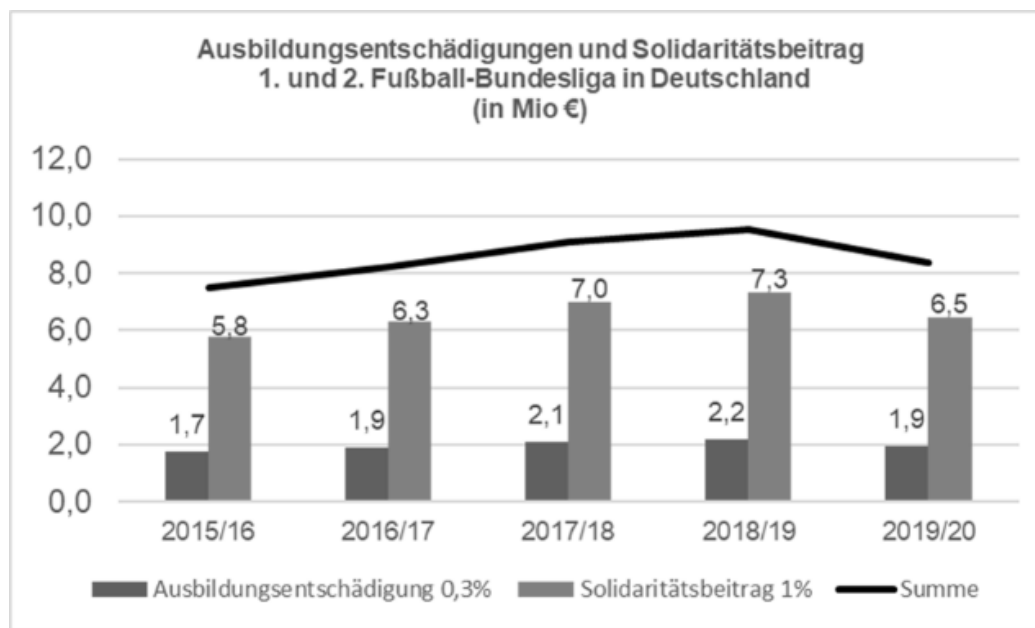
Investitionen in ein Nachwuchsleistungszentrum werden pro Saison auf zwei Hauptbereiche verteilt: Personal und Verwaltung. Der Bereich *Personal* umfasst sämtliche Gehälter für Spieler, Trainer und weitere Angestellte des Nachwuchsleistungszentrums. Unter *Verwaltung* fallen die Kosten für Instandhaltung des Nachwuchsleistungszentrums, Einkauf der Trainingsausrüstung, Investitionen in Trainingsgeräte, Fahrtkosten und Antrittsgelder (Laupus, 2019; Röttger, 2013). Investitionen im Personalbereich, wie vermehrte Festanstellungen, hochqualifizierte Trainer und Erweiterung des Trainerstabs durch Physiotherapeuten oder Spielanalysten, führen zu einer höheren Professionalität der Nachwuchsarbeit. Investitionen des Verwaltungsbereichs in neue Trainingstechnologien, Optimierungen der Nachwuchsleistungszentrum-Infrastruktur und in innovative Trainingsmethoden dienen ebenfalls der Erhöhung des Grads der Professionalisierung. Zusätzlich zu den genannten saisonalen Investitionen erfordern neue Nachwuchsleistungszentren hohe Infrastrukturaufwendungen (Laupus, 2019).

Einnahmen

Nachwuchsleistungszentren werden in der Regel in den Vereinen nicht als eigenständiges Profitcenter geführt, die Einnahmen aus der Nachwuchsarbeit fließen an den Verein. Entschädigungszahlungen und Beteiligungen bei (inter-)nationalen Transfers, bei Verkauf selbstausgebildeter Spieler sowie die Beteiligung an den durch die DFL verwalteten Medienerlösen können aber als Einnahmen zu Controllingzwecken einem Nachwuchsleistungszentrum zugeschrieben werden.

Basierend auf den Transfereinnahmen der Vereine der 1. und 2. Bundesliga von 2015/16 bis 2019/20 werden im Durchschnitt bei Verkauf eines Spielers pro Saison 2,2 Mio. Euro Ausbildungsentschädigung und 6,6 Mio. Euro Solidaritätsbeitrag fällig. Abbildung 13 zeigt die Ausbildungsentschädigungen und Solidaritätsbeiträge für den Betrachtungszeitraum, berechnet mit den weltweit ermittelten Prozentsätzen für Ausbildungsentschädigungen (0,3 %) und Solidaritätsbeiträgen (1 %) der Transfersumme.

Abbildung 13: Ausbildungsentschädigungen und Solidaritätsbeiträge pro Saison 1. und 2. Fußball-Bundesliga. Eigene Darstellung (DFL, 2016; DFL, 2017; DFL, 2018; DFL, 2019b; DFL, 2020a; DFL, 2021a).

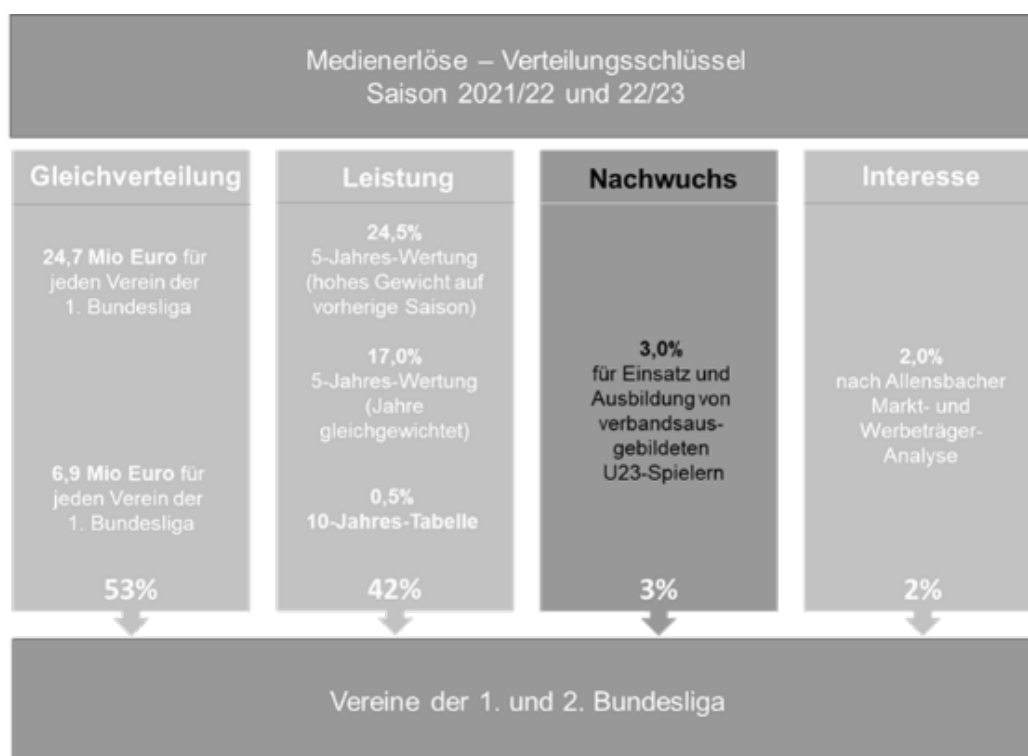


Durch den Verkauf von selbst ausgebildeten Spielern konnten in den vergangenen Jahren von nationalen Fußballvereinen Einnahmen in Höhe von zweistelligen Millionen summen erzielt werden (Rehm-Engel, 2019). Schalke 04 zum Beispiel erzielte durch den Verkauf einzelner im vereinseigenen Nachwuchsleistungszentrum ausgebildeter Spieler ins europäische Ausland (2016 in die englische Premier League: Leroy Sané für 52 Mio. Euro, 2018 in die französische Ligue 1: Thilo Kehrer für 37 Mio. Euro) und innerhalb Deutschlands (2015 Julian Draxler für 43 Mio. Euro) in Summe 130 Mio. Euro (o.A. Transfermarkt.de, 2021). Die Ausbildung von hochqualifizierten Spielern, die sich im europäischen Fußballarbeitsmarkt behaupten, kann hohe Einnahmen aus den Transfers für die Vereine und einen nicht unerheblichen Wert für den Kader des Vereins generieren. Damit gewinnt die Qualität der Ausbildung immer mehr an Bedeutung für die sportlichen und wirtschaftlichen Erfolge der Vereine mit Nachwuchsleistungszentrum (Rehm-Engel, 2019; McKinsey & Company, 2020).

Die Beteiligung an den Medienerlösen erfolgt zentral über die DFL. Alle erwirtschafteten Gelder der Übertragungsrechte werden nach einer Ende des Jahres 2020 beschlossenen Neuregelung eines bereits existierenden Schlüssels verteilt. Die Verteilung der

Medienerlöse basiert auf 4 Säulen: *Gleichverteilung*, *Leistung*, *Nachwuchs* und *Interesse* am Verein (Kicker, 2020). Abbildung 14 zeigt einen Überblick über die Regelungen für die Saisons 2021/22 bis 2022/23. Der Aspekt der Nachwuchsförderung hat im Vergleich zu der vorher bestehenden Regelung einen höheren Stellenwert bekommen. Es werden sowohl die Einsatzzeit junger Talente als auch deren Ausbildung honoriert (siehe dazu auch die Ausführungen zu Anreizsystemen der Verbände in Kapitel 2.1.3.2.2 dieser Arbeit). Die zu erwartenden Erlöse liegen bei ca. 1,1 Mrd. Euro für die Saisons 2021/22 bis 2023/24. In den Saisons 2023/24 und 2024/25 werden voraussichtlich (?) im Durchschnitt 47,5 Mio. Euro für die Nachwuchsarbeit an die Vereine der 1. und 2. Bundesliga verteilt werden, mehr als doppelt so viel als nach der für die Saisons 2017/18 bis 2020/21 bestehenden Systematik mit 23 Mio. Euro pro Saison (DFL, 2020a). Internationale Fernsehgelder aus der Champions League und Übertragungen von Spielen ins Ausland erhöhen die Medienerlöse insgesamt, eine Beteiligung für den Einsatz von Nachwuchsspielern existiert in diesem Zusammenhang nicht (DFL, 2020a).

Abbildung 14: Verteilungsschlüssel Medienerlöse



Die mit den Anreizsystemen verbundenen Einnahmen durch Entschädigungs- und Kompensationszahlung stehen nur in Ausnahmefällen in Relation zu den Ausgaben der über mehrere Jahre gehenden Spielerausbildung im Nachwuchsleistungszentrum (Walkebohr & Daumann, 2019). Laut der Studie der European Club Association (ECA, 2012) stellen Investitionen in die Nachwuchsarbeit eher ein Mittel zur Reduktion der finanziellen Risiken bei der Zusammenstellung des Kaders dar. Darüber hinaus zeigt eine Stu-

die für den Fußball in der Schweiz, dass eine hohe Stabilität und ein niedriges Durchschnittsalter des Kaders eine hohe Bedeutung für langfristig positive Ergebnisse der Nachwuchsarbeit haben (CIES, 2017b). Die Investition in ein Nachwuchsleistungszentrum ist – neben den Liga-Lizenzierungsaufgaben – eher als Langfristinvestment und damit als Teil einer nachhaltigen Strategie im europäischen Profifußball anzusehen (Wal-lebohr, 2019 nach Strack & Baier, 2009). Mehrere Studien über die Talententwicklung, wie zum Beispiel Abbott und Collins (2004), Regnier et al. (1993), haben darauf hingewiesen, dass der Prozess der Talentidentifikation grundsätzlich fehlerhaft ist, nationale Sportverbände aber in hohem Maße weiterhin beträchtliche Ressourcen darin investieren (Baker et al., 2010).

Eine betriebswirtschaftliche Betrachtung anhand des Return on Investment erscheint daher als erforderlich, um die Investitionsstrategien im Nachwuchsbereich zu bewerten.

Ziel der Nachwuchsleistungszentren ist es, wie beschrieben, junge Talente zu entdecken und durch systematische und intensive Förderung auf athletischer und sportlicher Ebene als auch auf intellektueller und persönlicher Ebene auf die Anforderungen des Lizenzfußballs vorzubereiten. Nachwuchsleistungszentren, ihre gewünschte Struktur und erforderliche Kompetenzfelder werden über Lizenzierungsstatuten durch den Deutschen Fußballbund (DFB) und der Deutschen Fußball Liga (DFL) vorgegeben. Eingebettet ist die Nachwuchsausbildung in einen Rahmen mit Regelungen der internationalen Fußball-Verbände (FIFA, UEFA) wie Ausbildungsentschädigungen, Solidaritätsbeiträge, Homegrown Player Rule und Finanzielles Fairplay, die Auswirkungen auf das ökonomische Umfeld der Nachwuchsleistungszentren haben. Die Wirkung der Ausbildung in den Nachwuchsleistungszentren auf den wirtschaftlichen Erfolg kann mittels Evaluation bewertet werden. Der theoretische Rahmen der Evaluationsforschung wird daher im folgenden Kapitel vorgestellt und es werden mögliche Evaluationsdesigns, -standards und Evaluationsmodelle aus der allgemeinen unternehmerischen Praxis näher betrachtet.

2.2 Evaluationsforschung

Lohnt sich das Engagement in der Nachwuchsausbildung des Profifußballs? Kann der tatsächliche Nutzen die Investitionen rechtfertigen? Um Antworten auf diese Fragen zu finden, ist es erforderlich, die mit der Einführung der Nachwuchsleistungszentren initiierten Talentförderprogramme im Verein zu begutachten. Der Wert und die Wirksamkeit der Nachwuchsarbeit in Bezug auf die Ziele des Untersuchungsobjekts sind nachzuweisen und zu optimieren (Mittag & Hager, 2000; Kauffeld, 2016). Hierfür bieten sich Evaluationsstudien an, sie sind erforderlich, um den Erfolg oder Misserfolg von Maßnahmen bzw. Programmen mit Hilfe empirischer Daten zu beurteilen (Kromrey, 2002). Der Begriff der Evaluation deckt zum einen eine spezifische *Fragestellung* ab, mit der ein Sachverhalt oder ein Programm nach vorgegebenen Kriterien bewertet wird, zum anderen ist damit auch ein spezifisches *Design* gemeint, mit dem das zu evaluierende Programm und die „bewirkten Effekte in methodisch kontrollierter Weise miteinander in

Beziehung gesetzt und aus ihrer Perspektive der Handlungsziele auf ihren Erfolg hin bewertet“ (Kromrey, 2002, S. 99) werden. Das Ziel dieser Arbeit ist die Entwicklung eines Evaluationskonzepts, das outcome-orientiert ist, Ergebnisse im Zeitverlauf erhebt, öffentlich zugängliche objektive Daten nutzt, Individual-Leistungs-Indikatoren und geeignete betriebswirtschaftliche Kennzahlen als Maßzahlen einsetzt und mit geringem Aufwand anzuwenden ist. Für das Design dieses Konzepts bieten sich diverse Bedeutungsfacetten, Vorgehensweisen und deren Stärken und Schwächen von Evaluationen an. Im folgenden Kapitel wird daher zunächst der Begriff *Evaluation* geklärt, des Weiteren werden die verschiedenen Evaluationsdesigns und -standards dargestellt sowie die Evaluation von Personalentwicklungsmaßnahmen und ihrer Resultate näher ausgeführt und diskutiert.

2.2.1 Evaluation – Begriffliche Klärung

Es gibt nicht DIE EINE Definition von Evaluation, auf die sich die Wissenschaft im Bereich der Evaluationsforschung einigen konnte. In der Literatur finden sich eine Vielzahl von Beschreibungen, Bedeutungen, Perspektiven und Vorgehensweisen zum Begriff *Evaluation* (Mittag & Hager, 2000). Während Rossi, P.H., Freeman, H.E. & Hofmann, G. (1988) Evaluation als „die systematische Anwendung *sozialwissenschaftlicher Forschungsmethoden* zur Beurteilung der Konzeption, Ausgestaltung, Umsetzung und des Nutzens sozialer Interventionsprogramme“ (Rossi et al., 1988, S.3) definieren, stellt (Balzer, 2005) Evaluation als einen *Prozess* dar, bei dem zielorientiert eine Bewertung eines Evaluationsgegenstands nach relevanten und begründeten Kriterien vorgenommen wird. (Widmer, T. & Rocchi, T. de, 2012) verstehen *Evaluation* als „...*wissenschaftliche Dienstleistung*, die sich mit der systematischen und transparenten Bewertung eines Gegenstands befasst.“ Spiel (2003) fokussiert in ihrer Definition auf eine wissenschaftsgestützte Bewertung von *Effektivität und Effizienz* von Evaluationsgegenständen unter Berücksichtigung geltender Standards. Döring und Bortz (2016) betonen in ihrer Definition die wissenschaftliche Vorgehensweise mittels sozialwissenschaftlicher Methoden unter Berücksichtigung der betroffenen Stakeholder-Gruppen und bestimmter Kriterien zur Bewertung eines Evaluationsgegenstands wie zum Beispiel Effektivität, Effizienz oder Akzeptanz. Demnach ist eine wissenschaftliche Evaluation „...den üblichen Prinzipien der Wissenschaftlichkeit sowie der Wissenschafts- und Forschungsethik verpflichtet, die für die besonderen Bedingungen der Evaluationsforschung zusätzlich in verbindliche Evaluationsstandards als Gütestandards für Evaluationsstudien gefasst wurden.“ (Döring & Bortz, 2016, S. 979). Darüber hinaus weisen sie darauf hin, dass besondere Anforderungen an die Qualität der Evaluierenden gestellt wird. Für weiterführende Einblicke in die Historie und Vielfalt der Evaluationsforschung sei auf die reichhaltige Literatur verwiesen (vgl. Hager et al., 2000; Kromrey, 2001; Rossi et al., 2003; Wottawa & Thierau, 2003).

Effektivität und Effizienz sind die Kriterien, die ergebnisorientierten Evaluationsmodellen in der Personalentwicklung zugrunde liegen (Döring & Bortz, 2016). Über diese Kriterien hinaus haben die Berücksichtigung der Prinzipien der Wissenschaftlichkeit

und Anwendung von Standards eine hohe Bedeutung für die in dieser Arbeit durchgeführte Studie. Daher legt die Autorin dieser Arbeit für die weiteren Ausführungen zur Evaluation des Gegenstands *Nachwuchsleistungszentrum* und zur Vorgehensweise die Definition von (Döring & Bortz, 2016, S. 979) zugrunde.

2.2.2 Evaluationsdesign

Für das Design von Konzepten zur Evaluation von Handlungsprogrammen findet sich in der Literatur zu Evaluationen im sozial- und humanwissenschaftlichen Bereich, ähnlich wie bei der Definition des Begriffs der Evaluation, eine Vielzahl von Aspekten und Prozessbeschreibungen (siehe dazu Mittag & Hager, 2000; Kromrey, 2002; Kauffeld, 2016; Döring & Bortz, 2016; Stockmann & Meyer, 2014). Patton (1997) führt 58 das Evaluationsdesign beeinflussende Zwecke an und weist dazu noch darauf hin, dass damit noch nicht die gesamte Bandbreite erfasst ist. Gemeinsam sind den Zwecken zur Beschreibung des Rahmens im Wesentlichen Ausführungen zu Zielsetzung, Evaluationsgegenstand, Stakeholder, Evaluationstyp, Methode, Erfolgsmaße und Untersuchungszeitpunkte (Mittag & Hager, 2000; Stockmann & Meyer, 2014; Kauffeld, 2016).

2.2.2.1 Evaluationszielsetzung

Evaluationsziele sind von den Zielen des Evaluationsgegenstands abzugrenzen. Während im Allgemeinen das Ziel einer Interventionsmaßnahme oder eines Handlungsprogramms ist, eine Veränderung der Teilnehmer hervorzurufen, besteht das Ziel einer Evaluation darin, die Wirksamkeit der Maßnahme bzw. des Programms zu prüfen (Mittag & Hager, 2000; Mittag & Hasselhorn, 2000). Darüber hinaus kann die Prüfung von Hypothesen in der Implementations- oder Durchführungsforschung Ziel von Evaluationen sein (Mittag & Hasselhorn, 2000). Der empirische Nachweis der Wirksamkeit von Programmen kann als zentrale Aufgabe der Evaluationsforschung angesehen werden (Mittag & Hager, 2000). Zur Präzisierung und zum weiteren Verständnis im Kontext des auszuwählenden Evaluationsdesigns scheint es erforderlich, im Folgenden die Begrifflichkeiten der Programmforschung wie *Programm*, *Programmwirkung*, *Programmwirksamkeit* (*Effektivität*) und *Programmeffizienz* zu erläutern und abzugrenzen.

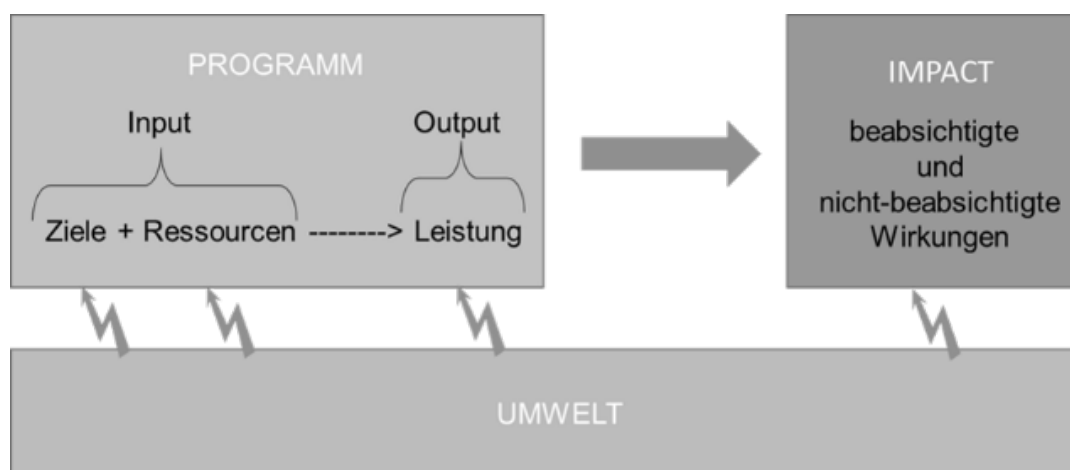
Exkurs Programmforschung:

Unter *Programm* wird in der Sozialwissenschaft ein komplexes Handlungsmodell verstanden, das auf das Erreichen spezifischer Ziele ausgerichtet ist, auf den Zielen abgestimmter Handlungsstrategien basiert und darüber hinaus Ressourcen, finanzieller und personeller Art, einsetzt (Kromrey, 2002 nach Mayntz, 1980). Ziele, Ressourcenbereitstellung und die Durchführung bestimmter Maßnahmen (Leistungserbringung) werden für ein Programm implementiert, um Wirkungen hervorzurufen, die beabsichtigt oder unbeabsichtigt sein können. Beeinflusst wird dieses Gefüge von sog. *Störvariablen* der Programmumwelt (Stockmann & Meyer, 2014). Die Zusammenhänge einer Programmevaluation können Abbildung 15 entnommen werden. Ein Programm ist demnach eine

Einheit, „die mit materiellen und personellen Ressourcen ausgestattet ist und in eine Organisation eingebettet ist, die wiederum Bestandteil eines größeren Systemzusammenhangs ist.“ (Stockmann & Meyer, 2014, S. 76).

Die Leistungserbringung bewirkt bei der Organisation oder ihrem Umfeld Veränderungen beabsichtigter oder nichtbeabsichtigter Art (Stockmann & Meyer, 2014).

Abbildung 15: Programmevaluation Zusammenhänge. Eigene Darstellung nach Stockmann & Meyer (2014).



Unter *Wirkung* wird eine intendierte Zustandsänderung des Individuums durch das Programm verstanden, die beobachtbar, beschreibbar, messbar und kommunizierbar ist und nach plausiblen nachvollziehbaren Kausalitäten bewertet werden kann (Schneider, 2011, S. 16). Wirkungen stellen beabsichtigte und unbeabsichtigte Effekte, die durch das Programm erzeugt werden, dar. Tornow (2005) definiert Wirkung als die beabsichtigte Folge eines Tuns. Schneider (2011) betont mit seiner Definition – Wirkung als „Ergebnis einer Ursache, einer Kausalität“ (ebd., S. 14) – einen Ursache-Wirkungs-Zusammenhang. Die Kausalität sei „in der sozialen Arbeit nicht eindeutig nachweisbar und höchstens durch aufwändige Rekonstruktionen annäherungsweise zu erkunden“ (Schneider, 2011, S. 19). Wirkungen sind als eine „Konstruktion von angenommenen Zusammenhängen“ (Schneider, 2011, S. 28) zu verstehen, die im Kontext des für die Untersuchung relevanten Umfelds betrachtet werden müssen. Bei der Überprüfung der Zielerreichung und der Wirkung einzelner Maßnahmen, wie z. B. Personalentwicklung oder Ausbildung, ist dieser Hintergrund zu berücksichtigen (Gellert-Beckmann, 2020).

Das vorrangige Ziel einer Programmevaluation bzw. -forschung ist nicht, das *Warum* der Wirkung anhand eines Wirkmodells (Prinzipien, Mechanismen, Prozesse) zu erklären, sondern den empirischen Nachweis der Wirksamkeit des Programms oder der Maßnahme zu bewerten (Patry & Perrez, 2000; Mittag & Hager, 2000). Der Grad, mit dem eine beabsichtigte Wirkung erreicht wird, wird als *Wirksamkeit* oder *Effektivität* bezeichnet. Evaluationen im sozialwissenschaftlichen Bereich untersuchen, ob – ungeachtet der

Ursachen im Einzelnen – eine Maßnahme oder ein Programm die gesetzten Ziele erreicht (Tornow, 2005). Evaluationen können sowohl der Wissenserweiterung dienen sowie zu Kontrollzwecken oder zur Entwicklungskontrolle eingesetzt werden (Kromrey, 2002). Jedes der genannten Ziele zieht spezifische Evaluationsdesigns und eingesetzte Methoden und Qualitätskriterien nach sich (Kromrey, 2002). Bei Evaluationen zur *Wissenserweiterung* im wissenschaftlichen Bereich (Forschungsparadigma) werden bei dem Vergleich der beobachteten Veränderungen mit den Zielen des Programms wissenschaftliche Gütekriterien angelegt und es wird versucht, methodisch eindeutige Zusammenhänge der Effekte des Programms unter Berücksichtigung der relevanten Rahmenbedingungen darzustellen (Kromrey, 2001).

Zentrale Bedeutung kommt dabei der Ermittlung desjenigen Anteils an den insgesamt beobachteten Veränderungen zu, „der allein auf die durchgeführte Interventionsmaßnahme zurückzuführen ist (Bestimmung der Nettowirkung eines Programmes“ (Mittag & Hager, 2000, S. 117; Phillips & Schirmer, 2008). Neben dem Programm können auch andere Faktoren bzw. Störvariablen zu den beobachteten Veränderungen geführt haben. Die Wirksamkeit eines Programms kann nur angemessen bewertet werden, wenn der Einfluss anderer Faktoren und Fehlerquellen an den beobachteten Veränderungen berücksichtigt wird (Mittag & Hager, 2000).

Mit einer Evaluation der *Programmeffizienz* werden Programme unter ökonomischen Gesichtspunkten untersucht (Mittag & Hager, 2000). Es werden Kosten und Wirkungen der Programme erfasst und einer Bewertung im Hinblick auf ihre Wirtschaftlichkeit unterzogen. In Zeiten knapper finanzieller Ressourcen ist es erforderlich, zu überprüfen, wie ein Programm auch unter Kostengesichtspunkten einzuschätzen ist. Zusätzlich zu den oben beschriebenen Wirksamkeitsanalysen sind diese Kosten-Effizienz-Betrachtungen (Kosten-Nutzen-Analysen) ebenfalls Grundlage für Entscheidungsprozesse, zum einen als Argumentationshilfe für den Einsatz des untersuchten Programms und zum anderen zur Entscheidungsfindung über eine Auswahl verschiedener Programme. Ausführungen hinsichtlich Kosten-Effizienz-Analysen sind in der Literatur zu Evaluationen selten und werden auch in der Praxis kaum durchgeführt (Mittag & Jerusalem, 1997).

Die Evaluation der Programmeffizienz umfasst die Ermittlung aller Programmkosten, wie z. B. Personal-, Betriebs- und Materialkosten, sowie die Überführung des Gesamtnutzens in monetäre Einheiten, um den Netto-Nutzen (Nutzen minus Kosten des Programms) zu kalkulieren. Eine vielfach eingesetzte Kennzahl für die Berechnung des Verhältnisses der Kosten und des Nutzens eines Programms ist der Return on Investment (Kauffeld, 2016). Der *Return on Investment (ROI)* ist als Standardbegriff zur Bewertung von Kapitalinvestitionen in der Wirtschaft bekannt und kann von Führungskräften eines Unternehmens interpretiert werden, um damit die Effizienz eines Programms zu beurteilen (Kauffeld, 2014; Phillips & Schirmer, 2008).

2.2.2.2 Evaluationsgegenstand

Als Evaluationsgegenstand wird der Untersuchungsgegenstand bezeichnet, „auf den sich die wissenschaftliche Evaluation bezieht“ (Döring & Bortz, 2016, S. 980). Die Gesellschaft für Evaluation e. V. (DeGEval) definiert den Evaluationsgegenstand als „das, was im Rahmen der Evaluation untersucht und bewertet wird und worauf sich mögliche Konsequenzen in Folge der Evaluation beziehen“ (DeGEval, 2021, S. 67). Die Begriffe Evaluationsobjekt, Evaluandum oder Evaluand werden synonym verwendet (Döring & Bortz, 2016). Gollwitzer und Jäger (2009) und Wottawa und Thierau (2003) stellen fest, dass eine breite Spanne an Evaluationsgegenständen wissenschaftlich evaluiert wird, zum Beispiel Bauvorhaben, Evaluationsprojekte selbst (Metaevaluation), Personen, Produkte, akademische Lehre und vieles andere mehr. Einzelne Interventionsmaßnahmen, größere Programme, die individuelle oder kollektive Veränderungen zum Ziel haben, oder auch Organisationen werden als typische Evaluationsgegenstände angesehen (Döring & Bortz, 2016).

2.2.2.3 Evaluationsstakeholder

Döring und Bortz (2016) sehen im Zusammenhang mit der Evaluationsforschung „alle direkt und indirekt von der Nutzung und Wirkung des Evaluationsgegenstandes **Betroffenen** („program beneficiary“ bzw. **Zielgruppe**: Kunden, Patienten, Teilnehmer sowie deren Angehörige, Freunde etc.) sowie alle an der Entwicklung, Umsetzung und Optimierung des Evaluationsgegenstandes **Beteiligten** („program developer/provider“: Produkt- und Maßnahmenentwickler, Maßnahmenverantwortliche, Dienstleistungspersonal, politische Entscheidungsträger etc.)“ (Döring & Bortz, 2016, S. 983, Hervorhebung durch Autoren) als Stakeholder an. Fitzpatrick et al. (2011) benennen die Personen oder Organisationen, die durch die Programmkonzeption und -umsetzung etwas zu verlieren oder zu gewinnen haben, als Stakeholder. Es werden ebenfalls Betroffene und Beteiligte unterschieden. Im Rahmen der Evaluationsforschung sind gemäß der Gütekriterien (Evaluationsstandards) zunächst die für den Evaluationsgegenstand relevanten Stakeholder, wie Auftraggeber, Programmverantwortliche und Zielgruppe des Programms, zu identifizieren und deren Interessen in den Forschungsprozess einzubeziehen (Döring & Bortz, 2016). Fehlt ein konkreter Auftraggeber, tritt bei der Evaluation von Programmen an seine Stelle das *Erkenntnisinteresse* an dem Programm und seine Wirkung, das unabhängig von den Vorgaben des Programms, aber mit Blick auf die potentiellen Nutzer*innen des Programms und der Evaluation untersucht, um konkrete Empfehlungen geben zu können (Schiffler & Hübner, 2000).

2.2.2.4 Evaluationstyp

Evaluationen können sowohl von internen als auch externen Spezialisten durchgeführt werden (Stockmann & Meyer, 2014; Döring & Bortz, 2016). *Interne Evaluationen* sind Evaluationen, die durch die gleiche Organisation durchgeführt werden, die auch das

Programm abgewickelt hat. Innerhalb dieser Klassifikation wird zwischen der *Selbstevaluation* und der *Fremdevaluation* differenziert. Während bei der internen Selbstevaluation die Evaluation durch Programmteilnehmer oder Mitarbeiter der Abteilung, die mit der operativen Durchführung des Programms betraut sind oder waren, erfolgt, wird die Evaluation bei der internen Fremdevaluation von Mitarbeitern aus einer anderen Abteilung der gleichen Organisation durchgeführt. Döring und Bortz (2016) weisen darauf hin, dass der Begriff Selbstevaluation vom Begriff Selbst-einschätzung abzugrenzen ist, da die Sammlung von Daten auch per Befragung und Beobachtung der Teilnehmer am Programm erfolgen kann. Vorteile von *internen Evaluationen* sind der geringe Aufwand, die Sachkenntnis über die Organisation und das Umfeld sowie die schnelle Umsetzung der Ergebnisse der Evaluation (Stockmann & Meyer, 2014). Nachteile der internen Evaluation – mangelnde Methodenkompetenz der Evaluierenden und deren nicht ausreichende Unabhängigkeit und Distanz – können dazu führen, dass Alternativen zu dem untersuchten Programm nicht wahrgenommen werden (Stockmann & Meyer, 2014).

Evaluationen durch externe Spezialist*innen, die nicht an dem untersuchten Programm beteiligt oder von ihm betroffen sind bzw. waren, werden als *Fremdevaluation* bezeichnet. Externe Evaluator*innen haben einerseits die Vorteile, dass sie andere Erfahrungen und besondere Methodenkompetenz sowie ihr Wissen aus anderen Evaluationen einbringen (Döring & Bortz, 2016) und ihnen innerhalb der untersuchten Organisation oft eine zusätzliche Legitimität und Einflussstärke zugesprochen wird. Andererseits kann die mangelnde Sachkenntnis der Externen über das Programm oder das Umfeld dazu führen, dass die Akzeptanz fehlt und die Umsetzung der Ergebnisse in der untersuchten Organisation nicht angenommen wird (Stockmann & Meyer, 2014).

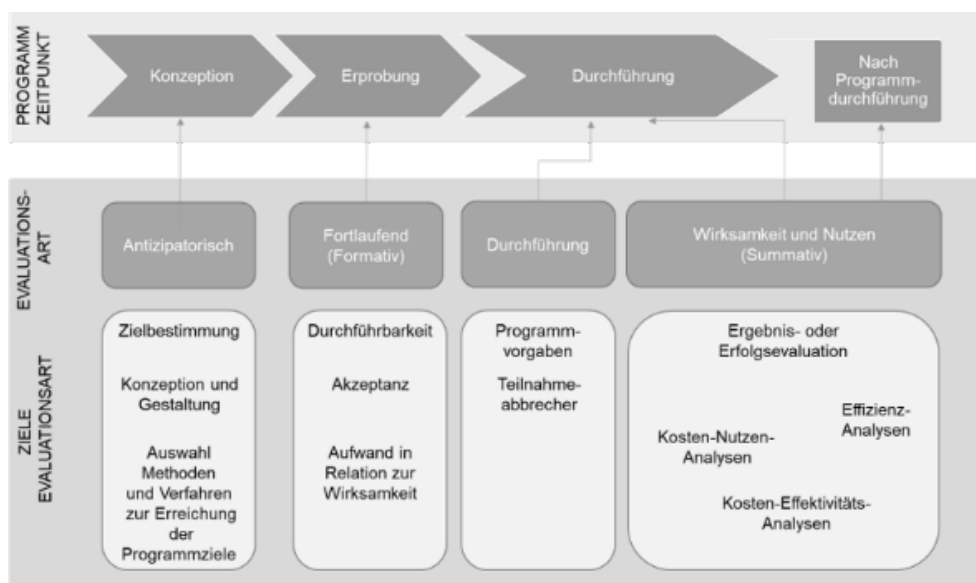
2.2.2.5 Evaluationszeitpunkt

Verschiedene Arten der Evaluation ergeben sich aus den unterschiedlichen Zeitpunkten ihrer Durchführung: vor oder während der Erprobung eines Programms, während der Durchführung der Interventionsmaßnahme und nach der Durchführung des Programms (Kauffeld, 2016; Mittag & Hager, 2000). Es wird daher unterschieden zwischen der (1) Evaluation der Programmkonzeption (antizipatorische Evaluation), der (2) formativen Evaluation als fortlaufende Evaluation zur eventuellen Modifikation des Programmkonzepts, der (3) Evaluation der Programmdurchführung zum Zwecke der Implementations- und Begleitforschung sowie der Prozessevaluation, der (4) Evaluation der Programmwirksamkeit insbesondere zur Ergebnis- oder Erfolgsevaluation und der (5) Evaluation der Programmeffizienz mit Kosten-Nutzen- und Kosten-Effektivitäts-Analysen (Mittag & Hager, 2000).

Abbildung 16 gibt einen Überblick über den Zusammenhang der genannten Evaluationsarten. Zielsetzung und Zeitpunkt der Evaluation bestimmen die Evaluationsstrategie. Ist es Ziel einer Evaluation, die Wirksamkeit und/oder die Effizienz nach der Programmdurchführung zu bewerten, bedient man sich einer summativen ergebnisorientierten Strategie. Mit der summativen Evaluation der Programmwirksamkeit wird zusammenfassend die Qualität und Wirksamkeit, sowie der Nutzen nach Abschluss einer

durchgeführten Maßnahme bzw. eines Programms bestimmt (Kauffeld, 2016; Stangl, 2021). Am Ende des Programms werden die Veränderungen, die durch die Maßnahme bewirkt wurden, erfasst. Evaluiert werden kurzfristige Ergebnisse, wie die Reaktion der Teilnehmer am Programm oder der Wissens-/Kompetenzzuwachs, und langfristige Ergebnisse, wie Veränderungen im Arbeitsverhalten und betriebliche Kennzahlen, z. B. dem Return on Investment (Kauffeld, 2016 nach G. G. Wang & Wilcox, 2006). Das Programm wird gesamthaft als Ursachenfaktor für die hervorgerufenen Veränderungen betrachtet und im Kontrast mit anderen Programmen untersucht. Die unterliegenden Mechanismen, die zu den Veränderungen (Input-orientiert) geführt haben, sind nicht Gegenstand der Untersuchung; es wird daher auch von einer *Black-Box-Evaluation* gesprochen (Döring & Bortz, 2016). Ziel einer summativen Evaluation ist eine Bilanzierung des Erfolgs eines Programms, ergebnisbezogen und outcome-orientiert. Das Ergebnis der Evaluation kann eine Überlegenheit eines Programms gegenüber anderen gleichartig intendierten Programmen darstellen und somit als Entscheidungsgrundlage für die Wahl eines bestimmten Programms dienen (Kauffeld, 2016).

Abbildung 16: Evaluationsarten in Abhängigkeit von Zielsetzung und Zeitpunkt. eigene Darstellung nach Mittag und Hager (2000) sowie Kauffeld (2016).



2.2.2.6 Evaluationskriterien

Mit der Spezifizierung der Messzahlen zur Bewertung, den Bewertungsstandards, den Instrumenten, die zum Einsatz kommen (z. B. Fragebögen, Interviews, Beobachtungen während oder nach der Programmdurchführung, Auswertung von Dokumenten und Daten (interne und externe)), und der Qualität der gesammelten Daten wird dieser Aspekt in der Evaluationsforschung beschrieben (Kauffeld, 2016).

Um eine wissenschaftlich fundierte Bewertung eines Evaluationsgegenstands durchführen zu können, sind klare und messbare Bewertungskriterien zu definieren sowie die relevanten Daten zu erheben und zu interpretieren. Kriterien sollten empirisch erfassbar sein und inhaltlich zur Beantwortung der Fragestellung beitragen sowie sich an den Zielen des Programms orientieren (Döring & Bortz, 2016; Sattler & Sonntag, 2016).

2.2.3 Evaluationsstandards

Um eine hohe Güte wissenschaftlicher Evaluationen eines Programms sicherzustellen, ist die Berücksichtigung allgemeiner und spezifischer Standards für die Evaluationspraxis erforderlich (Schiffler & Hübner, 2000; Döring & Bortz, 2016). Unabhängig vom Gebiet, auf dem eine Evaluation durchgeführt wird (wie zum Beispiel im Bereich sozialer Programme, in Pädagogik, Therapieforchung oder in Betrieben bzw. Organisationen), sind allgemeine Standards zur Evaluationsforschung aufgestellt worden (Rost, 2000). Die existierenden allgemeinen Standards der Grundlagenforschung, wie z. B. für die „Formulierung von Forschungsfragen [...], Literaturrecherche, [...] „Einbettung in einschlägige und konkurrierende Theorien, Versuchsplanung, Messtheorie, Testtheorie, Datenanalyse und Signifikanzprüfung“ (Rost, 2000, S. 130), werden auch bei der Evaluationsforschung angewendet (Rost, 2000). Es ist zu unterscheiden zwischen Standards, die für den Forschungsprozess anzulegen sind, und Standards, die für Evaluationsdurchführung zu berücksichtigen sind (Rossi et al, 1988; Mittag & Hager, 2000; Rost, 2000). Eine detaillierte Betrachtung der Prinzipien wissenschaftlichen Arbeitens im Rahmen der Evaluationsforschung erfolgt an dieser Stelle nicht. Im Rahmen der Beschreibung zur Methodik dieser Arbeit wird aber auf die relevanten einzuhaltenden Standards eingegangen. Fokus der weiteren Ausführungen zu diesem Thema ist der Katalog an Standards zur Evaluationsdurchführung.

Gute Evaluationsstudien orientieren sich an vier grundlegenden Standards zur Evaluationsdurchführung, die wiederum die generellen Qualitätskriterien der Wissenschaft, wie z. B. inhaltliche Relevanz, methodische und ethische Strenge und Präsentationsqualität, aufgreifen (Rost, 2000; Döring & Bortz, 2016). Die Standards des *Joint Committee on Standards for Educational Evaluation (JCSEE)* – eine Institution, die 1975 als Zusammenschluss von Berufsverbänden in den Vereinigten Staaten und Kanada gegründet wurde und sich mit der Qualität der Evaluierungspraxis befasst – gelten als die wohl wichtigsten Evaluationsstandards (Evaluationsstandards der JCSEE, T. Widmer, 2000; Schiffler & Hübner, 2000). Basierend auf den Arbeiten der *Evaluation Research Society (ERS)* hat das *JCSEE* u. a. Standards für Programmevaluierungen (1994), Personalevaluierungen (1988) und Evaluierungen von Studierenden (2003) entwickelt (Schiffler & Hübner, 2000; Stufflebeam, 2004). Die Gesellschaft für Evaluation e. V. (DeGEval) hat diese Standards für den deutschsprachigen Raum erstmalig 2001 mit den *Standards für Evaluation* umgesetzt. Ein regelmäßiger Revisionsprozess führte im Jahre 2016 zur Verabschiedung der im Folgenden erläuterten aktuellen Fassung der Standards für Evaluation (DeGEval, 2017b). Falls nicht anders ausgewiesen, basieren die folgenden Ausführungen auf der Quelle DeGEval (2017b).

Evaluationen sollen demnach – in Analogie zu den vier generellen Qualitätskriterien der Wissenschaft (siehe Anfang dieses Abschnitts) – vier grundlegenden Eigenschaften genügen: der *Nützlichkeit* (utility), der *Durchführbarkeit* (feasibility), der *Fairness* (propriety) und der *Genauigkeit* (accuracy) (Schiffler & Hübner, 2000).

Ziel des Standards *Nützlichkeit* ist es, zu gewährleisten, dass sich die Evaluation am erklärten Zweck und Informationsbedarf der Nutzer*innen ausrichtet und die Bewertungskriterien transparent zugänglich sind. Die Berücksichtigung der Interessen und Bedürfnisse der Stakeholder, die hohe Qualifikation der Evaluators*innen sowie die Qualität der Ergebnisberichte werden ebenfalls in dieser Kategorie geregelt (Döring & Bortz, 2016).

Ziel des Standards *Durchführbarkeit* ist es, zu gewährleisten, dass das Verfahren der Evaluation, einschließlich der Verfahren zur Beschaffung notwendiger Informationen, einerseits professionell und entsprechend den Erfordernissen umgesetzt wird und andererseits der sowohl zeitliche als auch finanzielle Aufwand für die Stakeholder in einem guten Verhältnis zum Nutzen der Evaluation gehalten wird. Darüber hinaus sollte eine hohe Akzeptanz der Stakeholder in Bezug auf die Abläufe und die Ergebnisse der Evaluation erreicht werden.

Ziel des Standards *Fairness* ist es, zu gewährleisten, dass Rechte, Pflichten und Würde der Stakeholder der Evaluation festgehalten werden und geschützt sind, der Evaluationsgegenstand umfassend und fair geprüft wird, eine unabhängige Durchführung und Berichterstattung erfolgt und die Ergebnisse allen Stakeholdern zugänglich gemacht wird.

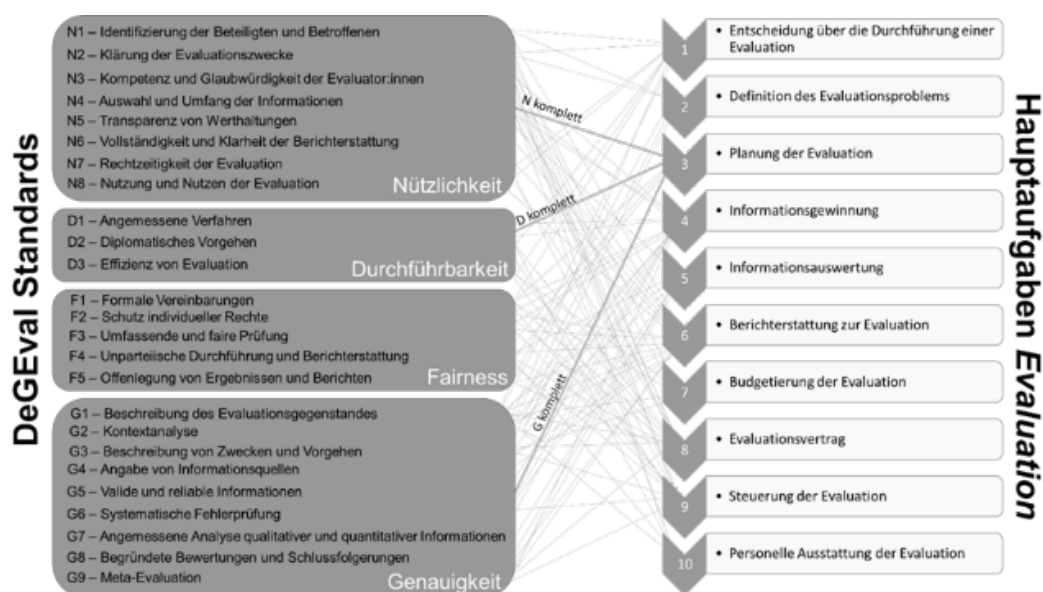
Ziel des Standards *Genauigkeit* ist es, zu gewährleisten, dass Konzept und Umsetzung des Evaluationsgegenstands, Kontext, Ziele, Fragestellung und Vorgehen (inkl. der eingesetzten Methoden) sowie Informationsquellen beschrieben, dokumentiert und analysiert werden, die Zuverlässigkeit der gewonnenen Daten und ihre Gültigkeit bezogen auf die Beantwortung der Evaluationsfragen sichergestellt sind, eine systematische Fehlerprüfung und angemessene Analyse aller hervorgebrachten Informationen erfolgt, die Bewertungen und Schlussfolgerungen auf spezifischen Kriterien und Zielwerten basieren und Schlussfolgerungen auf Grundlage der erhobenen und analysierten Daten begründet werden.

Darüber hinaus ist sicherzustellen, dass alle Dokumentationen zugänglich gemacht werden, um eine Metaevaluation durchführen zu können. Metaevaluationen schätzen ein, inwieweit die zu bewertende Evaluation in Übereinstimmung mit den Standards für gute Evaluationen durchgeführt wurde (Stockmann & Meyer, 2014). Die von der DeGEval festgelegten vier Standards mit ihren genannten untergeordneten einzelnen Standards dienen als Gütekriterien für die Bewertung der Evaluation (Metaevaluation). Nicht jeder Einzelstandard kann bei jeder Evaluation erfüllt sein. Vor Start einer Evaluation sind Überlegungen anzustellen, wie oben angeführte Standards sichergestellt und dokumentiert werden, um im Anschluss eine Bewertung der Evaluation durchführen zu können (Döring & Bortz, 2016). Scriven (2007) empfiehlt für die praktische Ausgestaltung der

Evaluation Checklistenverfahren, die die wesentlichen Elemente zur Erfüllung von Evaluationsstandards für verschiedene Aspekte einer Evaluation enthält. Als Beispiel sei hier eine *Evaluation Design Checklist* von Stufflebeam (2004) angeführt. Auch die DeGEval unterstützt die Professionalisierung der Evaluationspraxis mit einer checklistenähnlichen Einordnung der Einzelstandards zu den Hauptaufgaben einer Evaluation (siehe Abbildung 17). Es wird den Stakeholdern einer Evaluation empfohlen, die aufgeführten Standards bei der Ausgestaltung der konkreten Arbeitsschritte zu berücksichtigen, um insbesondere die Qualität einer Evaluation zu sichern (DeGEval, 2017b).

Das Feld der Personalentwicklungsmaßnahmen und -programme, wie auch bei der Nachwuchsausbildung im Fußball, ist durch eine hohe Komplexität und Mehrdimensionalität gekennzeichnet, die bei der Planung einer Evaluation Berücksichtigung finden müssen. Daher ist es unerlässlich, den Evaluationsprozess strukturiert und transparent zu gestalten. Die DeGEval (2017b) schlägt vor, nach der Entscheidung über die Durchführung einer Evaluation in einem ersten Schritt das Evaluationsproblem zu definieren und dann in die Planung der Evaluation konkret einzusteigen. Balzer und Beywl (2018, S. 31) empfehlen einen Evaluationsprozess in zehn Schritten: „1. Bestimmung des Evaluationsgegenstands; 2. Bestimmung der interessierten Akteure und die Rolle der Evaluierenden; 3. Bestimmung der Evaluationszwecke und Fragestellungen; 4. Festlegung der Bewertungskriterien; 5. Auswahl von Erhebungsdesign und -methoden; 6. Durchführung der Erhebung, 7. Datenauswertung, Interpretation und Bewertungssynthese; 8. Berichterstattung; 9. Nutzung der Ergebnisse; 10. Evaluation der Evaluation.“ Farrokhzad und Mäder (2014, 120f) schlagen ebenfalls eine Checkliste für Evaluationen vor, die allerdings nur sechs Themenfelder enthält:

Abbildung 17: DeGEval-Standards und ihre Anwendbarkeit auf die zehn Hauptaufgaben für eine Evaluation. Eigene Darstellung nach DeGEval (2017b).



„Evaluation fokussieren; Informationen gewinnen; Daten organisieren; Daten analysieren; Berichten der Informationen; Ausführung der Evaluation.“ Allen Vorschlägen ist gemein, dass spezifische zentrale Arbeitsschritte und die damit verbundenen Inhalte, Vorgehensweisen und Standards erforderlich sind, um wissenschaftlich fundierte und gegenstandsbezogene Evaluationsergebnisse zu erhalten (Pfeil & Müller, 2020).

2.2.4 Evaluation von Personalentwicklungsmaßnahmen

Für den Begriff Personalentwicklung existieren in der wissenschaftlichen Literatur diverse Definitionen, auch in der Praxis findet sich keine einheitliche Begriffsbestimmung, verdeutlicht hat das bereits Neuberger (1994) mit einer Auflistung von 18 unterschiedlichen Definitionen von Personalentwicklung. Sowohl „Berufsausbildung, Qualifizierungsmaßnahmen, Weiterbildung, Lernen am Arbeitsplatz, Seminarteilnahmen, aber auch Coaching, Wissenstransfer, Karriereplanung, lebensphasenorientiertes Lernen, optimale Arbeitsplatzgestaltung und viele andere Begriffe“ (Bartscher & Nissen, 2018, 52604/version-384860) werden unter Personalentwicklung verstanden. Es erscheint daher sinnvoll, zunächst eine Klärung des Begriffs *Personalentwicklung* herbeizuführen und ihre Ziele zu erläutern. M. Becker (2013) definiert Personalentwicklung als „... alle Maßnahmen der Bildung, der Förderung und der Organisationsentwicklung, die von einer Person oder Organisation zur Erreichung spezieller Zwecke zielgerichtet, systematisch und methodisch geplant, realisiert und evaluiert werden.“ (M. Becker, 2013, S. 5). Laut Solga et al. (2011) beinhaltet Personalentwicklung „sämtliche Maßnahmen zur systematischen Förderung der beruflichen Handlungskompetenz von Menschen, die in einer und für eine Organisation arbeitstätig sind.“ (Solga et al., 2011, S. 19). Bezugnehmend auf die angeführten Definitionen werden im Folgenden die Begriffe *Bildung* und *berufliche Handlungskompetenz* erläutert. Bildung umfasst „Wissensbestände, Kompetenzen und Qualifikationen“ [...] „die als Humankapital auf dem Arbeitsmarkt angeboten werden“, [...] und „Wissen über gesellschaftliche Vorgänge und die Kompetenz, gesellschaftliche Vorgaben sinnvoll verarbeiten zu können“ (R. Becker, 2017, S. 5). Die Entwicklung der Gesamtheit der Fähigkeiten folgt einem längerfristigen permanenten Entwicklungsprozess und wird von individuellen Veranlagungen und sozialem Umfeld des Individuums geprägt. „Berufliche Handlungskompetenz umfasst alle Fähigkeiten, Fertigkeiten und Wissensbestände, die eine Person, ein Team oder eine Organisation bei der Bewältigung konkreter sowie vertrauter als auch neuartiger Arbeitsaufgaben handlungs- und reaktionsfähig machen und sich in der erfolgreichen Bewältigung konkreter Arbeitsanforderungen zeigen“ (Kauffeld & Paulsen, 2018, S. 14). Es ist das Vermögen eines Individuums, die beruflichen Herausforderungen im spezifischen Arbeitsmarktkontext selbstorganisiert zu meistern.

Ziel von Nachwuchsleistungszentren im deutschen Profifußball ist es, junge Talente für den Arbeitsmarkt *Nationaler und internationaler Lizenzfußball* auszubilden. „Im Rahmen der übergeordneten Ziele *Weltspitze in Bundesliga und Nationalmannschaften* sowie *Optimale Ausschöpfung des Talentpools* des DFB und der DFL, sollen die besten Talente durch eine systematische Förderung auf spätere Anforderungen im Lizenzfußball vorbereitet werden“

(Lanwehr et al., 2021). Um eine optimale Ausbildung für die verschiedenen Entwicklungsverläufe der Talente anzubieten, sind professionelle Ausbildungsprogramme etabliert worden, die zur spielerischen und persönlichen Entwicklung des Talents beitragen (DFB, 2021c). Damit entsprechen die Programme der Nachwuchsleistungszentren im weiteren Sinne den oben angeführten Definitionen zu *Personalentwicklung*. In vielen Bereichen der Soziologie, insbesondere bei der Analyse von formalen Organisationen und Beschäftigung, wird Bildung als kausale Variable berücksichtigt und ihr Outcome, z. B. an Leistung, Erfolg und Performance, gemessen (R. Becker, 2017). In der Wirtschaft sehen sich Personalarbeit, Ausbildung und Personalentwicklung vermehrt einem stärker werdenden Rechtfertigungsdruck ausgesetzt. Es wird vom Management und vom Umfeld erwartet, nachzuweisen, dass die Maßnahmen und Programme die angestrebten Ziele erreicht haben, erfolgreich sind und ein Nutzen für das Unternehmen oder die Gesellschaft geschaffen wird (Sonntag & Stegmaier, 2016).

Im Kontext *Personalentwicklung* umfasst Evaluationsforschung eine empirisch fundierte Bewertung von Personalentwicklungsmaßnahmen. Der Fokus der Evaluation kann auch in diesem Kontext sowohl auf der Konzeption der konkreten Durchführung, der Wirksamkeit als auch der Kosten-Nutzen-Relation eines Personalentwicklungsprogramms liegen (Sattler & Sonntag, 2016). Evaluationsergebnisse von Personalentwicklungsmaßnahmen können laut (J. D. Kirkpatrick & Kirkpatrick W., 2016) die Kosten des Programms legitimieren, als fundierte Entscheidungsgrundlage für die Weiterführung oder Beendigung des Programms dienen und auch zur Identifizierung von Optimierungspotenzialen beitragen.

Zur Qualitätssicherung und -steigerung von Personalentwicklungsmaßnahmen bzw. Programmen sind methodisch fundierte *Evaluationen* erforderlich. Zur Erhaltung der Wettbewerbsfähigkeit einer Organisation bzw. eines Unternehmens wird von (Sattler & Sonntag, 2016) empfohlen, eine zielgerichtete Evaluation von Personalentwicklungsmaßnahmen durchzuführen. F. G. Becker (2005) fordert eine Evaluation sogar als die „*Conditio sine qua non* einer professionellen Personalarbeit“ (S. 48), da eine Datenbasis zur Beurteilung und Optimierung der Personalentwicklungsmaßnahmen geschaffen werden muss (Solga, 2011). Salas et al. (1997) heben hervor, dass Organisationen und Unternehmen, die beträchtliche Ressourcen (Personal, Finanzen) in die Ausbildung ihrer Arbeitskräfte investieren, wissen sollten, welchen Wert die Ausbildung hat und ob die Ausbildung effektiv ist (siehe dazu auch Hochholding et al., 2008 und Kauffeld, 2010). Viele Organisationen oder Unternehmen führen jedoch keine systematischen Evaluationen durch, obwohl die hohen Kosten bzw. die betriebswirtschaftliche Relevanz der Personalentwicklung gesehen werden. Gründe dafür sind häufig, dass obligatorische Evaluationen in der Personalentwicklungsstrategie nicht von Beginn an verankert sind, Evaluationsexpertise im Unternehmen fehlt oder der zeitliche und finanzielle Aufwand, im Rahmen einer Kosten-Nutzen-Betrachtung eines Evaluationsprojektes, als zu hoch erachtet wird (Kabst & Wehner, 2010; Sattler & Sonntag, 2016).

Für die Programmforschung im Bereich der Personalentwicklung werden einerseits allgemeine Evaluationsstrategien, -standards und -methoden aus anderen Disziplinen auf

den Programmkontext übertragen, andererseits spezifische Evaluationsstrategien und -modelle entwickelt und optimiert (Hochholdinger et al., 2008; Rowold, 2008). Evaluationsstudien, die die Wirksamkeit eines konkreten Programms untersuchen, stellen den Erfolg und den Nutzen des überprüften Programms und nicht den Erfolg und Nutzen des Evaluationskonzepts selbst in den Fokus der Untersuchung (Saborowski & Muellerbuchhof, 2010; Lang & Otto, 2010). Die Wirksamkeit und der Erfolg von Personalentwicklungsmaßnahmen werden nach der Durchführungsphase evaluiert (Kontrollfunktion, Summative Evaluation). Ergebnisse dienen als Entscheidungsgrundlage, um Aussagen zur Eignung der Maßnahme treffen zu können und Ansatzpunkte zur Weiterentwicklung zu identifizieren (Gollwitzer & Jäger, 2009). In aktuellen Programmstudien wird überwiegend die summative Herangehensweise zur Überprüfung der Wirksamkeit unter systematischer Anwendung verschiedener Messinstrumente und -methoden angewandt (Tirre, 2011).

In der Literatur finden sich zahlreiche verschiedene Modelle zur Evaluation von Personalentwicklungsmaßnahmen, als das am häufigsten eingesetzte ist das Vier-Ebenen-Modell von Kirkpatrick zu nennen (Lanwehr et al., 2021). Um die Entwicklung eines Evaluierungsmodells für die Ausbildung in einem Nachwuchsleistungszentrum besser nachvollziehen zu können, werden die gängigsten Modelle im Folgenden erläutert und darüber hinaus in den Kontext des Ziels der vorliegenden Arbeit gesetzt. Die vorgestellten Modelle werden daher insbesondere im Hinblick auf ihren Beitrag bezüglich der Bewertung der Effizienz und Effektivität von Personalentwicklungsmaßnahmen diskutiert.

2.2.4.1 Modelle für die Evaluation von Personalentwicklungsmaßnahmen

Zu den wichtigsten theoretischen Konzepten und Modellen zur Evaluation von Personalentwicklungsmaßnahmen gehört das Ende der 1950er Jahre entwickelte und im Laufe von Jahrzehnten weiterentwickelte Vier-Ebenen-Modell von (Kirkpatrick, 1959_1, 1959_2, 1960_1, 1960_2, 1967, 1998). Dieses Modell sowie dessen implizite Wirkungsannahmen bilden die Grundlage für die Entwicklung von Evaluationsmodellen in der Personalentwicklung (siehe dazu P. B. Warr et al., 1970; Hamblin, 1974; Brinkerhoff, 1987; Kaufman & Keller, 1994; Phillips & Phillips, 1994-2001; Kellner, 2006). Auch Autoren wie Kraiger et al. (1993), Holton (1996), Alliger et al. (1997) und Day et al. (2001) haben Evaluationsmodelle für Trainings- und Personalentwicklungsmaßnahmen entwickelt. Das Kirkpatrick-Modell ist in dem Bereich aber nach wie vor das Evaluationsmodell, das am bekanntesten ist und am häufigsten eingesetzt wird (Alliger et al., 1997; Kraiger & Jung, 1997; Salas & Cannon-Bowers, 2001; Goldstein & Ford, 2002; Reio et al., 2017).

2.2.4.1.1 Das Vier-Ebenen Modell von Kirkpatrick

Die von Kirkpatrick ursprünglich als Richtlinien zur praktischen Durchführung von Trainingsevaluationen aufgestellten Wirkungsannahmen (?) avancierten zum Modell (Kraiger, 2003). Kirkpatrick (1959_1) unterscheidet bei der Evaluation von Trainings

die vier folgenden Ebenen: Reaction (Teilnehmerzufriedenheit), Learning (Wissens-transfer), Behavior (Praxistransfer) und Results (Unternehmenserfolg) (siehe dazu Abbildung 1). Es werden für jede Ebene Evaluationskriterien spezifiziert und Instrumente zu ihrer Erhebung vorgeschlagen. Im originären Modell stehen die vier Ebenen in einem hierarchisch-kausalen Abhängigkeitsverhältnis. Annahme ist, dass der Informationsgehalt nach jeder Ebene steigt und das Ergebnis der jeweils vorherigen Ebene eine notwendige, aber nicht hinreichende Bedingung für das Ergebnis auf der nächsthöheren Ebene darstellt, es wird damit von positiven Korrelationen zwischen den Ebenen ausgegangen (D. L. Kirkpatrick & Kirkpatrick, 2006). Im Folgenden werden die vier Ebenen des Kirkpatrick-Modells vorgestellt und die Häufigkeit ihrer Anwendung in Theorie und Praxis aufgezeigt.

Ebene 1 – Reaction – Teilnehmerzufriedenheit

Die Überprüfung der Wirksamkeit eines Programms beginnt auf Ebene 1 mit der Erhebung der *Teilnehmerzufriedenheit*, häufig untergliedert nach Zufriedenheit mit dem Trainer, den Inhalten und der Didaktik etc. (D. L. Kirkpatrick & Kirkpatrick, 2006; Kauffeld, 2010). Das subjektive Erleben des Programmteilnehmers wird in der Regel mittels eines standardisierten Fragebogens (sog. Happy-Sheets-Befragungen) gemessen (Kauffeld, 2010; Solga, 2011). D. L. Kirkpatrick und Kirkpatrick (2006) betonen eine hohe Bedeutung dieser Ebene für die Beurteilung des Erfolgs der Trainingsmaßnahme und die vergleichsweise einfache Evaluation der Reaktionsebene.

Ebene 2 – Learning – Wissenstransfer

Die Überprüfung des Inhalts eines Programms erfolgt auf der Ebene 2 mit der Untersuchung dessen, was die Mitarbeiter durch die Teilnahme an dem Programm gelernt haben. Bezogen auf Kirkpatrick's Definition des Lernens „als das Ausmaß, in dem die Teilnehmer ihre Einstellung ändern, ihr Wissen verbessern und/oder ihre Fähigkeiten als Ergebnis der Teilnahme am Programm steigern“ (D. L. Kirkpatrick & Kirkpatrick, 2006, S. 22), wird die Annahme geäußert, dass erworbene Kenntnisse und Fähigkeiten zwangsläufig mit Verhaltensänderungen oder Leistung am Arbeitsplatz gleichzusetzen sind, was wissenschaftlich aber nicht belegt ist (Reio et al., 2017 nach Strunk, 1999). D. L. Kirkpatrick und Kirkpatrick (2006) betonen, dass die Bewertung des Lernens wichtig ist, da es ohne Lernen keine Verhaltensänderung gäbe. Aus der Literatur geht hervor, dass Evaluierungen der Ebene 2 nach wie vor häufig verwendet werden, um Personalentwicklungsprogramme zu bewerten und so nachzuweisen, dass Teilnehmer des Programms Wissen und Fähigkeiten erworben haben, die umsetzbar sind (Reio et al., 2017).

Ebene 3 – Behavior – Praxistransfer

Die Überprüfung der Verhaltensänderung der Programmteilnehmer wird der Ebene 3 zugeordnet, indem gemessen wird, inwieweit die neu erworbenen Kenntnisse und Fähigkeiten angewendet werden (D. L. Kirkpatrick, 1960_1). Ohne Anwendung des Er-

lernten entstehen keine endgültigen Resultate auf der organisatorischen Ebene. Die Bewertung des Praxistransfers stellt sich als komplizierter, schwieriger und zeitaufwändiger als die Bewertung der Ebenen 1 und 2 dar, sodass sie weniger häufig durchgeführt wird (D. L. Kirkpatrick & Kirkpatrick, 2006).

Ebene 4 – Results – Unternehmenserfolg

Die Überprüfung der tatsächlichen Veränderung der organisatorischen Leistung geht auf der Ebene 4 vonstatten, dort müssen zur Bewertung des Resultats Unternehmenskennzahlen erhoben werden (D. L. Kirkpatrick & Kirkpatrick, 2005). Dazu werden Kennzahlen (KPIs) betrachtet, die für die jeweilige Organisation als Indikatoren des Unternehmenserfolgs angesehen werden, wie z. B. Kostenreduktion oder die Steigerung der Produktivität (Solga, 2011). Darüber hinaus werden auf dieser Ebene auch weitere Faktoren wie Arbeits- und Organisationsklima, Kundenzufriedenheit und Leistungsqualität überprüft (Phillips & Schirmer, 2008). In der Literatur wird diese Ebene als die am schwierigsten zu bewertende eingeschätzt (D. L. Kirkpatrick, 1960_2; D. L. Kirkpatrick, 1998; Phillips, 1996; Werner & DeSimone, 2005). D. L. Kirkpatrick und Kirkpatrick (2005) erklären, dass die Messung von Ergebnissen anhand von z. B. geringerem Turn-Over, verbesserter Produktivität, verbesserter Qualität, Verringerung von Verschwendung, Verringerung des Zeitaufwands, Umsatzsteigerung, Kostensenkung, Gewinnsteigerung und Investitionsrentabilität (ROI) leichter ist als eine Ergebnismessung von Veränderungen z. B. im Führungsverhalten, in der Kommunikation, der Motivation oder des Zeitmanagements (D. L. Kirkpatrick, 1998).

Seit der Etablierung des Vier-Ebenen-Modells erfolgt in der Forschung eine kritische Diskussion des Modells mit seinen immanenten Annahmen. Alliger und Janak (1989) kommen auf der Basis von zwölf zu dem Zeitpunkt verfügbaren relevanten Studien zu dem Ergebnis, dass die vier Ebenen nur schwach korrelieren. Eine Bestätigung dieser Ergebnisse erfolgte durch eine weitere Analyse auf der Basis von 34 Studien durch Alliger et al. (1997). Colquitt et al. (2000) haben 106 empirische Studien analysiert und nachgewiesen, dass die Abhängigkeiten zwischen den vier Ebenen lediglich geringfügig ausgeprägt sind. Obwohl die Wirkungsannahmen damit empirisch nicht nachweisbar sind, wird in vielen, von anderen Autoren entwickelten Modellen weiterhin von einer kausalen Verknüpfung ausgegangen (Holton, 1996; Cannon-Bowers et al., 1995). Eine empirische Überprüfung der Wirkungsannahmen durch Gessler und Sebe-Opfermann (2011) ergab ebenfalls, dass „die Einschätzung, die Zufriedenheit der Teilnehmenden sei ein hinreichendes (Output-)Maß, um die Qualität erbrachter Dienstleistungen“ im Rahmen eines Programms „unzutreffend ist, da die Zufriedenheitswerte keinen Zusammenhang aufweisen mit dem Lernerfolg, respektive Transfererfolg und damit keinen Aufschluss geben über den tatsächlich erreichten Lernerfolg, respektive Transfererfolg. Die derzeitige Praxis der Weiterbildungs-evaluation sowie die Vorgaben der normativen Qualitätsmanagementsysteme sind zur Erfassung des Outputs von Weiterbildung bislang unzureichend und wären entsprechend weiterzuentwickeln“ (Gessler & Sebe-Opfermann, 2011, S. 277).

Im Folgenden werden die weiteren auf Basis dieses Vier-Ebenen-Modells entwickelten Modelle zur Evaluation von Personalentwicklungsmaßnahmen näher betrachtet.

2.2.4.1.2 Brinkerhoff's six-stage model (1987)

Brinkerhoff (1987) fordert über das Kirkpatrick-Modell hinaus eine systematische Evaluation, die bereits Teil der Planung, der Entwicklung und der Durchführung des Programms zur Personalentwicklungsmaßnahme ist. Sein sechsstufiges Evaluierungsmodell beginnt mit der Bedarfsanalyse und der Ermittlung der *Programmziele*. Es folgt auf der zweiten Stufe die *Programmgestaltung*, im Anschluss daran wird – ähnlich zur Ebene 1 von Kirkpatrick – die *Programmdurchführung* evaluiert. Stufe 4 ist identisch mit der Ebene 2 von Kirkpatrick, sie bewertet das *Lernen*. Auf Stufe 5 erfolgt die Bewertung des *Verhaltens*, ähnlich wie auf der Ebene 3 von Kirkpatrick. Auf der Stufe sechs wird – wie auf der Ebene 4 von Kirkpatrick – bewertet, wie sich das Gelernte auf die *Ergebnisse* auswirkt. Annahme dabei ist, dass alle aufgelegten Programme ein positives Ergebnis, einen positiven Nutzen für die Organisation oder das Unternehmen hervorbringen (Brinkerhoff, 1987). Dieses Six-Stage-Modell ist im Wesentlichen deckungsgleich mit dem Vier-Ebenen-Modell von Kirkpatrick, auch wenn mit der Stufe 1 die Funktionen der Programmgestaltung bewertet werden (Reio et al., 2017). Stufe 6 sieht neben der Bewertung nicht-monetärer Ergebnisse auch die Feststellung des wirtschaftlichen Erfolgs mittels eines Kosten-Nutzen-Vergleichs des untersuchten Programms vor (Brinkerhoff, 1987).

2.2.4.1.3 IPO-Modell (Bushnell, 1990)

Auch Bushnell (1990) hat aus seiner Kritik an Kirkpatricks Modell – es werde lediglich die Phase nach der Durchführung eines Programms berücksichtigt – heraus ein Evaluierungsmodell entwickelt, das einen Ansatz zur Bewertung der Ausbildungseffizienz darstellt (Reio et al., 2017). Das Modell verfügt über drei Stufen – Input, Prozess und *Output/Outcome* (IPO). Alle Elemente, die sich auf die Wirksamkeit des Programms auswirken können, z. B. die Kompetenz des Ausbilders oder Infrastruktur, werden als *Input* gesehen. Auf der Stufe *Prozess* wird das Programm geplant, gestaltet, entwickelt und durchgeführt. Der *Output* bezieht sich auf den kurzfristigen Nutzen, wie er mit den ersten drei Ebenen Kirkpatricks (Teilnehmerzufriedenheit, Wissens- und Praxistransfer) evaluiert wird. Darüber hinaus wird Kirkpatricks vierte Ebene (Unternehmenserfolg) als *Outcome* (langfristiger Nutzen für die Organisation) einbezogen, zu dem z. B. Rentabilität, Kundenzufriedenheit und Produktivität gehören (Bushnell, 1990; Reio et al., 2017). Mit dem IPO-Modell soll vorrangig festgestellt werden, inwieweit die Ziele des Programms auf individueller und organisationaler Ebene erreicht werden und welche Optimierungen erforderlich sind (Bushnell, 1990). Das Modell liefert nach Bomberger (2003) formative und summative Informationen, die über Kirkpatricks Modell hinausgehen, da es versucht, den Wert der Ausbildung in finanzieller Hinsicht festzustellen.

2.2.4.1.4 CIPP-Modell: Stufflebeam (1971)

Das CIPP-Modell besteht aus den vier Komponenten Context-, Input-, Process- und Product-Evaluation und konzentriert sich auf die kontinuierliche Verbesserung des Programms, die Rechenschaftspflicht, die Kosten und den Programmbedarf. Ziele und Vorgaben werden als *Context* verstanden, *Input* bezieht sich auf die Programmstrukturierung und -gestaltung, *Process* konzentriert sich auf die Umsetzung des Programms und das *Product* bezieht sich auf die Ergebnisse des Programms (Stufflebeam, 1971). Im Rahmen der Product-Evaluation erfolgt eine Analyse des Umfangs und aller Arten der Wirkungen der Personalentwicklungsmaßnahmen. Dabei werden die Resultate des durchgeführten Programms im Verhältnis zu den Programmzielen, den Kosten, übergeordneten Werten oder alternativen Programmen bewertet (Döring & Bortz, 2016).

2.2.4.1.5 CIRO-Modell: Warr, Bird und Rackham (1970)

Mit dem CIRO-Modell wird durch P. B. Warr et al. (1970) ein ebenfalls vierstufiges Modell etabliert. Es besteht aus den Stufen Context, Input, Reaction und Outcome. Auf der Stufe *Context* werden Informationen zur aktuellen Situation erfasst, um Programmbedarf und -ziel festzustellen. Dies ist der Context-Komponente des CIPP-Modells sehr ähnlich. Für die Evaluierung des *Inputs* werden Informationen über mögliche Programmressourcen genutzt. Diese Stufe ähnelt der Input-Komponente des CIPP-Modells. Die Evaluation der *Reaction* bewertet die Reaktion der Teilnehmer, ähnlich wie bei Kirkpatrick's Ebene 1 (reaction, Teilnehmerzufriedenheit). Mit der Evaluation des *Outcome* werden Informationen über die Ergebnisse des Programms beschafft, die sich in unmittelbare, mittlere und endgültige Ergebnisse kategorisieren lassen. Diese Stufe ähnelt Kirkpatrick's Ebenen 2, 3 und 4 (learning, behaviour, results). Der Schwerpunkt dieses Modells liegt eher auf den ersten drei Stufen Context, Input und Reaction als auf der Bewertung der Ergebnisse (P. B. Warr et al., 1970).

2.2.4.1.6 Fünf-Level-Modell: Kaufman & Keller (1994)

Kaufman und Keller (1994) erweitern das Vier-Ebenen-Bewertungsmodell von Kirkpatrick auf insgesamt fünf Level, zudem wird die erste Ebene von Kirkpatrick ergänzt um Befähigung. Die Level 2, 3 und 4 werden mit Akquisition, Anwendung und organisatorische Ergebnisse bezeichnet, Level 5 wird hinzugefügt, um eine Messung der gesellschaftlichen Folgen eines Programms zu ermöglichen. Das erweiterte Modell ermöglicht damit eine Berücksichtigung des Wertes und des Nutzens des Programms sowie der potentiellen gesellschaftlichen Folgen (Kaufman & Keller, 1994; Kaufman, R., Keller, J., Watkins, R., 1996; Reio et al., 2017 nach Bledsoe, 2000).

2.2.4.1.7 ROI-Theorie und -Modell: Phillips (1998)

Auch das ROI-Theorie und -Modell mit ebenfalls fünf Levels von Phillips (1998) basiert auf Kirkpatrick's Modell. Die ersten drei Bewertungsebenen von Kirkpatrick werden von

Phillips übernommen, die vierte Ebene wird erweitert um die Qualität der Leistung. Zusätzlich installiert Phillips ein fünftes Level, mit der der organisatorische Nutzen des Programms durch Konvertierung in monetäre Einheiten bestimmt werden kann (Phillips, 1998). Die Ergebnisse des Programms werden in Geldwerte umgerechnet und mit den Kosten der Ausbildung verglichen, um die tatsächliche Rendite der Investitionen in das Programm (ROI) zu ermitteln (Phillips, 1999). Phillips und Schirmer (2008) bezeichnen dieses Level als die „aussagekräftigste Stufe der Evaluation“ (S. 28), da mit der ROI-Analyse eine gleichwertige Beurteilung aller Investitionen des Unternehmens – auch die in Personalentwicklungsmaßnahmen – stattfinden kann (Phillips & Phillips, 1994-2001).

2.2.4.1.8 VOI: Kellner (2006)

Das fünfstufige Modell von Phillips wird von Kellner (2006) um eine weitere Stufe ergänzt, den VOI (Value of Investment). Die sechste Stufe des Modells dient der qualitativen Auswertung der Personalentwicklungsmaßnahme (Messbarkeit der mittel- und langfristigen Ergebnisse). Über das Instrument *Value Finder* wird der Nutzen des gesamten Programms partizipativ von den Stakeholdern des Nutzens beschrieben. Der qualitative Nutzen einer Personalentwicklungsmaßnahme wird als Kriterium für den langfristigen Erfolg einer Bildungsmaßnahme gegenüber dem monetären Nutzen priorisiert. Darüber hinaus werden auf Stufe 6 auch das Betriebsklima, die Unternehmenskultur sowie die Kommunikation der Ziele und Werte des Unternehmens bewertet (Birgmayr, 2011). Sie schließt damit – über den ROI hinaus – „eine kritische Analyse der qualitativen Wertschöpfung ein und vermittelt dadurch das gesamte Spektrum der Auswirkungen der Weiterbildungsmaßnahmen“ (Kellner, 2006, S. 12).

2.2.4.2 Einordnung der Modelle im Hinblick auf das Ziel der Arbeit

Die Wirksamkeit von Personalentwicklungsmaßnahmen wird an allgemeinen organisationalen Leistungskriterien wie Produktivitätskennzahlen oder Bilanzkennzahlen festgemacht (Höft, 2001). Das Modell von Kirkpatrick bildet einen logischen Rahmen, um Ergebnisse und Auswirkungen von Personalentwicklungsmaßnahmen aus der Sicht des Individuums und einer Organisation zu analysieren (Setaro, 2001). Es bietet allerdings nur einige allgemeine Richtlinien statt spezifischer Techniken zur quantitativen Bewertung von Personalentwicklungsmaßnahmen, selbst auf der vierten Ebene *Results* (Kirkpatrick 1959_1, 1959_2, 1960_1, 1960_2, 1967, 1998). Kirkpatrick's Modell wird von vielen Forschern und Praktikern im Bereich der Evaluierung von Personalentwicklungsmaßnahmen kritisiert. Alliger und Janak (1989) finden das Modell von Kirkpatrick irreführend, da der Eindruck entstehe, dass das Erreichen von Ergebnissen auf höherer Ebene auch Ergebnisse auf niedrigerer Ebene voraussetzt. Holton (1996) kritisiert, dass das Vier-Ebenen-Modell unvollständig sei und die kausalen Beziehungen zwischen den vier Ebenen nicht spezifiziert werden. Spitzer und Conway (2002) weisen auf das Fehlen von Instrumenten hin, die eine Steigerung der Geschäftsergebnisse aufzeigen könnten. Phillips (1999) hält es für unabdingbar, quantitativ zu bewerten, ob sich die Maßnahme gelohnt hat, und erweitert daher das Kirkpatrick-Modell um die fünfte Stufe *Return on*

Investment (ROI). Varianten zum ROI, wie z. B. der Return on Development Investment (RODI) von Cascio et al. (2019) und Avolio et al. (2010), die Effectsize von Meuse et al. (2009), und den weiter gefassten Value of Investment (VOI) von Kellner (2006), ermöglichen umfangreichere betriebswirtschaftliche Analysen. Für diese Indikatoren liegen wenige empirische Daten vor, und wenn, dann überwiegend aus dem amerikanischen Raum (McGovern et al., 2001).

Mehrheitlich erforscht sind die Reaktions-, Lern- und Verhaltensebene des Kirkpatrick-Modells, die Ergebnisebene dagegen wird in der Forschung vernachlässigt (P. B. Warr et al., 1970; Brinkerhoff (1987); Kaufman & Keller, 1994). van Buren und Erskine (2002) stellen fest, dass die Häufigkeit der Messung der Erfolgsmaße in Organisationen für Ebene 1 bei 78 %, für Ebene 2 bei 32 %, für Ebene 3 bei 9 % und für Ebene 4 lediglich bei 7 % liegt. Aktuelle Untersuchungen von 199 Organisationen in den USA zeigen, dass die Anzahl der Evaluationen aller vier Ebenen wächst: Ebene 1 evaluieren demnach 88 %, Ebene 2 83 %, Ebene 3 60 %, Ebene 4 35 % und Ebene 5 (ROI) 15 % (ATD Research, 2016). Alliger et al. (1997) stützen die Ergebnisse von Arthur et al. (2003): In ca. 15 % der Studien wurde auf der Ebene Reaction – Teilnehmerzufriedenheit, in ca. 50 % Studien auf der Ebene Learning – Wissenstransfer und in ca. 35 % auf der Ebene Behavior – Praxistransfer evaluiert, zur Ebene Results – Unternehmenserfolg werden allerdings keine Angaben gemacht (siehe Abbildung 18). Arthur et al. (2003) folgern, dass in der Praxis überwiegend die Evaluation der Ebene 1 aufgrund der relativ einfachen Datengewinnung auf dieser Ebene ermöglicht wird.

Abbildung 18: Häufigkeit von Evaluationen der Ebenen des Kirkpatrick Modells in Praxis und Wissenschaft im Vergleich. Eigene Darstellung nach van Buren und Erskine (2002) und Arthur et al. (2003).

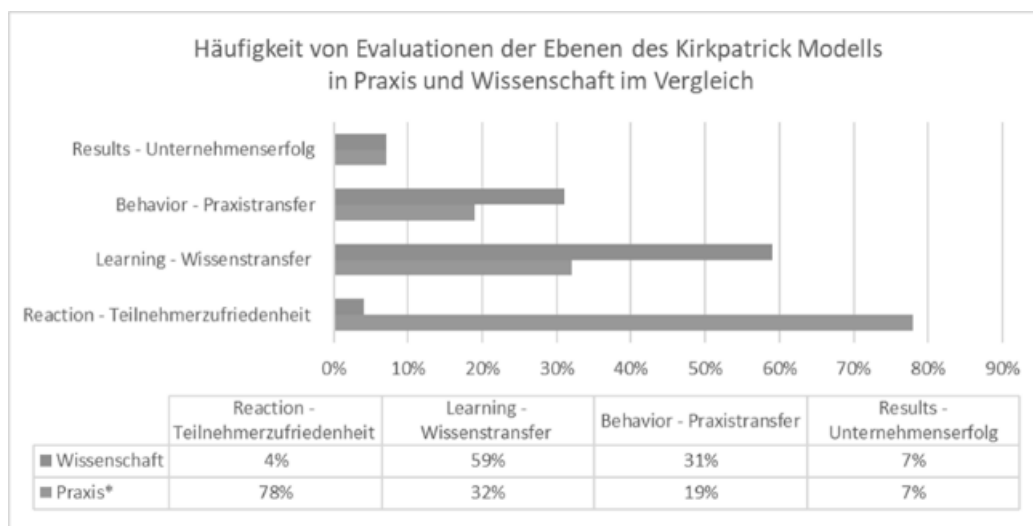


Abbildung 19 zeigt das Vier-Ebenen-Modell von Kirkpatrick im Vergleich zur Abbildung 1 erweitert um die fünfte Ebene ROI durch Phillips mit Angaben zur Häufigkeit der Messung der Erfolgsmaße in Organisationen (Praxis). Phillips und Schirmer (2008) betonen, dass die Messung des ROI häufig auf Selbsteinschätzungen der Stakeholder basiert und damit lediglich eine subjektive Wahrnehmung des finanziellen Nutzens darstellt, statt beständig und umfassend für die Organisation ermittelt worden zu sein.

Die Evaluation von Personalentwicklungsmaßnahmen wird darüber hinaus bezüglich (Ressourcen-)Aufwand und Ertrag kritisch hinterfragt (Süßmair, 2007). In einem Großteil der Evaluationen werden die Ebenen Reaction und Learning über Selbsteinschätzungen der Teilnehmer erhoben, während auf der Ebene Behavior auf Fremdurteile gesetzt wird.

Abbildung 19: Das Vier-Ebenen-Modell von Kirkpatrick, erweitert um die Ebene ROI durch Phillips mit Angaben zur Häufigkeit der Messung der Erfolgsmaße in Organisationen (Praxis). Eigene Darstellung nach Phillips (2003), Kirkpatrick (1967, 1994), van Buren und Erskine (2002) und Döring und Bortz (2016).



Zur Einschätzung der durch das Programm hervorgerufenen Veränderungen des Verhaltens der Programmteilnehmer, werden z. B. andere Trainer, Kollegen aus dem gleichen Unternehmen oder gezielt geschulte Beobachter eingesetzt (Rowold, 2008). Derartige Verfahren erfordern einen hohen zeitlichen und auch personellen Aufwand (Tirre, 2012). Der Autor empfiehlt zur Generierung von validen Ergebnissen auf der Verhaltensebene, bei der Erhebung von Erfolgskriterien auf Fremdurteile zu setzen. Ein praktikables standardisiertes Evaluationskonzept bietet sich hier an, um den Aufwand bei der Erhebung gering zu halten (Tirre, 2012). Aus der Praxis heraus wird die Evaluation von Personalentwicklungsmaßnahmen mittels des Kirkpatrick-Modells als aufwändig

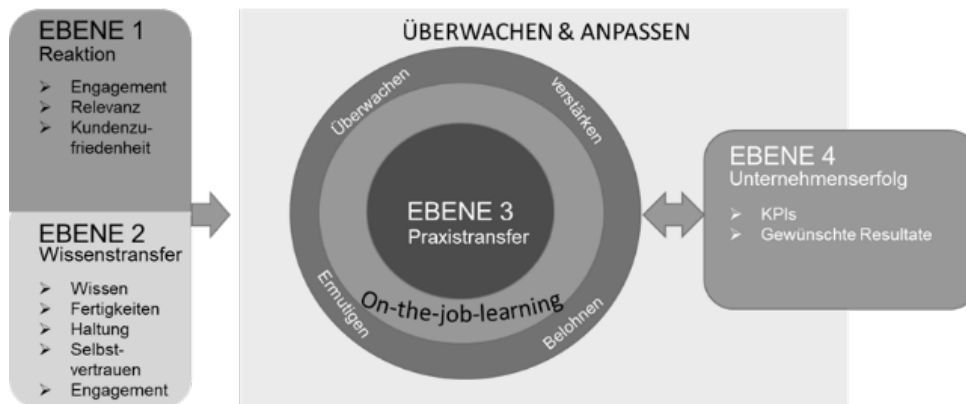
eingeschätzt. Personalentwickler verstehen den Aufwand für Personalentwicklungsmaßnahmen aber als Investition, wobei die Ergebnisse gemessen werden sollten, um Programme zu verbessern, Ausbildungserfolge nachzuweisen und zu zeigen, wie die Maßnahmen der Personalentwicklung dazu beitragen, dass die Unternehmensziele erreicht werden.

Trotz aller Kritik an Kirkpatrick's Modell und der Versuche von Forschern, ihre eigens entwickelten Modelle von dem Vier-Ebenen-Modell abzugrenzen, haben die Mehrzahl der in der Literatur zu findenden Evaluationsmodelle einen starken Bezug zu den Abgrenzungen und Inhalten der vier Ebenen des Modells. Gülpen (1996) führt Plausibilität, die Einfachheit und die Anwendungsfreundlichkeit als Gründe für die Popularität des Kirkpatrick-Modells an. Das Modell wird daher in großem Umfang weltweit zur Evaluation von Personalentwicklungsprogrammen verwendet (Bates, 2004; Kauffeld, 2016). D. L. Kirkpatrick (1998) rät, mit der Evaluation auf Ebene 1 zu beginnen und anschließend mit den anderen Ebenen der Reihe nach fortzufahren. Laut Gülpen (1996) spezifiziert die jeweilige Forschungsfrage das Evaluationsziel und muss bei jeder Evaluation definiert werden. Besteht das Interesse einer Trainingsevaluationsstudie lediglich in einer *Erhebung der Lernleistung* und nicht in der Untersuchung möglicher weiterer Auswirkungen auf der Transfer- und/oder Organisationsebene, sind die Auswertungen auf der Ebene *Transfer Organisation* nicht brauchbar. Sie tragen nicht weiter zur Klärung des Ziels der Studie bei.

Kirkpatrick nimmt an, dass die Evaluationsebenen nacheinander qualitativ immer hochwertigere Informationen liefern. Die vier Ebenen bilden eine Art Evaluationspyramide und die vierte Ebene stellt dabei die höchste und wertvollste Stufe dar. Ziel des Evaluationsprozesses ist es demnach, ein Training so „hoch“ wie möglich, d. h. auf der möglichst höchsten Ebene zu evaluieren. Als hierarchisch höchste Evaluationsebene sollte daher die Ebene gewählt werden, die das in einer Bedarfsanalyse spezifizierte Ziel der Personalentwicklungsmaßnahme messen kann.

J. D. Kirkpatrick und Kirkpatrick W. (2016) haben auf Basis der grundlegenden Arbeiten Kirkpatrick's das Vier-Ebenen-Modell zum *New World Kirkpatrick Model* entwickelt (siehe Abbildung 20). Das New World Kirkpatrick Model reflektiert die Modernisierungen der Arbeitswelt und die Veränderungen in der Bildungslandschaft. Darüber hinaus soll das neue Modell dazu veranlassen, dass Evaluationen nicht mehr überwiegend auf Ebene 1 und 2 stattfinden, sondern auch mit dem Vorurteil aufgeräumt wird, dass Evaluationen auf Ebene 3 und 4 zu aufwändig oder zu schwierig seien (J. D. Kirkpatrick & Kirkpatrick W., 2016). Der Fokus liegt auf der Ebene 4, dem Ergebnis des Personalentwicklungsprogramms, das eine verbesserte Leistung und/oder Qualität der Teilnehmer repräsentiert. Unter der Annahme, dass jede Organisation nur *ein* Ebene-4-Resultat („the true target of organizational results“, S. 13) aufweist, das langanhaltend und langfristig ist, wird darauf hingewiesen, dass es herausfordernd ist, dieses eine Resultat zu identifizieren und zu operationalisieren. Das Kriterium zur Messung des Ergebnisses sollte genau das beinhalten, was die Organisation zur Wirtschaft oder zur Gesellschaft beiträgt (J. D. Kirkpatrick & Kirkpatrick W., 2016).

Abbildung 20: The New World Kirkpatrick Model. Eigene Darstellung nach J. D. Kirkpatrick & Kirkpatrick W., 2016, S. 11.



Die Ebenen *Reaction* und *Learning* würden als eindimensionale Konstrukte betrachtet, obwohl mehrere Studien die Mehrdimensionalität sowohl der Reaction-Ebene (Alliger et al., 1997; Morgan & Casper, 2000; P. Warr et al., 1999; P. Warr & Bunce, 1995) als auch der Learning-Ebene (Alliger et al., 1997; Kraiger et al., 1993) nachweisen konnten (Kraiger, 2003). Darüber hinaus fehlten Handlungsempfehlungen, wie die Ergebnisse der Evaluation mit einer strategischen Entscheidungsfindung über Personalentwicklungsmaßnahmen verbunden werden können (Kraiger & Jung, 1997). Delahaye (2000) wiederum kritisiert das Modell von Kirkpatrick als „outputorientiert“ (S. 339) und dass es sogenannte „Vorhersagefaktoren“ (S. 354) nicht berücksichtigt. Diese seien aber erforderlich, um ein bestimmtes Ergebnis vorhersagen und das Programm optimieren zu können. Rae (2002) fordert eine Stufe 0, mit der der Evaluationsprozess beginnt, bevor die Personalentwicklungsmaßnahme startet.

Den in diesem Kapitel vorgestellten Evaluationsmodellen ist gemein, dass das Kirkpatrick-Modell als Basis für die Entwicklung dieser dient. Der Schwerpunkt der Kritik an dem Modell liegt auf der mangelnden Evaluation der *Results*, also der outcome-orientierten Messung der Wirksamkeit des Personalentwicklungsprogramms. Der größte Beitrag des Kirkpatrick-Modells liegt darin, dass es ein einheitliches Vokabular für Evaluationen durch Praktiker in Personalabteilungen geschaffen hat (G. G. Wang et al., 2002). Personelle Ressourcen, Konkurrenzsituation sowie die Marktgegebenheiten des Unternehmens erfordern aber eine individuelle Konzeption bezüglich des Evaluationsprozesses und Methoden und Techniken, die auf die Ziele abgestimmt sind (Nork, 1991). Diese Aussage von Nork (1991) führt zu dem in dieser Arbeit entwickelten innovativen Evaluationskonzept. Es ist eine individuelle Konzeption unter Berücksichtigung der personellen Ressourcen in den Vereinen und in den Fußballverbänden, der Konkurrenzsituation unter den Nachwuchsleistungszentren sowie den besonderen Gegebenheiten der Fußballbranche und den dort vorliegenden Datenbeständen. Die Konzeption und die Anwendung des Evaluationskonzepts in dieser Arbeit folgt im Prinzip den beschriebenen Modellen und bezieht sich wegen der auf den betriebswirtschaftlichen Erfolg fokussierten Fragestellung auf das ROI-Theorie und -Modell von Phillips (1998).

Eine qualitative Auswertung der Personalentwicklungsmaßnahmen, wie im VOI-Modell von Kellner (2006) gefordert, erfolgt nicht.

2.2.5 Evaluation der Resultate von Personalentwicklungsmaßnahmen

Für Werner und DeSimone (2005), D. L. Kirkpatrick (1998) und Phillips (1996) ist die Ebene 4 – die Result-Ebene – die wichtigste, aber auch die am schwierigsten zu bewertende Ebene. Auf dieser Ebene versuchen Organisationen, die tatsächlichen organisatorischen Veränderungen aufgrund von Personalentwicklungsmaßnahmen zu messen und diesen Veränderungen einen monetären und/oder numerischen Wert beizumessen. Es gibt nur wenige Studien, die die Auswirkungen von Personalentwicklungsmaßnahmen anhand einer organisationalen Ergebnisgröße, wie z. B. Produktivität oder hervorbrachte Qualität, direkt messen. Im Bereich der Management- und Organisationsforschung werden aber zunehmend Resultate von Personalentwicklungsmaßnahmen im Rahmen eines ganzheitlichen Human Resource Managements untersucht (Alagaraja, 2013; Sonntag & Stegmaier, 2016).

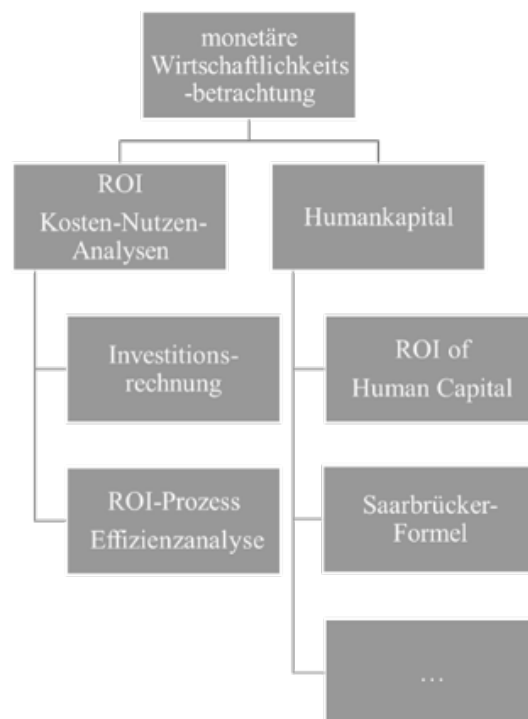
2.2.5.1 Messen des Ergebnisses von Personalentwicklungsmaßnahmen

Personalentwicklungsmaßnahmen stellen eine Investition in Humankapital dar, daher wird häufig die *Humankapitaltheorie* als erklärungsrelevante Grundlage zum besseren Verständnis der Messung des Ergebnisses von Personalentwicklungsmaßnahmen herangezogen. Ausbildung und Personalentwicklungsmaßnahmen stellen aus wirtschaftswissenschaftlicher Sicht die Bildung bzw. Produktion von Humankapital dar (Wallebohr, 2019). Laut Jaeger (2004) umfasst das Humankapital das gesamte personenbezogene Wissen und alle Fähigkeiten und Fertigkeiten des Individuums. Die *Humankapitaltheorie* besagt, dass durch Investitionen in die Personalentwicklung und durch den Einsatz von Humankapital in der Wirtschaft, die Produktivität des Individuums und damit des gesamten Unternehmens gesteigert wird (G. S. Becker, 1993). Mitarbeiter mit mehr Kenntnissen oder Fähigkeiten weisen ein höheres Humankapital aus, sind produktiver und erwirtschaften in der Regel mehr für das Unternehmen (Diebolt et al., 2017). Investitionen in die Personalentwicklungsmaßnahmen werden unterteilt in direkte Kosten für Weiterbildungsmaßnahmen und indirekte Kosten für den zeitlichen Aufwand an den Personalentwicklungsmaßnahmen. Dabei folgt diese Investitionsentscheidung den gleichen Prinzipien (z. B. Kosten-Nutzen-Vergleich) wie alle anderen Investitionsentscheidungen in einem Unternehmen (Phillips & Schirmer, 2008). Es wird betriebswirtschaftlich beurteilt, ob sich die Investition in den Bereich gelohnt hat. Im Fokus von *Human Capital Management* steht der Wert des Personals insgesamt statt eines Wertes des Individuums. Bewertet werden insbesondere immaterielle Vermögenswerte, wie „geistiges Eigentum, Patente, Lizenzen, Copyrights, Kundenstamm, Lieferantenbeziehungen, Nutzungseffizienz der IT-Infrastruktur, Image und daraus resultierender Wert von Marken sowie vor allem der Wert der Mitarbeiter, das Human Capital“ (Scholz et al., 2004, S. 16). In der *Produktionstheorie* ist Gegenstand der Diskussionen das – bei gegebener Produktionstechnologie – gewinnmaximierende Unternehmen (Produzent). Zwischen der

Menge der eingesetzten Produktionsfaktoren (Input) und der Produktmenge (Output) besteht ein Zusammenhang. Die *Kostentheorie* wiederum beschreibt die Zusammenhänge zwischen den Produktionskosten und dem produzierten Output (zu den Grundlagen der Produktions- und Kostentheorie siehe Fandel, 2005; Schweitzer & Küpper, 2013). Fragestellungen zur Effektivität und Effizienz von Produktionsverfahren können im Allgemeinen anhand der genannten Theorien beantwortet werden (Wallebohr, 2019). Das Ausmaß, in dem eine Organisation ihre Ziele erreicht, stellt die Effektivität dar (Output orientiert an der Fragestellung: Tun wir die richtigen Dinge?). Die Effizienz hingegen berücksichtigt auch die Menge an Ressourcen, die für die Produktion des gewünschten Outputs eingesetzt werden (Input-/Output-Relation). Ziel ist es, die Produktionstechnologie bzw. das Ausbildungsprogramm auszuwählen, das einen hohen Output mit minimalen Kosten erzeugt (Minimalkostenkombination; Fandel, 2005). Die Produktionstheorie sieht für die Messung der Effizienz Produktionsfunktionen vor. Für die vorliegende Arbeit sind diese nicht zielführend, sodass eher die grundsätzlichen Zusammenhänge zwischen Input- und Outcome sowie die Effizienz und Effektivität beschrieben werden.

Monetäre Wirtschaftlichkeitsbetrachtungen von Personalentwicklungsmaßnahmen werden einerseits über Kosten-Nutzen-Analysen und andererseits über Humankapitalbewertungen durchgeführt (siehe Abbildung 21) Die Ansätze zur Evaluation von Personalentwicklungsmaßnahmen und ausgewählte Evaluationskriterien werden im nächsten Kapitel näher ausgeführt.

Abbildung 21: Ansätze zur Evaluation von Personalentwicklungsmaßnahmen. Eigene Darstellung nach J. Müller (2017) und J. Littkemann et al. (2012).



2.2.5.2 Evaluationskriterien für Personalentwicklungsmaßnahmen

Verschiedene Modelle zur Ermittlung eines *Human-Capital-Wertes* für die gesamte Belegschaft eines Unternehmens zeigen Scholz et al. (2011). Die Autoren kategorisieren insgesamt 47 unterschiedliche Modelle in Marktwertorientierte Ansätze – spekulativ (z.B. Markt-/Buchwert-Differenz, Human Capital Market Value); Accounting-orientierte Ansätze – kalkulativ (z. B. Human Resource Accounting, Lernzeitbasierte Wissensbilanz); Indikatorenbasierte Ansätze – kumulativ (z. B. HR Scorecard, Human Capital Value, CIPD Framework); Value Added-Ansätze – derivativ (z. B. Market Value Added, Knowledge Capital); Renditeorientierte Ansätze – relativ (z. B. Human Capital Pricing Model, ROI of Human Capital) und Ertragspotenzialorientierte Ansätze – integrativ (z. B. Value Creation Index, Saarbrücker Formel). Diese Ansätze wurden hinsichtlich ihrer Eignung zur Erfüllung der von den Autoren aufgestellten neun Postulate zu den Gestaltungsgrundsätzen eines Modells bewertet und führen anhand angestrebter Ergebnisse zu einer Entscheidungsmatrix, die die für ein Unternehmen geeignete individuelle Human-Capital-Bewertungsmethode zu identifizieren hilft (siehe Abbildung 22).

Personalentwicklungsmaßnahmen, die auf erhöhte Verkäufe, reduzierte Unfälle, geringere Fluktuation, gesunkene Kosten oder gesteigerte Produktion abzielen, können in der Regel bewertet werden (D. L. Kirkpatrick & Kirkpatrick, 2006). Geschäftliche Erfolgskriterien kategorisieren Phillips und Schirmer (2008) hinsichtlich der Bewertung in harte und weiche Kriterien. Zu den harten Kriterien zählen sie Produktivitätskennzahlen (produzierte Quantität, verkaufte Einheiten usw.) und Qualitätsindikatoren (Fehlerhäufigkeit, Ausschuss usw.), die beide objektiv messbar und in monetäre Werte konvertierbar sind. Diese harten Kriterien sind in der Regel in den Unternehmen leicht verfügbar, dem Topmanagement vertraut und können in Analysen einfach eingesetzt werden (Solga, 2011; Phillips & Schirmer, 2008). Weiche Kriterien hingegen, wie z. B. Arbeits- und Organisationsklima sowie Kundenzufriedenheit, lassen sich eher über psychologische Messmethoden objektivieren. Eine Monetarisierung ist schwierig. Um eine hohe Glaubwürdigkeit bzw. große Akzeptanz einer Umwandlung der Einschätzung von weichen Daten in einen monetären Wert im Unternehmen zu erreichen und die Genauigkeit der Daten zu demonstrieren, sind einheitliche Prozesse und Regeln erforderlich (Phillips & Schirmer, 2008). Auswirkungen von Personalentwicklungsmaßnahmen werden zunehmend an Indikatoren des Unternehmenserfolgs gemessen, die den Entscheidern im Unternehmen vertraut sind (Solga, 2011). Die Resultate werden nicht direkt am Ende der Personalentwicklungsmaßnahmen evaluiert, sondern nachdem die Teilnehmer über längere Zeit Gelegenheit hatten, die erworbenen Kompetenzen umzusetzen (Solga, 2011).

Eine Investition enthält eine klare Erwartungshaltung an ihren Return. In den Unternehmen wird weltweit mehr Kostenverantwortlichkeit von mittlerweile allen Bereichen erwartet. Davon ist mittlerweile auch die Personalentwicklung betroffen (Phillips & Schirmer, 2008). Das Topmanagement von Unternehmen fordert Nachweise über den

Abbildung 22: Human-Capital Bewertung-Entscheidungsmatrix. Eigene Darstellung nach Scholz et al. (2011)

	? Methode intransparent	i (ausschließlich unternehmensintern durchführbar)	e (zusätzlich unternehmensextern durchführbar)
€ Human Capital Ergebnis als Geldgröße	Accounting for the future ICM Model Skandia Navigator Total Value Creation Value Explorer	Calculated Intangible Value Economic Value Added Entgeltbarwert-Ansatz Human Capital Value Added Human Economic Value Added Human Resource Accounting Investor-Assigned Market Value Knowledge Capital Scoreboard Knowledge Capital Scoreboard Workonomics	Human Capital Market Value Market Value Added Markt-/Buchwertdifferenz Marktwert/Mitarbeiter Quotient Saarbrücker Formel
Z Human Capital Ergebnis als Zahl	CIPD Framework Competence x Commitment Employee-Value-Index HR Scorecard Human Asset Worth Human Capital Index Human-Capital-Indikator Human Capital Pricing Model Human-Capital-Value Human-Potential-Index Humatics Intellectual Capital Audit Intellectual Capital Index Intellectual Capital Navigator Kennzahlenbasierte Wissensbilanz Lernzeitbasierte Wissensbilanz Personalwert Summenmodell des Humankapitals Value Creation Index	Intangible Assets Monitor ROI of Human Capital Tobin's q Weiterentwickelter Value Added Intellectual Coefficient	Markt-/Buchwert-Relation Originärer Value Added Intellectual Coefficient
V Human Capital Ergebnis verbal	Aries IC-Rating		

Erfolg der Aufwendungen für die Personalentwicklungen. Bildungscontrolling, Messung von immateriellen Werten, ROI für Personalentwicklungsmaßnahmen oder Humankapitalmessung sind nur einige der Instrumente, mit denen – wie mit Berechnungsmodellen im Produktionsbereich – die Resultate der Personalentwicklungsprogramme

bewertet werden (Kühl, 2006). Personalentwicklung wird nicht mehr nur als Kostenfaktor betrachtet, sondern als eine gleichrangige Investition in die strategische Zukunft des Unternehmens gesehen. Wie bei anderen Investitionen im Unternehmen auch ist der Return on Investment (ROI) die Kennzahl, die den monetären Nutzen der Personalentwicklungsmaßnahmen ins Verhältnis zu den Kosten setzt und damit den Erlös des eingesetzten Kapitals bestimmt (Phillips & Schirmer, 2008). Um stetiges Wachstum zu generieren, sind Unternehmen gezwungen, die ihnen für die Personalentwicklungsmaßnahmen zur Verfügung stehenden Ressourcen effektiv einzusetzen (Grossman & Salas, 2011). Insbesondere bei Ressourcenknappheit sind der *added value* oder der *chain of impact* von Qualifizierungsmaßnahmen nachzuweisen und ggf. zu optimieren (Schermuly et al., 2012). Die Evaluation des Erfolgs von Personalentwicklungsmaßnahmen ist ein wesentliches Instrument, um zu belegen, dass die Personalentwicklungsmaßnahmen eine notwendige Maßnahme zur Sicherung der Wettbewerbsfähigkeit und Sicherung des Unternehmenserfolgs ist (Kauffeld, 2016). Schwierigkeiten bei der Analyse geschäftlicher Resultate bestehen darin, dass die vom Topmanagement bevorzugten harten betriebswirtschaftlichen Kennzahlen von Einflüssen abhängig sind, die sich nicht durch die Teilnehmer der Personalentwicklungsmaßnahmen kontrollieren lassen und infolgedessen auch nicht durch Personalentwicklungsmaßnahmen oder -Programme veränderbar sind (Kriterienkontamination; Goldstein & Ford, 2002). „Ein Evaluationskriterium gilt als kontaminiert, wenn es Aspekte beinhaltet oder von Einflüssen abhängig ist, die sich der Veränderbarkeit durch das Personalentwicklungsmaßnahmen-Programm entziehen.“ (Solga, 2011, S. 384). Rowold und Mönninghoff (2005) zeigen, wie Personalentwicklungsmaßnahmen effektiv mittels einer Kosten-Nutzen-Analyse (unter Nutzung der HC Bridge™ nach Boudreau & Ramstad, 2002) evaluiert werden können, sodass die Ergebnisse vom Management der Organisationen anerkannt werden. Cascio et al. (2019) empfehlen ebenfalls den Einsatz von Modellen und Techniken zur Identifizierung und Bewertung von Humankapitalinvestitionen, da Kosten-Nutzen- und Kosten-Effektivitäts-Analysen nützliche Instrumente seien, um den Nutzen im Verhältnis zu den Kosten von Programmen oder Investitionen zu bewerten. Die Kosten-Nutzen-Analyse ermöglicht einen Vergleich der absoluten Werte der Erträge aus unterschiedlichen Programmen oder Organisationen (Cascio et al., 2019). Die Ermittlung des ROI für Personalentwicklungsmaßnahmen wurde als wichtigster Trend im Personalwesen bezeichnet (ASTD, 2002; Schermuly et al., 2012; Cascio et al., 2019). Im Folgenden werden Grundlagen zur Ermittlung des ROI, der ROI-Prozess und in Kapitel 2.2.5.2 die Ermittlung des ROI für Personalentwicklungsmaßnahmen erläutert.

Cascio et al. (2019) beschreiben den ROI als Wert der Rendite zu einem bestimmten Zeitpunkt in der Zukunft (angemessen abgezinst, um das Risiko und die Inflation widerzuspiegeln), geteilt durch die für die Investition erforderlichen Mittel (ebd., S. 54). Der ROI gibt den monetären Ertrag jeder einzelnen investierten Geldeinheit an (Kauffeld, 2016). Mit dem ROI kann die Ertragskraft eines Unternehmens unabhängig von der Finanzierungsstruktur (Eigen- oder Fremdkapital) ermittelt werden (Gladen, 2014). Ziel des ROI ist die Steuerung der knappen Ressource *Kapital* im Hinblick auf die

Ergiebigkeit des Kapitals je Kapitaleinheit. Die Ermittlung des ROI wird in Unternehmen häufig zur Beurteilung von Investitionen angewendet, um die Rentabilität für bestimmte Investitionsobjekte festzustellen und deren Risiko zu minimieren (Gabler Wirtschaftslexikon, 2022). Der ROI ist unter den *Statischen Investitionsrechenverfahren* einzuordnen. Es erfolgt eine „Abstraktion von der zeitlichen Struktur des durch das Investitionsprojekt generierten Zahlungsstroms. Stattdessen ermitteln sie Erfolgsgrößen für ein exemplarisches Durchschnittsjahr“ (Gabler Wirtschaftslexikon, 2022). Dabei wird auf Prinzipien der Kostenrechnung sowie der Erfolgsrechnung zurückgegriffen. Bei den statischen Investitionsrechenverfahren handelt es sich um einperiodige Modelle. Sie besitzen hohe praktische Relevanz, da sie vergleichsweise einfach in der Anwendung sind und einen relativ geringen Informationsbedarf haben (Gabler Wirtschaftslexikon, 2022). Die zeitliche Struktur der durch das Investitionsprojekt generierten Ein- und Auszahlungen wird bei *Dynamischen Investitionsrechenverfahren* in die Bewertung einbezogen (Mehrperiodenmodelle). „Dynamische Rechenverfahren berücksichtigen den Umstand, dass gleich hohe Zahlungen zu verschiedenen Zeitpunkten unterschiedlich bewertet werden“ (Gabler Wirtschaftslexikon, 2022).

Der *ROI of Human Capital (HCROI)*, wie von Fitz-enz (2009) vorgestellt, wird von Scholz et al. (2011) in der Entscheidungsmatrix den Kategorien *Z* (Human Capital zusätzlich als Zahl) und *i* (unternehmensintern durchführbar) zugeordnet. Der HCROI als nicht-monetäre Kennzahl basiert auf finanzwirtschaftlichen Inputs und repräsentiert den Wert für das Gesamtunternehmen. Die Daten werden unternehmensintern generiert und eventuell gemäß der gebotenen Berichtspflichten extern publiziert (Scholz et al., 2011). Der Berechnung des HCROI erfolgt auf Basis des Umsatzertrags abzüglich der Differenz von Aufwand und Löhnen/Gehältern in Relation zu Löhnen/Gehältern.

Die *Saarbrücker Formel* stellt einen ertragspotentialorientierten integrativen Ansatz dar, mit dem eine Bewertung des Human Capital als Geldgröße unter Berücksichtigung (zusätzlicher) externer Inputgrößen vorgenommen wird (Scholz et al., 2011). Sie berücksichtigt marktliche Referenzwerte der im Unternehmen vorhandenen Belegschaft (HC-Wertbasis), den Wert, der im Zeitablauf verloren geht (HC-Wertverlust), den Wert, der durch Personalentwicklungsmaßnahmen kompensiert wird (HC-Wertkompensation) und die HC-Wertveränderung, eine Wertmehrung oder -minderung durch Motivation (Scholz et al., 2004).

Ziel der Arbeit ist u. a. die Entwicklung eines Evaluationskonzepts, das mit *geringem Aufwand leicht anzuwenden ist*. Durch Komplexitätsreduktion sollen Nachvollziehbarkeit und Transparenz der zu entwickelnden Indikatoren sichergestellt werden (Scholz et al., 2004). Der ROI ist eine Methodik, die sich in der Praxis unter verschiedenen Rahmenbedingungen bewährt hat. Mit dem ROI werden genau die Daten ermittelt, mit denen Manager und Entscheider hinsichtlich der Bewertung ihrer Investitionen vertraut sind. Auch die Implementierung des Prozesses und der Prozess selbst gilt als nicht ressourcenintensiv (Phillips & Schirmer, 2008). Empirische Studien weisen nach, dass der ROI eine große Akzeptanz erfährt und als Kennzahl für den strategischen Erfolg für aussagekräftig gilt (Hansmann & Kehl, 2000).

Für die vorliegende Evaluation wird daher der ROI für Personalentwicklungsmaßnahmen als Messinstrument des monetären Erfolgs der Personalentwicklungsmaßnahmen in Nachwuchsleistungszentren eingesetzt und im nächsten Kapitel näher erläutert.

2.2.5.3 ROI für Personalentwicklungsmaßnahmen

Ansätze zur ROI-Messung in der Personalentwicklung basieren auf einem 1919 entwickelten Buchhaltungsmodell der Firma DuPont zur dezentralen Finanzkontrolle (Koontz et al. (1984). „Der ROI für Personalentwicklungsmaßnahmen kann interpretiert werden als eine Schätzung der Anzahl an Geldeinheiten, die man für eine in eine Trainingsmaßnahme investierte Geldeinheit zurückbekommt oder als Prozentsatz der Investition, den man zurückerhalten wird“ (Kauffeld, 2016, S. 127). Nach Bott et al. (2003) kann die allgemeine Formel des ROI für die Ermittlung des ROI für Personalentwicklungsmaßnahmen folgendermaßen übersetzt werden:

Formel 1: ROI für Personalentwicklungsmaßnahmen

$$ROI (\%) = \frac{\text{Nettonutzen der Personalentwicklungsmaßnahmen}}{\text{Kosten der Personalentwicklungsmaßnahmen}} \times 100$$

Der Nettonutzen ist die Differenz zwischen der Steigerung des Ergebnisses gemessen an dem ausgewählten Evaluationskriterium (z. B. Erlöse) und den Kosten der Personalentwicklungsmaßnahmen. Um den ROI zu berechnen, werden zunächst direkte und indirekte Kosten für Materialien, Personal, Trainer, Räumlichkeiten und die Durchführung sowie Opportunitätskosten aufsummiert. Anschließend erfolgt die Identifizierung eines für die Messung des Programmnutzens geeigneten Kriteriums. Die Veränderung dieses Kriteriums wird monetär bewertet. Die Auswertung des ROI erfolgt unter der Annahme, dass alle Parameter konstant bleiben und der errechnete Wert im monokausalen Zusammenhang zu der Personalentwicklungsmaßnahme steht. Diese Bewertung ist häufig mit Schwierigkeiten verbunden, da zu berücksichtigen ist, dass die Veränderungen auch andere Ursachen als die Personalentwicklungsmaßnahme haben können und die Bewertung von weichen Kriterien – wie Arbeitszufriedenheit, Arbeitnehmer-Commitment – in Geldeinheiten oft stark konstruiert werden muss. In einem weiteren Schritt wird der (Netto-)Nutzen der Personalentwicklungsmaßnahmen anhand der Differenz der Kosten von der Veränderung subtrahiert. Abschließend erfolgt die Division des Nettonutzens durch die Kosten der Personalentwicklungsmaßnahme (Hagen & Wenning, 2007).

Der ROI für Personalentwicklungsmaßnahmen ist als ergebnisorientierte Kennzahl zur Messung der Effektivität insbesondere dann das Instrument der Wahl, wenn die Resultate der Personalentwicklungsmaßnahmen ohne großen Aufwand in monetären Einheiten

ten ausgedrückt werden sollen (Kauffeld, 2016). Bedenken gegenüber dem ROI für Personalentwicklungsmaßnahmen werden von Personalentwicklern, Entscheidungsträgern und Wissenschaftlern geäußert. Personalentwickler empfinden den Prozess zur Ermittlung des ROI als zu aufwändig bei der Datenerfassung und -analyse und als zu ressourcenbindend (Personal und Kosten). Entscheidungsträger von Unternehmen haben aber ein starkes Interesse daran, die Investitionen in Personalentwicklungsmaßnahmen mit der gleichen Kennzahl zu bewerten wie die übrigen Investitionen des Unternehmens, erwarten aber ebenfalls einen kostengünstigen, einfachen Prozess. Wissenschaftler fordern ein Vorgehen, das auf der Basis von Modellen, Formeln, Annahmen und Theorien genaue Werte und beständige Ergebnisse ermittelt.

Viele Daten in der Praxis beruhen auf Schätzungen, sodass die Gültigkeit der ROI-Berechnung infrage gestellt wird. Sind Schätzungen für Inputgrößen erforderlich, sind sie gemäß ihres Risikos zu korrigieren und anzupassen (Phillips & Schirmer, 2008). Kauffeld (2016) führt einige der umstrittenen Punkte an: Häufig lägen keine Leistungskennwerte für die Teilnehmer oder die Gesamtorganisation vor, der Kausalzusammenhang zwischen den Personalentwicklungsmaßnahmen und der Veränderung des Erfolgskriteriums ist fraglich, die Anwendung der erlernten Fertigkeiten und Fähigkeiten sei von vielen anderen Faktoren abhängig, häufig sei die Evaluation zu zeitnah nach Abschluss der Personalentwicklungsmaßnahme, die Kosten seien nicht für jede Organisation gleich, auch wenn z. B. das gleiche Programm durchgeführt werde, Effekte variabler Kosten und Steuern sowie die Inflation würden nicht berücksichtigt, es werde häufig nur eine Teilnehmergruppe evaluiert und die Fluktuation von Mitarbeitern werde bei der Ermittlung des ROI außer Acht gelassen.

Auch wenn oft eine genaue Zuschreibung des Erfolgs zu den Personalentwicklungsmaßnahmen nicht ermittelt werden kann, schaffen ROI-Messungen einen Mehrwert, indem sie Objektivität, Eindeutigkeit, Transparenz und Vergleichbarkeit sicherstellen (Bergmann & Meurer, 2003; Cascio et al., 2019; . Kauffeld, 2016). Messungen der Rentabilität von Investitionen in Menschen sind „zentral für den strategischen Erfolg praktisch aller Organisationen“ (Cascio et al., 2019, S. 1).

2.3 Erfolgsfaktorenforschung

Erfolg ist ein allgegenwärtiger Begriff in unserer Gesellschaft, insbesondere im wirtschaftlichen Bereich. Für jedes Unternehmen hat Erfolg jedoch eine andere Bedeutung und verschiedene Definitionen. Das können sein: Erhöhung des Umsatzes, des Absatzes, des Gewinns, des Aktienkurses, der Kundenwahrnehmung, der Mitarbeiterzufriedenheit oder das Verhindern von Wettbewerbsnachteilen. Das Ergebnis von Aktivitäten ist Erfolg (positiv) oder Misserfolg (negativ) und tritt dann ein, wenn es mit einem zuvor definierten Ziel verglichen wird. Erfolg stellt einen Wert dar, der sich in Abhängigkeit von getroffenen Entscheidungen und daraus resultierenden Aktivitäten unterschiedlich ausprägt. Der Erfolg ist damit eine abhängige Variable (Müller-Hagedorn et al., 2012).

Sowohl der betriebswirtschaftliche Erfolg eines Unternehmens als auch ein gesellschaftlicher Erfolg lassen sich definieren.

Erfolg kann quantitativ oder qualitativ sein. Quantitativ messbar sind absolute oder relative Erfolgsgrößen. Absolute Erfolgsgrößen sind zum Beispiel Umsatz, Deckungsbeitrag, Bilanzgewinn, das betriebswirtschaftliche oder steuerliche Ergebnis oder aber auch Artikel- oder Auftragserfolg sowie Betriebs- oder Finanzerfolg. Relative Erfolgsgrößen setzen Erfolgsgrößen in Bezug zueinander, wie entsprechende Vorjahres- bzw. Vorperiodenzahl, oder in Bezug zu weiteren Größen, wie Umsatz im Verhältnis zur Fläche oder zur Mitarbeiteranzahl (Müller-Hagedorn et al., 2012). In der Betriebswirtschaftslehre wird Erfolg als „das i.d.R. in monetären Größen erfasste bzw. ausgedrückte Ergebnis des Wirtschaftens; ermittelt durch Erfolgsrechnung“ (Gabler Wirtschaftslexikon, 2018) definiert. Jedoch ist dies nur eine der vielen möglichen Auffassungen von Erfolg.

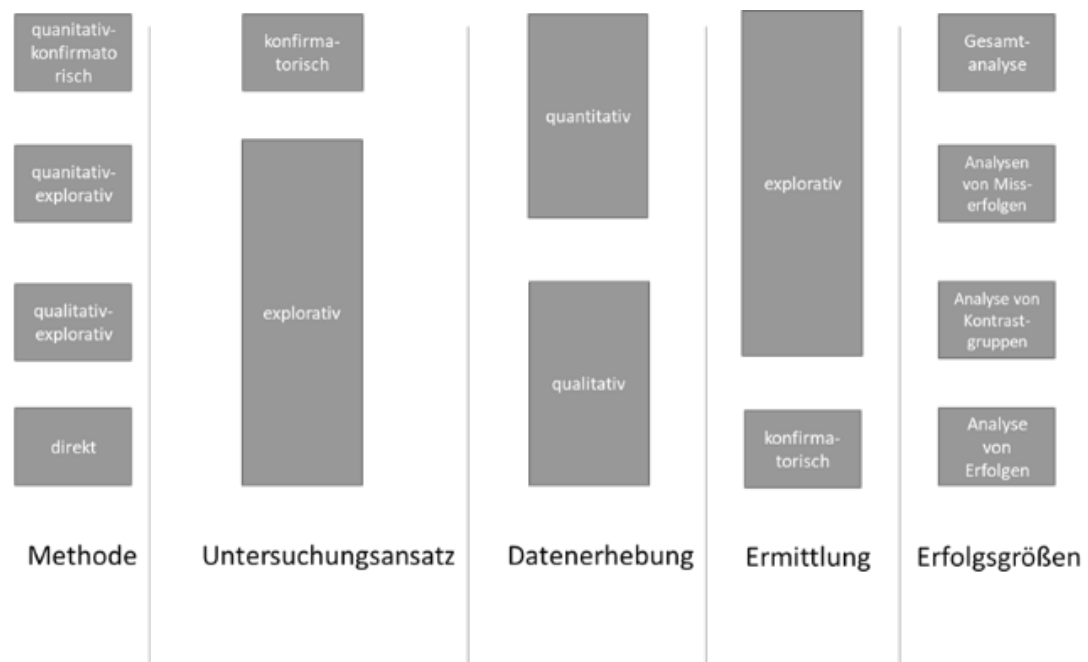
Ziel unternehmerischen Handelns ist es, Erfolg zu haben. Erfolgsfaktorenforschung untersucht die Ursachen für Erfolg und Misserfolg. Es wird versucht, allgemeingültige Erfolgsfaktoren für Unternehmen zu identifizieren und daraus Empfehlungen für Entscheider zu generieren. Grundannahme dabei ist, dass es eine begrenzte Zahl an Variablen gibt, die den Erfolg eines Unternehmens beeinflussen (Haenecke, 2002; Schmalen et al., 2006). Statistische Studien der Erfolgsfaktorenforschung analysieren, welcher Wirkungszusammenhang zwischen Erfolgsfaktoren (als unabhängige Variablen) und Erfolgsgrößen besteht (René Röderstein, 2009). Die Erfolgsfaktorenforschung hat vor allem im Strategischen Management eine große Bedeutung. Das Erkenntnisinteresse in der Wissenschaft liegt in der Erklärung von Erfolgen und Verhinderung von Misserfolgen. Darüber hinaus stellt die Erfolgsfaktorenforschung der Praxis nützliches Wissen zur Verfügung (Herr, 2007).

Es bestehen zahlreiche (wechselseitige) Verbindungen zwischen den bestimmenden Faktoren. Voraussetzungen für die Evaluation von Erfolgsfaktoren sind eine klare Abgrenzung und ein einheitlicher Bezugsrahmen. Zur Evaluierung von Erfolgsfaktoren existieren verschiedene methodische Ansätze. Nach Haenecke (2002) lassen sich fünf methodische Herangehensweisen der empirischen Forschung unterscheiden, die sich grundsätzlich in die direkte und die indirekte Ermittlung spalten. In Abhängigkeit von Einsatz der Erfolgsgröße, Vorgehen bei der Ermittlung der Daten, der Datenerhebung und dem Untersuchungsansatz, kann die Methodik direkt, qualitativ-explorativ, quantitativ-explorativ oder quantitativ-konfirmatorisch sein (siehe Abbildung 23).

Quantitativ-konfirmatorische Erfolgsfaktorenstudien werden in der Regel eingesetzt, um bereits theoretisch und empirisch analysierte Wirkungszusammenhänge zu überprüfen (Haenecke, 2002). In dieser Art der Studie kommen im Gegensatz zu explorativen Studien weniger Variablen in Betracht (Grünig et al., 1996). Der konfirmatorische Ansatz zeichnet sich dadurch aus, dass sich die Ergebnisse häufig detaillierter beschreiben lassen, da in Vorbereitung der Studie die wichtigsten Zusammenhänge bereits bekannt sind. Der konfirmatorische Ansatz stellt hohe Anforderungen an die Stichprobengröße (mit $n \geq 200$), da ansonsten das Risiko besteht, dass Zusammenhänge als nicht signifikant eingestuft werden (Backhaus et al., 2000).

Als Ursprung der Erfolgsfaktorenforschung gilt die PIMS (Profit Impact of Marketing Studies)-Studie (Lange, 1982; Erläuterungen zu PIMS siehe Kapitel 3.1 in dieser Arbeit). Peters und Waterman (1982) wählten anhand von Erfolgsindikatoren Unternehmen aus, die sie auf Basis dieser Indikatoren als erfolgreich kategorisierten und hinsichtlich gemeinsamer Merkmalsausprägungen untersuchten.

Abbildung 23: Methoden der Erfolgsfaktorenforschung. Eigene Darstellung nach Haenecke (2002).

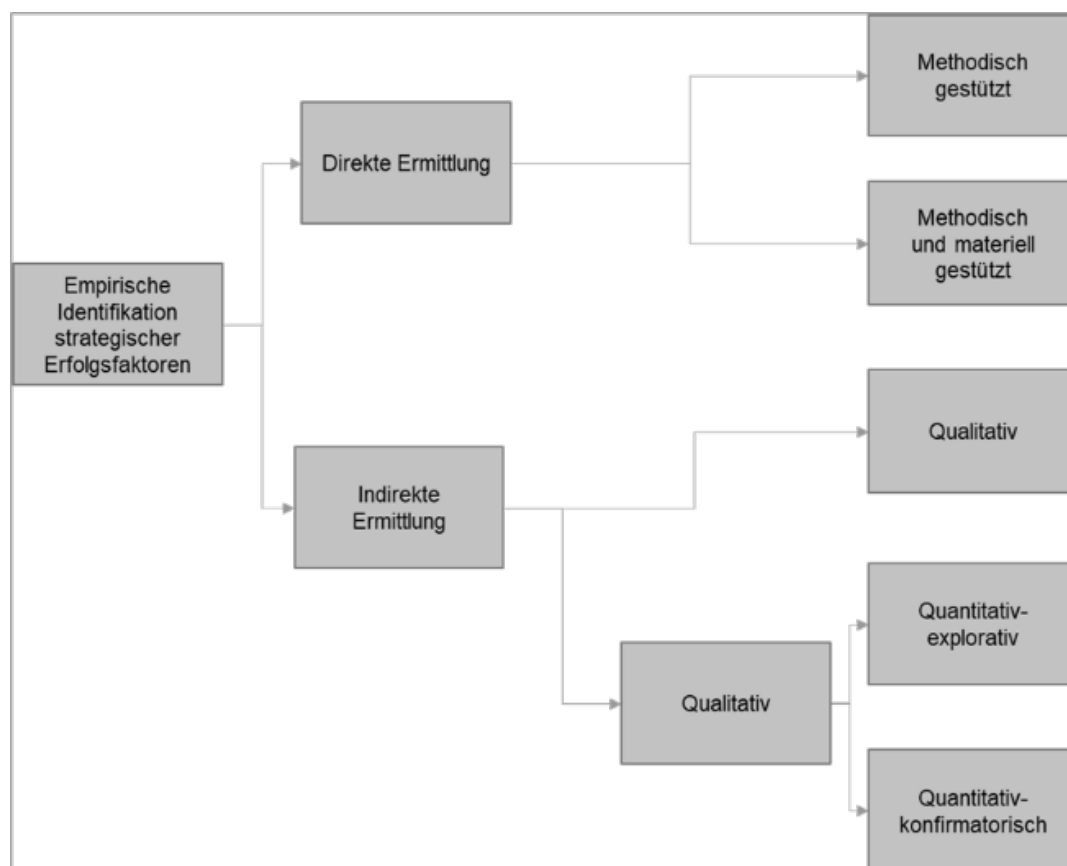


Die Forschungsmethoden lassen sich in direkte und indirekte Ermittlungen von Erfolgsfaktoren unterscheiden (siehe Abbildung 24). Durch direkte Befragung von Experten nach den Erfolgsfaktoren werden diese qualitativ und methodisch gestützt ermittelt. Eine indirekte Ermittlung erfolgt über den Einsatz von quantitativen Methoden oder ebenfalls über den Einsatz qualitativer Forschung. Mit der indirekten Ermittlung von Erfolgsfaktoren wird geprüft, ob und welcher Zusammenhang zwischen möglichen Erfolgsfaktoren (unabhängige Variable) und einer Messgröße (abhängige Variable) besteht.

Der Einsatz quantitativer Methoden scheint grundsätzlich besser geeignet, da alle Objekte hinsichtlich eines oder mehrerer Merkmale analysiert werden, um unter Anwendung statistischer Verfahren signifikante Zusammenhänge und Erfolgsfaktoren aufzudecken. Ein umfassender Datensatz, der sowohl eine Vielzahl von Fällen und eine entsprechend hohe Anzahl und vollständige Daten zu den relevanten Indikatoren beinhaltet, ist Voraussetzung für die Anwendung dieser Methodik. Mit einer ausreichenden

Fallanzahl (Nachwuchsleistungszentren, N=54) und quantifizierbaren Daten kann diese Methodik im Rahmen der vorliegenden Untersuchung angewendet werden.

Abbildung 24: Methoden zur Identifikation von Erfolgsfaktoren. Eigene Darstellung nach Haenecke (2002).



Das Ziel der Erfolgsfaktorenforschung besteht darin, Unternehmenserfolg auf bestimmte Einflussfaktoren zurückzuführen. Zahlreiche Studien in der Erfolgsfaktorenforschung wurden dahingehend kritisiert, dass es nicht möglich sei, eindeutige Erfolgsfaktoren zu identifizieren und nur schwer erkennbar sei, welchen Anteil ein bestimmter Erfolgsfaktor vom Unternehmenserfolg erklärt (Abfalter, 2010). Eine Vielzahl miteinander im Zusammenhang stehender interner und externer Variablen bestimme den Unternehmenserfolg, und dabei könne die Erfolgswirksamkeit der einzelnen Variablen nicht isoliert werden (Haenecke, 2002).

Nicolai und Kieser (2002) initiierten mit ihren Aussagen, dass einzelne Erfolgsfaktoren an Wert und Bedeutung verlören, sobald sie identifiziert und von Unternehmen implementiert würden und sich nicht mehr voneinander unterscheiden ließen, in der deutschen betriebswirtschaftlichen Forschung einen wissenschaftlichen Diskurs bezüglich der Kritik von Erfolgsfaktorenstudien (vgl. dazu www.dialog-erfolgsfaktorenforschung.de). Darüber hinaus kritisieren sie, dass zu viele Variablen einer Analyse zugrunde gelegt werden, um einige wenige signifikante Ergebnisse zu erzielen.

2.4 Evaluation des Erfolgs im Fußball

Ziel der Nachwuchsleistungszentren im Fußball ist es wie bereits benannt, talentierte Fußballspieler auf die Anforderungen im Lizenzfußball vorzubereiten und mit der Verpflichtung der Sportler für den eigenen Vereinskader oder mit einem Transfer zu einem anderen Fußballverein sportlich und ökonomisch erfolgreich zu sein (DFB, 2020; DFL, 2019b). Viemann et al. (2005) zeigen, dass Sportorganisationen dem wirtschaftlichen und sportlichen Erfolg große Bedeutung beimessen. Woran aber kann Erfolg in diesem Bereich festgemacht werden? Zur Beantwortung dieser Frage wird zunächst der Begriff *Erfolg* definiert. Des Weiteren geht dieses Kapitel auf den aktuellen Forschungsstand zur Beurteilung des Erfolgs im Fußball ein.

2.4.1 Definition von Erfolg

Im Sport gilt generell als Erfolg, wenn ein einzelner Wettkampf, ein Turnier oder eine Meisterschaft gewonnen wird. Im Fußball kann also von Erfolg beim Gewinn eines Spiels gesprochen werden. Gemessen wird er an der relativen Stärke des Gegners, daher kann auch ein Unentschieden oder eine knappe Niederlage als Erfolg interpretiert werden (Wallebohr & Daumann, 2019). Punkte, die im Rahmen einer Spielzeit oder im Wettbewerb auf (inter-)nationaler Ebene (z. B. Meisterschaft, Pokal, Champions League) errungen werden, gelten als die gebräuchlichste Maßzahl, um den Erfolg eines Vereins auf sportlicher Ebene darzustellen (Klimmer, 2003). Der ökonomische Erfolg ist das Ergebnis der wirtschaftlichen Tätigkeit eines Unternehmens (oder auch Vereins) während eines Betrachtungszeitraumes. Der Erfolg kann als Differenz von Wertentstehung und Werteverzehr in einer Periode positiv (Gewinn) oder negativ (Verlust) sein. „In der gleichen Periode kann der Erfolg als Differenz beider Größen positiv (Gewinn) oder negativ (Verlust) sein“ (Duden, 2016). Um Erfolg langfristig zu sichern und zu beurteilen, ob gemessen an den eigenen Erfolgsmaßstäben die Ziele erreicht wurden, ist eine umfassende Evaluation erforderlich (Sontag, 2011).

In einigen Profisportarten wird eine zunehmende Ökonomisierung wahrgenommen, insbesondere der Profifußball erlebt eine unvergleichliche Ökonomisierung (Follert, 2018). Lizenzspielerabteilungen der Vereine sind als Kapitalgesellschaften aufgestellt, die entweder an der Börse notiert sind oder sich im Besitz des Vereins oder von Investoren und Unternehmen befinden; das wiederum führt dazu, dass an der Fußballbranche Beteiligte ökonomische Aspekte in den Fokus stellen (Follert, 2018). Im professionellen Teamsport ist der wirtschaftliche Erfolg stark vom sportlichen Erfolg des Vereins abhängig (Jörn Littkemann & Kleist, 2002). Die Erreichung von Rentabilitätszielen und finanzieller Sicherheit ist insbesondere in der Fußballbranche direkt mit der Erreichung von Sachzielen, wie der Güterherstellung und des -vertriebs, verbunden (Chadwick & Parker, 1996). In der deutschen Fußball-Bundesliga kann schon *ein* sportlicher Misserfolg innerhalb einer Saison zu einem wirtschaftlichen Einbruch führen, wenn nicht aus erfolgreichen Spielzeiten entsprechende Rücklagen gebildet werden konnten (Jörn Littkemann & Kleist, 2002).

Der Forschungsstand einiger in der Praxis eingesetzten Erfolg messenden Indikatoren auf sportlicher und auf ökonomischer Ebene wird im Folgenden beschrieben. Darüber hinaus wird aufgezeigt, wie der Erfolg eines Nachwuchsleistungszentrums bewertet werden kann.

2.4.2 Der sportliche Erfolg von Vereinen

Kennzahlen wie z. B. die Tabellenplatzierungen der Vereine in ihren Ligen, die erreichten Punktzahlen, die über mehrere Jahre in einer ewigen Tabelle dokumentiert sind oder einfache Zählungen von Spielererfolgen und -leistungsdaten sind gängige Praxis, um den sportlichen Erfolg von Vereinen und Spielern dazustellen (Wallebohr & Daumann, 2019). Jörn Littkemann und Kleist (2002) stellen fest, dass sich der Spielerfolg, gemessen an der erreichten Punktzahl aus den Spielergebnissen, insbesondere auf das Spielsystem und die Spielerleistung zurückführen lassen. Der DFB führt eine ewige Tabelle zum sportlichen Erfolg der Lizenzmannschaften, die an den Meisterschaftsrunden der Fußball-Bundesliga in Deutschland teilgenommen haben. Die Berechnung des sportlichen Erfolgs erfolgt auf Basis der absolvierten Spiele, geschossenen und erhaltenen Tore sowie der erreichten Punktzahl seit 1963. Ab der Saison 95/96 gilt die Drei-Punkte-Regel (drei Punkte pro Sieg, ein Punkt pro Unentschieden), davor galt die Zwei-Punkte-Regel (Bundesliga, 2021). Der Verein mit der höchsten Punktzahl ist sportlich am erfolgreichsten. Als Top 3 sind Bayern München (Platz 1), Borussia Dortmund (Platz 2) und Werder Bremen (Platz 3) nennen (DFB, 2021b).

Den europäischen Fußball vertritt als Dachverband, mit 55 Nationalverbänden, die Union der Europäischen Fußballverbände (UEFA). Die UEFA führt zur Messung des sportlichen Erfolgs von Vereinen eine Klubkoeffizienten-Rangliste basierend auf den Ergebnissen der Vereine, die in den letzten fünf Jahren an der UEFA Champions League oder UEFA Europa League teilgenommen haben. Auf der Basis dieser Rangliste wird entschieden, an welche Stelle die Vereine bei der Auslosung der entsprechenden Europapokalwettbewerbe gesetzt werden. Die Summe aller Punkte, die in den letzten fünf Jahren erzielt wurden, oder ein Verbandskoeffizient bestimmen den Koeffizienten des Vereins. Ein Punktesystem für die Teilnahme an der UEFA Champions League sowie der UEFA Europa League schlagen sich in der Klubkoeffizienten-Rangliste der Vereine nieder. Darüber hinaus wird – basierend auf den gleichen Regeln – eine 10-Jahres-Klubkoeffizienten-Rangliste geführt (UEFA, 2021).

Um den individuellen sportlichen Erfolg von Spielern zu dokumentieren, werden die vom Spieler auf dem Spielfeld erbrachten Leistungen mittels verschiedener Kennzahlen erfasst. Seit einigen Jahren stellen Dienstleister im Bereich Datenbereitstellung, wie z. B. Optasports, Live- oder historische Daten, Widgets für visuelle Daten oder innovative analytische Metriken als Problemlösungen zur Erfassung und Messung von Fußballdaten zur Verfügung (Martinez-Arastey, 2018). Die von Optasports erhobenen Daten fließen zum großen Teil in die Publikation der Leistungsdaten der Spieler auf Transfermarkt.de ein (Haidl, 2021). Die Impire AG, seit der Saison 2011/2012 von der DFL zur

Erhebung der offiziellen Spieldaten eingesetzt (DFL, 2011), dokumentiert die Spielerleistungen für jeden einzelnen Bundesligaspieler auf der Basis jeder einzelnen Spielsequenz jedes einzelnen Bundesligaeinsatzes und rechnet diese Daten in ein Punktesystem um, sodass die spielerische Leistung über einzelne Spielpositionen vergleichbar und intersubjektiv nachprüfbar ist (Rosar et al., 2013). Jörn Littkemann und Kleist (2002) bedienen sich in ihrer Studie zum sportlichen Erfolg in der Fußball-Bundesliga der Leistungsbewertung pro absolviertem Bundesliga-Spiel des Kicker-Sportmagazins. Die Bewertung durch das Kicker-Sportmagazin erfolgt an jedem Bundesligaspieltag für jeden Spieler, der mindestens 30 Minuten im Einsatz war (Eschweiler & Vieth, 2004; zur Objektivität der verwendeten Maßzahl siehe Rosar et al., 2013). Die sportlichen Erfolge der Spieler, auf der sich viele wissenschaftliche Untersuchungen beziehen, werden mit objektiven Kennzahlen, wie z. B. Anzahl Einsätze in der Bundesliga und der (Jugend-)Nationalmannschaften, Anzahl geschossener Tore (bei Torwarten die Spiele ohne Gegentor), erreichte Titel (Torschützenkönig, (Vize-)Weltmeister, Pokalsieger, Meister), gelaufene Kilometer u. v. a. m., gemessen.

2.4.3 Der ökonomische Erfolg von Vereinen

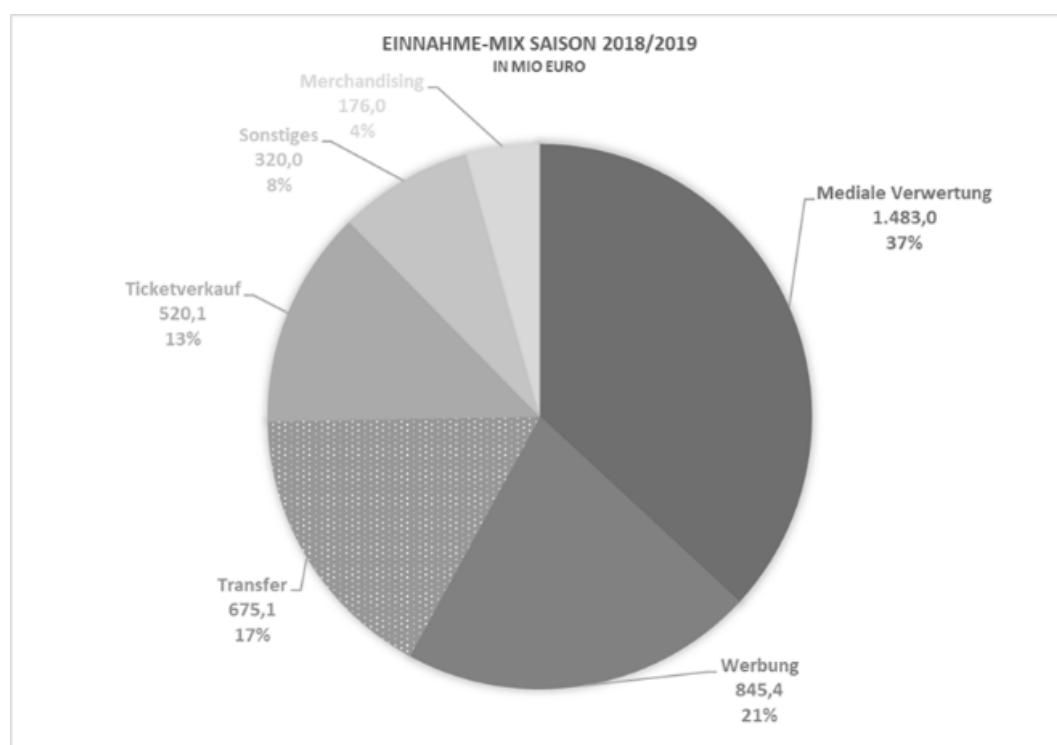
Zwischen dem sportlichen Erfolg der Vereine und deren ökonomischem Erfolg existiert ein direkter Zusammenhang (G. Schewe & Littkemann, 1999; Jörn Littkemann & Kleist, 2002). Die Deutsche Fußball Liga – ein Unternehmen, das die Aktivitäten des Lizenzfußballs der 1. und 2. Bundesliga in Deutschland repräsentiert – veröffentlicht jährlich einen Wirtschaftsreport zur ökonomischen Gesamtsituation des deutschen Lizenzfußballs (DFL, 2020a). Der ökonomische Erfolg des deutschen Lizenzfußballs wird jeweils kumuliert für die 1. und 2. Bundesliga anhand diverser gängiger betriebswirtschaftlicher Kennzahlen, wie z.B. Umsatz und Erlöse, ausgewiesen. Auch die Anzahl der Vereine, die mit einem Jahresüberschuss abschließen können, die insgesamt erzielte Eigenkapitalquote sowie die Höhe des fiskalischen Effekts – Steuern und Abgaben an die Finanzbehörden und der Sozialversicherungen – werden als Indikatoren für den ökonomischen Erfolg angeführt (DFL, 2020a).

Die Vereine der 1. und 2. Bundesliga erzielen ihre Einnahmen aus den Bereichen *Ticketverkauf* für (inter-)nationale Spiele, *Werbung* (Haupt- und Trikotsponsoren sowie stadion- und vereinsgeborenen Rechte), *mediale Verwertung* (Vermarktung von Medienrechten an (inter-)nationalen Spielen, *Transfers* (Transfer- und Ausbildungsentschädigungen), *Merchandising* und *Sonstiges* (Überlassung von Nutzungsrechten, Catering, Vermietung und Verpachtung, Mitgliedsbeiträge u. ä.) (DFL, 2020a). Eine Übersicht über die prozentuale Verteilung der Bereiche auf die Gesamteinnahmen beispielhaft für die Saison 2018/19 zeigt Abbildung 25. Einnahmen durch Werbung und mediale Verwertung mit zusammen 58 % machen den wesentlichen Anteil der Einnahmen aus. Transfers und Ausbildungsentschädigungen schlagen mit 17 %, der Ticketverkauf mit 13 %, Sonstiges mit 8 % und Merchandising mit 4 % zu Buche.

Eschweiler und Vieth (2004) haben in ihrer Arbeit zu den Preisdeterminanten bei Spielertransfers in der Fußball-Bundesliga untersucht, inwiefern die Höhe von *Ablösesummen*

in der Fußball-Bundesliga von branchen-, vereins- oder spielerspezifischen Determinanten bestimmt werden. Das von ihnen entwickelte Notenmodell verwendet für die spielerspezifischen Determinanten sowohl objektive Daten, wie sie Optasports erhebt, als auch die Beurteilungsnote des Kicker-Sportmagazins. Die Abbildung 26 zeigt inhaltlich eine Übersicht der verwendeten unabhängigen Variablen im Notenmodell. Ergebnis der Studie von Eschweiler und Vieth (2004) ist, dass die Qualität eines Spielers den größten Einfluss auf

Abbildung 25: Einnahme-Mix 1. und 2. Bundesliga in der Saison 2018/19. Eigene Darstellung nach DFL (2020a).

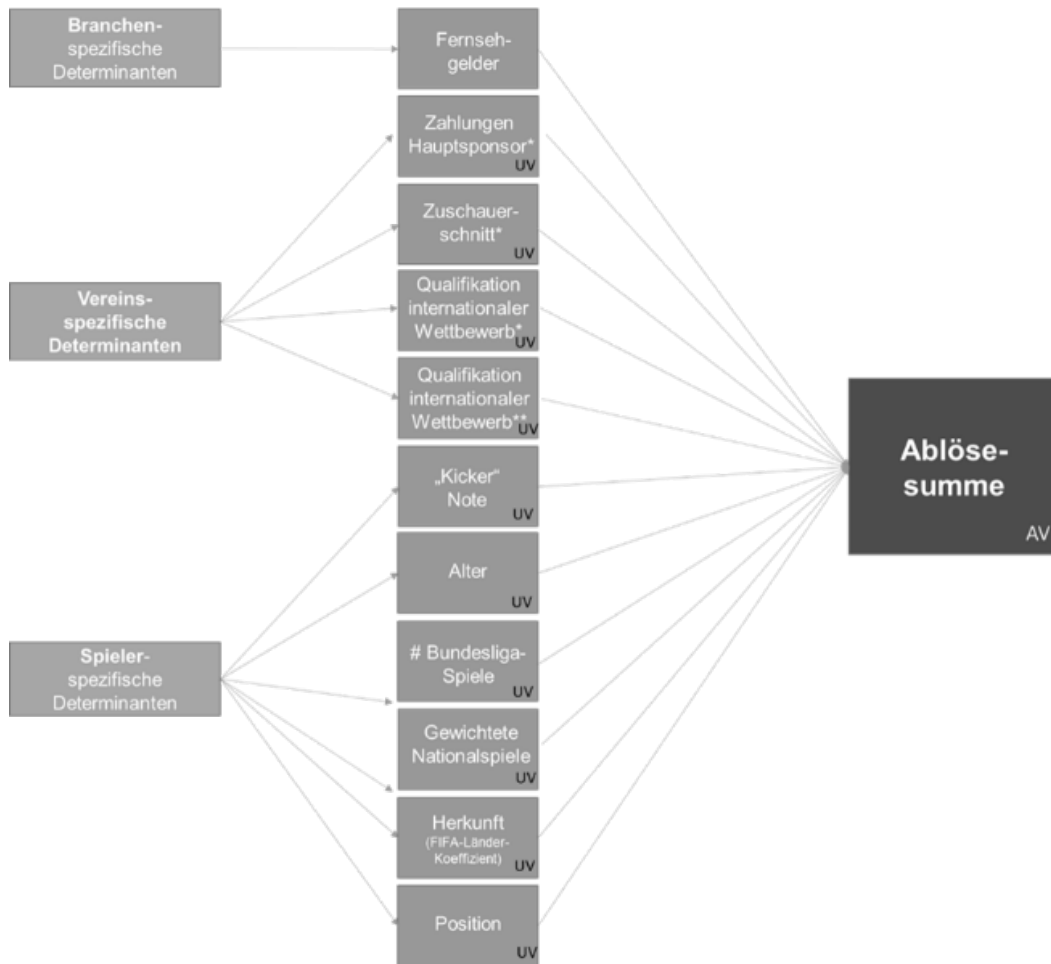


die Messung des Erfolgs eines Spielers anhand seines Transferpreises hat und im Wesentlichen den Transferpreis bestimmt. Die Summe der Transferpreise aller Spieler wird von Eschweiler und Vieth (2004) als eine Kennzahl für die Messung des ökonomischen Erfolgs eines Vereins interpretiert.

Der *CIES-Transferwert* versucht ebenfalls, den ökonomischen Erfolg eines Spielers darzustellen. Auf individueller Spielerebene gehen Alter, Position, verbleibende Vertragsdauer und der Restbuchwert (vom aktuellen Verein gezahlte Transfergebühr dividiert durch den Prozentsatz der Vertragsjahre seit der Vertragsunterzeichnung) sowie die Leistungen des Spielers (z. B. Einsatzzeiten in den verschiedenen nationalen und internationalen Wettbewerben und Nationalmannschaften) in die Bewertung ein. Darüber hinaus werden auch das Niveau der Liga, in denen der Spieler spielt bzw. gespielt hat, sowie der sportliche Erfolg der Vereine des Spielers berücksichtigt (CIES, 2020). Die

Summe der Transferpreise aller Spieler eines Vereins kann als eine Kennzahl für die Messung des ökonomischen Erfolgs eines Vereins interpretiert werden.

Abbildung 26: Übersicht verwendete unabhängige Variablen im Notenmodell. Eigene Darstellung nach Eschweiler und Vieth (2004).



* laufender Verein; ** verkaufender Verein; UV = unabhängige Variable; AV = Abhängige Variable

Einen weiteren Ansatz, den ökonomischen Erfolg eines Vereins zu messen, verfolgt Transfermarkt.de mit der Aufsummierung der Marktwerte aller Spieler zu einem *Marktwert-Verein-Gesamt Transfermarkt.de*. Besucher und Experten der Internetplattform Transfermarkt.de legen nach öffentlichen Diskussionen unter Einbeziehung von individueller Leistung, internationalem Prestige und Karriereverlauf des Spielers sowie dem Status des Vereins als Teil der Bewertung den Marktwert für einen Spieler fest (Lanwehr et al., 2021). Die Summe der Marktwerte aller Spieler wird ebenfalls als eine Kennzahl für die Messung des ökonomischen Erfolgs eines Vereins interpretiert.

2.4.4 Erfolg in der Nachwuchsausbildung

Vorrangiges Ziel der Nachwuchsausbildung ist es, junge, talentierte Spieler an den Lizenzfußball heranzuführen (DFB, 2021c). Die DFL gibt an, dass durchschnittlich ca. 5.600 talentierte Spieler im Alter von 12 bis 19 Jahren in den Nachwuchsleistungszentren ausgebildet werden und in 279 Mannschaften spielen (DFL, 2019a). Für die Einschätzung des sportlichen Erfolgs der Nachwuchsausbildung werden häufig Kennzahlen eingesetzt, die die Anzahl der vom Nachwuchsleistungszentrum hervorgebrachten Profifußball-Spieler misst. Aber eine einheitliche Definition einer Kennzahl für die Messung des Erfolgs in der Nachwuchsausbildung fehlt, zu unterschiedlich sind die beschriebenen Selektions- und Auswertungskriterien (Stügelmaier, 2016; Güllich, 2014; Meylan et al., 2010). Einige Nachwuchsleistungszentren sehen einen Erfolg darin, wenn ein Spieler aus der eigenen Jugend die eigene U23 verstärkt oder in einem anderen Verein einen Vertrag der vierthöchsten Spielklasse erhalten hat, andere auch dann, wenn ein Spieler aus der U23 (mit 24 Jahren) in den Lizenzfußball wechselt (Stügelmaier, 2016). Güllich (2014) dagegen legt als Erfolgskriterium einen Vertrag in einer der zwei Lizenzligen nach der U19 als Maßstab fest. Laut Sportschau.de (2020) schafft es nur einer von zehn jungen Talenten aus einem Nachwuchsleistungszentrum eines Bundesligaverbands, Profifußballspieler zu werden. Aus einer Untersuchung der U19-Kaderlisten eines Nachwuchsleistungszentrums für die Jahre 1999-2013 geht hervor, dass von 323 Spielern insgesamt 48 Spieler in einer der drei Lizenzligen ihre Profikarriere starteten (15 %), davon 25 in der 1. Bundesliga (8 %), 18 in der 2. Bundesliga (6 %) und fünf in der Liga 3 (2 %; Stügelmaier, 2016). Eine Berechnung der Erfolgsquote in Abhängigkeit von der Verweildauer des Talents im Nachwuchsleistungszentrum ist nicht bekannt, auch eine Studie der Erfolgsquote von U15- oder U17-Spielern fehlt (Stügelmaier, 2016).

Das *International Centre for Sports Studies* (CIES) hat in einer Studie mit der Swiss Football League (SFL) die Ausbildungsqualität der Vereine anhand der Kennzahl *Anzahl an Big-5-Spielern* gemessen. Für die Analyse berücksichtigt wurden Spieler, die im Herbst des entsprechenden Jahres der Beobachtung im Kader der 1. Mannschaft stehen und in der betreffenden Saison in der Meisterschaft einer der Big-5-Ligen zum Einsatz gekommen sind bzw. die zumindest in den beiden vorangegangenen Saisons jeweils Einsätze in der höchsten Liga absolviert haben. Als Ausbildungsverein wird der Verein definiert, in dem der Spieler im Alter zwischen 15 und 21 Jahren mindestens drei Saisons gespielt hat. Ein Spieler kann mehrere oder gar keinen Ausbildungsverein haben. Den prozentualen Anteil im Verein ausgebildeter Spieler bilden die Spieler, deren aktueller Verein auch dem Ausbildungsverein entspricht (CIES, 2017b).

Im Rahmen einer Studie zum aktuellen Stand der Jugendakademien bzgl. der administrativen Situation der Vereine, ihrer Vision, Mission, Strategie, Werte und Philosophie in der Saison 2016/2017 hat die *European Club Association* (ECA) neben anderen, input-orientierten Kennzahlen die Qualität der Ausbildung über die Anzahl von Spielern, die Profifußballspieler geworden sind, gemessen (ECA, 2019). Eine Definition, wann ein Spieler als Profifußballspieler kategorisiert wird, ist nicht ersichtlich. Weitere Faktoren, die die Anzahl der hervorgebrachten Spieler hinsichtlich ihrer Qualität bewerten, werden

nicht angeführt und das, obwohl eine der Empfehlungen aus einem ECA-Bericht 2012 besagt, dass sich die Vereine auf Qualität statt auf Quantität konzentrieren sollten (ECA, 2012).

Gängige verwendete Indikatoren, wie die Ewige Tabelle des DFB, der UEFA-Klubkoeffizient, die Anzahl der Big-5-Spieler der CIES, die Anzahl hervorgebrachter Profifußballspieler der ECA, der CIES-Transferwert, die Ablösesumme nach Eschweiler & Vieth und der Marktwert von Transfermarkt.de können nur näherungsweise den Erfolg von Nachwuchsleistungszentren messen.

Die *Ewige Tabelle des DFB* als auch die *Clubkoeffizienten-Ranglisten der UEFA* lassen keine direkten Rückschlüsse auf die Qualität der Nachwuchsausbildung zu, da Vereine, die über ein Nachwuchsleistungszentrum verfügen, aber nicht an europäischen Wettbewerben teilnehmen, hierin nicht gelistet sind. Darüber hinaus werden die Erfolge der gelisteten Vereine nicht nur durch Spieler, die in dem zugehörigen Nachwuchsleistungszentrum selbst ausgebildet wurden, sondern auch durch zugekaufte, zu anderen Zeiten und in anderen Nachwuchsleistungszentren Ausgebildete erzielt. Der *UEFA Clubkoeffizient* bewertet den Erfolg eines Vereins anhand der erreichten Punktzahlen in europäischen Wettbewerben. Da nur 15,1 % der Spieler eines Vereins in den deutschen Profiligen in dem zu diesem Verein gehörenden Nachwuchsleistungszentrum ausgebildet worden sind, erscheint diese Bewertung als nicht aussagekräftig genug, um den Erfolg eines Nachwuchsleistungszentrums zu messen. Mit einem Anteil der zum Kader hinzugeholten Spieler von 84,9 % kann der Erfolg eher diesen Spielern zugeordnet werden (CIES, 2018a). Ein Rückschluss auf den Erfolg eines Nachwuchsleistungszentrums ist aber damit nicht möglich (Wallebohr & Daumann, 2019). Die Kennzahlen *CIES Anzahl Big-5-Spieler* und *ECA Anzahl hervorgebrachter Profifußballspieler* zielen bereits auf die Bewertung der individuellen Performance eines Spielers und nicht des Vereins ab. Die Einstufung der Qualitäten eines Spielers schwingt bei den genannten Indikatoren implizit mit, da der Spieler so gute Kompetenzen entwickelt hat, dass er einen Profivertrag erhält und das eventuell sogar in einer der Top-Ligen der sogenannten Big-5-Länder (Spanien, England, Deutschland, Italien und Frankreich). Einerseits ermöglichen diese spezifischen Merkmale eine genauere Messung der hervorgebrachten Spieler, andererseits wird aber die individuelle Stärke eines Spielers lediglich abgebildet über die Tatsache, dass er in einer der Ligen der Big-5-Länder eingesetzt wird. Die individuelle Qualität eines Spielers kann damit aber nicht ausreichend bewertet werden (Lanwehr et al., 2021). Eine Bewertung der Stärke des Vereins bzw. der jeweiligen Liga wird ebenfalls nicht berücksichtigt (Wallebohr & Daumann, 2019). Der *CIES-Transferwert*, die *Ablösesumme nach Eschweiler & Vieth* und der *Marktwert von Transfermarkt.de* beziehen jeweils spieterspezifische Determinanten in die Bewertung ein, berücksichtigen aber zu einem großen Teil in ihren Kalkulationen subjektive Einschätzungen, wie insbesondere des Geschehens auf dem Transfermarkt, die Prognosen zukünftiger Einnahmen aus Merchandising oder die Teilnahme an internationalen Wettbewerben – die Qualität des Spielers aber wird monetär bewertet. Das hohe Maß an Subjektivität auf mehreren Ebenen (Interpretationen von Gerüchten, Erzeugung prospektiver Daten, die wiederum abhängig sind von den Erfahrungen

des Einzelnen) verzerrt die Evaluation der Spielerqualität und damit die Bewertung der Leistung der Nachwuchsleistungszentren. Damit fehlt nach wie vor eine zugleich valide und effiziente Methode (Lanwehr et al., 2021). Wallebohr und Daumann (2019) verfolgen für eine ähnlich gelagerte Fragestellung in der Schweiz einen ähnlichen Ansatz, jedoch fehlt eine Validierung der Indikatoren anhand geeigneter Kriterien (Lanwehr et al., 2021). Eine Übersicht der gängigen *output*-orientierten Indikatoren und eine Einschätzung ihrer Eignung zur Messung des Erfolgs von Nachwuchsleistungszentren zeigt Tabelle 2.

Die Beurteilung der Eignung der Indikatoren erfolgt anhand der Anforderungskriterien, die an das mit dieser Arbeit zu entwickelnde Evaluationskonzept gestellt werden – Messung über einen Zeitraum, Zugänglichkeit der Daten, Grad der Objektivität der Daten, Selektion, Berücksichtigung individueller Spielerqualität, Aufwand, Komplexitätsgrad, Ökonomische Kennzahl und Berücksichtigung der Ausbildungszeiten in einem Nachwuchsleistungszentrum (vgl. dazu Ziel der Arbeit Kapitel 1.3). Keiner der Indikatoren kann sämtliche Bewertungskriterien in hohem Maße erfüllen, sie erfüllen lediglich teilweise die dieser Arbeit zugrundeliegenden Anforderungen (siehe grün markierte Felder in der Tabelle 2). Eine Kennzahl *Spielerqualität* auf Basis objektiver Daten ist bisher nicht bekannt.

Tabelle 2: Übersicht der gängigen output-orientierten Indikatoren und eine Einschätzung ihrer Eignung zur Messung des Erfolgs von Nachwuchsleistungszentren. Eigene Darstellung nach Lanwehr et al. (2021).

Outputorientierte Indikatoren						
Indikator	UEFA°	CIES°	ECA°	CIES°	Eschweiler & Vieth	Transfermarkt.de
Maßzahl	Club-koeffizient	Anzahl an Big-5*- Spielern	Anzahl von Spielern, die Profifußball-spieler geworden sind	Transferwert	Ablöse-summe	"Marktwert"
	Quotient auf Basis der Ergebnisse der Vereine in der UEFA Championsleague und UEFA Europaleague oder Verbandskoeffizient	Spieler, die nach ihrer Ausbildung in einem Verein der Big-5-Ligen Profifußballspieler wurden	Spieler, die nach ihrer Ausbildung in einem Verein Profifußballspieler wurden	Schätzung auf Basis der in den letzten 6 Jahren abgeschlossenen Profispielerverträge	Notenmodell auf Basis von branchen-, vereinspezifischen und spieterspezifischen Determinanten	Berücksichtigung von Alter, Leistung, Zukunftsperspektive, real existierende Nachfrage, Prestige, Marketingtechnische Aspekte
Berechnungsbasis	Spielergebnisse	Anzahl	Anzahl	Algorithmus	Algorithmus	Schwarmintelligenz / Schätzung
Kriterien	Bewertung					
Berücksichtigung eines Zeitraums	Längsschnitt möglich	Längsschnitt möglich	Längsschnitt möglich	Längsschnitt möglich	Längsschnitt möglich	Längsschnitt möglich
Zugänglichkeit der Daten	öffentlich zugänglich	öffentlich zugänglich	öffentlich zugänglich	teilweise öffentlich zugänglich	teilweise öffentlich zugänglich	öffentlich zugänglich
Grad der Objektivität der Daten	hoch: Spielergebnis-orientiert	hoch: Anzahlorientiert	hoch: Anzahlorientiert	mittel: enthält Inflation; Preiseffekte des Transfermarktes fließen ein	mittel: enthält Schätzungen und Preiseffekte des Transfermarktes; aber hohe Korrelation zu Marktwerten	niedrig: subjektive Wahrnehmung der Bewerter, Preiseffekte des Transfermarktes fließen mit ein, aber hohe Korrelation zu tatsächlichen Marktwerten soweit bekannt
Selektion	nur ermittelbar, wenn Verein auch international spielt	Profifußballer in Big-5-Liga (ja/nein zum Erhebungszeitpunkt)	Definition "Beginn Profifußballspieler" nicht eindeutig	nur Spieler, deren Daten in allen Quellen vorhanden sind	nur Spieler, deren Daten in allen Quellen vorhanden sind	alle Spieler, die in den betrachteten Ligen spielen
Berücksichtigung individueller Spielerqualität	nein: der sportliche Erfolg des Vereins steht im Vordergrund	nein: nur implizit über den Profivertrag in Big-5	nein: nur implizit über einen Vertrag als Profi	ja - Über individuelle Leistung, internationalen Status des Spielers, Karriereverlauf sowie dem sportlichem Niveau des Vereins als Teil der Bewertung	ja - Annahme hoher Transferwert = hohe Qualität	ja: Über individuelle Leistung, internationales Prestige des Spielers, Karriereverlauf sowie den Status des Vereins als Teil der Bewertung
Aufwand	einfache Erhebung der Daten	einfache Erhebung, wenn Quellen identifiziert	einfache Erhebung, wenn Quellen identifiziert	aufwändige Datenerhebung: viele Datenquellen Verein: Sportliches Niveau, Wirtschaftliches Niveau;	aufwändige Datenerhebung: viele Datenquellen, komplexer Algorithmus	sehr aufwändig: Schätzungen nach umfangreichen Diskussionen und Festlegung durch einen Moderator
Komplexitätsgrad	gering	gering	gering	hoch: komplexer Algorithmus: komplexes Variablen-system: Spieler: Vertragslänge, Alter, Internationaler Status, Karriereverlauf, Leistung;	hoch: komplexes Preisdeterminanten-System (16 Elemente) in 3 Kategorien: Branche, Verein, Spieler	sehr hoch: komplexes Variablen-system: Spieler: Alter, Verletzungsanfälligkeit, Interesse des Marktes, Karriereverlauf, Leistung, Potential Vermarktung, Status im Verein, Höhe Gehalt, Position, Internationales Prestige; Verein: Status u.v.a.m.; Standardisierung über Leitlinien
Ökonomische Kennzahl	nein	nein	nein	EURO	EURO	EURO
Berücksichtigung Ausbildungszeit Spieler in einem Nachwuchsleistungszentrum	nein	nein	nein	nein	nein	nein

*UEFA = Union of European Football Associations; CIES = International Centre for Sports Studies, Neuchâtel, Switzerland; ECA = European Club Association

3 Empirische Untersuchung – Betriebswirtschaftlicher Erfolg von Nachwuchsleistungszentren

Das umfassende Ziel der Arbeit ist die Entwicklung eines Evaluationskonzepts, das am **Outcome** orientiert ist, die Resultate **im Zeitverlauf** erhebt, **öffentlich zugängliche objektive Daten** nutzt, **Individual-Leistungs-Indikatoren** der Teilnehmer berücksichtigt, mit **geringem Aufwand** leicht anzuwenden ist und mittels geeigneter **betriebswirtschaftlicher Kennzahlen** den Erfolg von Nachwuchsleistungszentren im deutschen Profifußball bewerten kann.

Um dieses Ziel zu erreichen, ist es erforderlich, das gedankliche Grundgerüst, das die relevanten Komponenten und ihre Beziehungen untereinander zeigt, in einem mentalen Modell abzubilden. Ein Modell stellt eine vereinfachte Abbildung eines Teils der Realität in bildlicher oder grafischer Form sowie einer mathematischen Beschreibung dar (Helfrich, 2016). Im Hinblick auf die Zielsetzung eines Modells erfolgt die Reduktion der Komplexität der Realität (Abstraktion) auf die relevanten bzw. wichtigen Komponenten und ihre Beziehungen (Helfrich, 2016).

3.1 Mentales Modell

Der Entwicklung des Evaluationskonzepts in dieser Arbeit liegt ein einfaches mentales Modell mit den als relevant bewerteten Komponenten zugrunde (siehe Abbildung 27). Das Gestaltungsfeld wird situativ beeinflusst von Komponenten, die entweder direkt die Konzipierung der Evaluation, die Strukturen oder den betriebswirtschaftlichen Erfolg eines Nachwuchsleistungszentrums im deutschen Profifußball betreffen. Der Grad der Digitalisierung der Evaluation, die Art und Weise von Datenerhebungen und die anzulegenden Evaluationsstandards beeinflussen unmittelbar die Gestaltung des Evaluationskonzepts. Die nationalen Lizenzierungsbestimmungen erfordern die Bildung von Nachwuchsleistungszentren und ihre Strukturen (Kosten), während durch die internationalen Regularien der FIFA und der UEFA rund um die Betreuung der Talente und Entschädigungszahlungen für Ausbildung und Solidaritätsbeiträge potentielle Einnahmen generiert werden können (Monetarisierung des Nutzens). Die allgemeine Wirtschaftskrise sowie die Einnahmeeinbußen während der COVID-19-Pandemie beeinflussen sehr stark die finanzielle Ausstattung der Vereine, die die Nachwuchsleistungszentren führen.

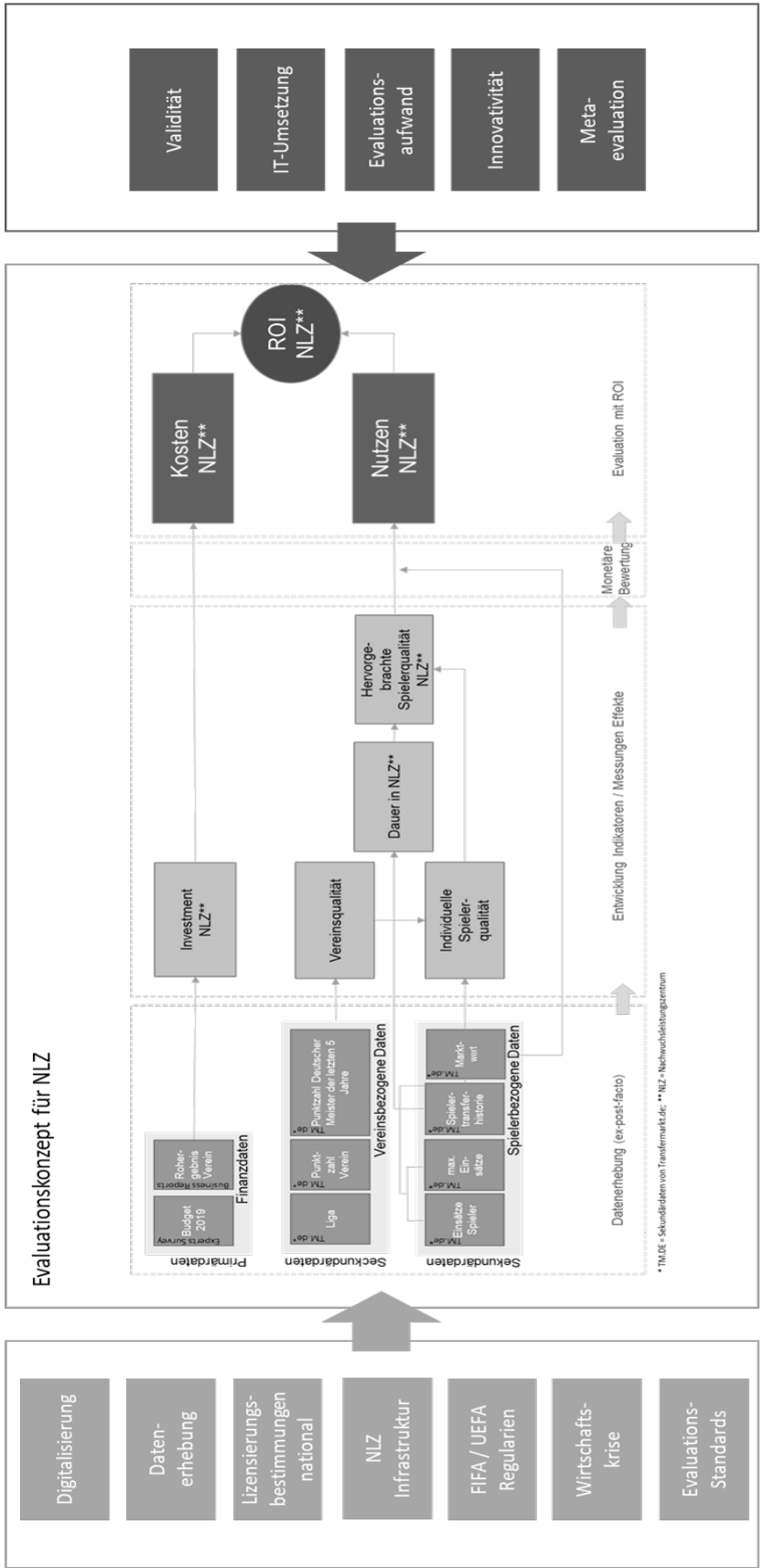
Das Gestaltungsfeld im mentalen Modell visualisiert die verwendeten Variablen in ihrem wechselseitigen Zusammenhang. Aus der grafischen Darstellung geht hervor, welche Variablen erforderlich sind sowie genutzt und konstruiert wurden, um die Evaluation des betriebswirtschaftlichen Erfolgs von Nachwuchsleistungszentren im deutschen Profifußball mittels des ROI zu ermöglichen. Ein Unternehmenserfolg ist auf bestimmte Erfolgsfaktoren oder Maßnahmen des Managements zurückzuführen. Daraus werden

in der Beratungspraxis seit langem Empfehlungen für die Unternehmensführung abgeleitet (Nicolai & Kieser, 2002). Schwaninger und Kaiser (2007) verstehen unter einem Erfolgsfaktor eine „Einflussgröße, die mit dem Unternehmenserfolg korreliert, sodass ein kausaler Zusammenhang abgeleitet werden könnte, wenn auch nicht bewiesen werden kann“ (S. 152).

Zur Identifizierung potentieller Erfolgsfaktoren und der Unternehmensleistung gibt es zahlreiche quantitativ-empirische Studien (Nicolai & Kieser, 2002). Diese sogenannten PaaV-Studien (**P**erformance **a**ls **a**bhängige **V**ariable) erforschen einen statistisch relevanten Zusammenhang zwischen operationalisierten Variablen und dem ROI eines Unternehmens. Zugrundeliegende Annahme ist, dass der ROI als Maßzahl des Erfolgs die zentrale abhängige Variable ist und die unabhängigen Variablen beobachtbar und messbar sind (Schmid, 2012). Erfolgsfaktorenforschungen kommen zu dem Ergebnis, dass nur einige wenige Bestimmungsgrößen zur maßgeblichen Erklärung beitragen (Schmid, 2012). Das Forschungsprogramm PIMS – Profit Impact of Market Strategies – identifiziert acht zentrale branchenübergreifende Bestimmungsgrößen des ROI, die ca. 70-80 % der Rentabilitätsunterschiede erklären können (Schmid, 2012; vgl. zum PIMS-Forschungsprogramm Meffert, 1994). Neben dem relativen Marktanteil, der Wachstumsrate des Marktes, der Innovationsrate, der vertikalen Integration und dem Kundenprofil sind die relative Qualität der Produkte, die Produktivität und die Investmentintensität wichtige Erfolgsfaktoren (Waibel & Käppeli, 2006). Die Erfolgsfaktorenforschung ist bezüglich ihrer Möglichkeiten, die Ursache für Erfolg wissenschaftlich zu belegen, allerdings in der Kritik (Nicolai & Kieser, 2002).

Waibel und Käppeli (2006) zeigen in einem gesamtheitlichen Modell qualitative Ursache-Wirkungs-Zusammenhänge zwischen den Schlüsselfaktoren für den Erfolg. Dabei unterscheiden sie zwischen Führungskräften als Hebel, Indikatoren (Ziel- und Messgrößen) sowie den Rahmenbedingungen (Schmid, 2012). Basis für die Messung der Wirksamkeit der Maßnahme *Nachwuchsförderung* im Nachwuchsleistungszentrum sind drei der acht genannten gültigen branchenübergreifenden Erfolgsfaktoren für den ROI: die relative Qualität der Produkte, die Produktivität und die Investmentintensität. Die zur Ermittlung des ROI für Nachwuchsleistungszentren ausgewählten und konstruierten Variablen sollen die drei genannten Erfolgsfaktoren für den ROI repräsentieren. Den größten negativen Einfluss auf den ROI hat eine hohe Investmentintensität, die weitgehend davon abhängig ist, in welcher Branche man tätig ist; eine hohe Produktivität und eine hohe relative Qualität wirken auf alle Finanzdaten, insbesondere bei einem geringen Marktanteil (Waibel & Käppeli, 2006). Der Erfolg des Unterfangens – Konzeption und Anwendung eines innovativen Evaluationskonzepts zur Messung des betriebswirtschaftlichen Erfolgs von Nachwuchsleistungszentren im deutschen Profifußball – wird anhand der Erfolgskriterien, Validität der genutzten und konstruierten Daten, dem Grad der IT-Unterstützung (Digitalisierungsgrad) bei der Ermittlung der Daten und bei der Anwendung des Konzepts, dem Evaluationsaufwand für Vereine und Verbände, der Innovativität und dem Ergebnis einer Metaevaluation bewertet.

Abbildung 27: Mentales Modell – ROI für Nachwuchsleistungszentren.



Um das Ziel des Unternehmens *Nachwuchsleistungszentrum*, eine „qualitativ hohe Ausbildung talentierter Nachwuchsspieler in den verschiedenen Altersklassen“, zu gewährleisten (DFL, 2020c) erfordern die Vorgaben bezüglich der infrastrukturellen Voraussetzungen sowie weiteren strukturellen Bedingungen, wie z. B. Ausführungen zum Aufbau des Trainingsgeländes, die Personalausstattung, die Verträge der Spieler mit den Vereinen, ein Investment in das Nachwuchsleistungszentrum (vgl. dazu Anhang V zur Lizenzierungsordnung des DFL e.V., DFL, 2021c). Ein Investment beinhaltet die klare Erwartungshaltung an eine angemessene Rendite der unternehmerischen Tätigkeit, gemessen am Erfolg im Verhältnis zum eingesetzten Kapital (Phillips & Schirmer, 2008).

Die Bewertung der Programmeffizienz mittels Kosten-Nutzen-Analyse setzt den Nachweis der Programmwirksamkeit anhand der Ergebnisse nach Abschluss der Aktivitäten des Unternehmens (*outcome*-orientierte Evaluation) voraus (Mittag & Hager, 2000). Eine Evaluation der Programmkonzeption, wie die Ausarbeitung und Gestaltung der geplanten Maßnahmen bzw. der Konzeption des Programmes als solches, oder eine begleitende Evaluation der Programmdurchführung wird nicht vorgenommen. Es handelt sich bei dieser Untersuchung um eine sog. *Black-Box-Evaluation*, mit der lediglich die Ergebnisse der Maßnahmen im Nachwuchsleistungszentrum (Wirksamkeit) analysiert werden, eine nähere Betrachtung der Kausalprozesse erfolgt nicht (Döring & Bortz, 2016).

3.2 Planung der Evaluation

Der Zweck der Evaluation des betriebswirtschaftlichen Erfolgs von Nachwuchsleistungszentren ist es, die Kosten und den Nutzen der Aktivitäten von Nachwuchsleistungszentren zu vergleichen. Eine Reflexion des dieser Studie zugrundeliegende Evaluationskonzepts erfolgt im Sinne einer Qualitätssicherung anhand der eingangs aufgestellten Kriterien: *Outcome*-Orientierung, Nutzung öffentlich zugänglicher Daten, Höhe des Aufwands in der Anwendung, Verwendung betriebswirtschaftlicher Kennzahlen und individueller Leistungsindikatoren, Analyse im Zeitverlauf und Leichtigkeit der Anwendung des Konzepts (Döring & Bortz, 2016). Abbildung 28 zeigt einen Überblick über die zugrunde gelegten Bewertungskriterien. Die Metaevaluation soll Erkenntnisse darüber liefern, ob es Optimierungsbedarf gibt und inwieweit dieses Konzept für die Praxis der Evaluation von Nachwuchsleistungszentrum gut geeignet ist. Evaluationsforscher empfehlen die Metaevaluation von Anfang an in die Evaluationsforschung einzuplanen, sie sollte den Evaluationsprozess begleiten (Döring & Bortz, 2016 nach Stufflebeam, 2001 und Stufflebeam & Shinkfield, 2007). Metaevaluation kann auf unterschiedliche Weise vorgenommen werden: selbst- oder fremdgesteuert, durch interne oder externe Personen, als formative und summative Metaevaluation (Döring & Bortz, 2016).

Abbildung 28: Kriterien zur Planung der Evaluation des betriebswirtschaftlichen Erfolgs von Nachwuchsleistungszentren.



Gängige Praxis ist es, Evaluationsstandards heranzuziehen und anhand von empirischen Daten zu überprüfen (Empfehlungen dazu von der Deutschen Gesellschaft für Evaluationen e. V.). Dazu sind die für die Bewertung der Evaluationsarbeit anzulegenden Kriterien sowie die Dokumentation der relevanten Daten festzulegen (Döring & Bortz, 2016). Wie bereits ausgeführt, haben die Messung der Effektivität und der Effizienz einen hohen Stellenwert bei der Evaluation von Personalentwicklungsmaßnahmen. Die Effektivität bildet die Wirksamkeit ab, sie stellt das Ausmaß dar, in dem eine Maßnahme die angestrebten Ziele bzw. Wirkungen erreicht hat. Je nach Ziel der Evaluation wird die unmittelbare Leistung bzw. Produktivität einer Personalentwicklungsmaßnahme (Output) die Wirksamkeit für eine bestimmte Gruppe (*Outcome*) oder die langfristige Wirksamkeit auf überindividueller oder gesellschaftlicher Ebene (Impact) gemessen. Fokus der vorliegenden Studie ist die Evaluation der Effektivität der Nachwuchsarbeit für die Gruppe der Nachwuchsleistungszentren. Das zugrundeliegende mentale Modell verdeutlicht die *Outcome-Orientierung* dieser Studie.

Es können unterschiedliche Weisen, den Erfolg einer Maßnahme zu messen, angewendet werden. Zum einen kommen Methoden zum Einsatz, die sich auf subjektive Einschätzungen z. B. von Unternehmensangehörigen oder Marktteilnehmern stützen, zum anderen Methoden, die sich objektiver Daten bedienen (Roediger, 2011). In Abhängigkeit von der Forschungsrichtung werden unterschiedliche Konzepte zur Messung des Erfolgs zugrunde gelegt. Während z. B. die Finanzforschung hauptsächlich objektive Daten, wie Daten aus dem Rechnungswesen, nutzt, greift z. B. die Organisationsforschung primär auf subjektive Einschätzungen von Experten oder des Managements zurück, da objektive Daten in vielen Fällen kaum zur Verfügung stehen (Schoenberg, 2006). Im Gegensatz dazu kann für die vorliegende Studie zum überwiegenden Teil auf öffentlich verfügbare Daten zurückgegriffen werden. Die *Nutzung öffentlich zugänglicher Daten* soll ein höheres Maß an Objektivität der Messung im Gegensatz zur Subjektivität der Selbsteinschätzung der Nachwuchsleistungszentren sicherstellen (siehe zur bisherigen Evaluationspraxis zum Zwecke des Qualitätsmanagements des DFB/der DFL Kapitel 3.5.1.2).

An der Spitze von betriebswirtschaftlichen Kennzahlensystemen steht der Return on Investment (ROI). Unternehmen bzw. deren Investoren steuern über diese betriebswirtschaftliche Kennzahl ihre Investitionsentscheidungen mit dem Ziel einer Maximierung von Kapitalrenditen. Die Verwendung des ROI als Evaluationskriterium erfüllt damit die diesbezügliche Anforderung als Teil des zu entwickelnden Evaluationskonzept.

Dem Zweck der Evaluation folgend sind für die Evaluation die zu sammelnden Daten, die Sammlungsmethode zur Leistungserfassung, die Art der Analyse und der formale Evaluationsbericht zu beschreiben (Phillips & Schirmer, 2008). Die zu sammelnden Daten und die Methode zur Leistungserfassung werden im Laufe dieser Arbeit in Kapitel 3.5 – Durchführung der Datenerhebung – eingehend beschrieben, anschließend wird die Art der Analysen erläutert.

3.3 Forschungsdesign

Ausgehend von der Forschungsfrage „Wie kann der betriebswirtschaftliche Erfolg von Nachwuchsleistungszentren im deutschen Profifußball valide beschrieben werden?“ wird das Forschungsdesign der Studie festgelegt. Das Forschungsdesign ist grundlegend für die Studie und wesentlich für die Validität der erzielten Ergebnisse, da es die Umsetzung der theoretischen Forschungsfrage in eine empirische Studie bestimmt (P. Stein, 2014). Die folgende Beschreibung des Forschungsdesigns orientiert sich an den neun Klassifikationskriterien für Untersuchungsdesigns nach Döring und Bortz (2016). Einen Überblick über die gewählten Varianten für diese Studie zeigt die Abbildung 29 (siehe blau markierte Begrifflichkeiten).

Im weiteren Verlauf dieses Kapitels wird die Wahl der Varianten des Untersuchungsdesigns dieser Studie näher erläutert.

3.3.1 Wissenschaftstheoretischer Ansatz der Studie

Für die wissenschaftliche Untersuchung und das zugrundeliegende Forschungsdesign muss eine Zuordnung zu einem wissenschaftsorientierten Paradigma erfolgen (Döring & Bortz, 2016). Das Forschungsdesign dieser Studie basiert im Wesentlichen auf einer *quantitativen Studie*, mit der die monetäre Wirksamkeit der Personalentwicklungsmaßnahmen in Nachwuchsleistungszentren und deren Entwicklung im Zeitverlauf analysiert werden. Die Ausprägungen der Variablen des Modells werden erhoben und die Messwerte statistisch ausgewertet. Über standardisierte Methoden soll die Forschungsfrage beantwortet (Perkhofer, 2016) und aus der daraus resultierenden wissenschaftlichen Aussage soll eine allgemeingültige Aussage getroffen werden.

Abbildung 29: Überblick über das Forschungsdesigns der vorliegenden Studie. Eigene Darstellung nach Döring und Bortz (2016).

Kennzeichen des Forschungsdesigns		Varianten des Forschungsdesigns	
1	Wissenschaftstheoretischer Ansatz der Studie	Quantitativ	
		Qualitativ	
		Mixed-Methods	
2	Erkenntnisziel der Studie	Grundlagen- wissenschaftlich	
		Anwendungs- wissenschaftlich	unabhängig Auftrag
3	Gegenstand der Studie	Empirisch	Originalstudie Replikationsstudie
		Methoden	
		Theorie	Review / Forschungsüberblick Metaanalyse
4	Datengrundlage bei empirischen Studien	Primäranalyse	
		Sekundäranalyse	
		Metaanalyse	
5	Erkenntnisinteresse bei empirischen Studien	Explorativ	
		Deskriptiv	
		Explanativ	
6	Bildung und Behandlung von Untersuchungsgruppen bei explanativen Studien	Experimentell	
		Quasi-Experimentell	
		Nicht-Experimentell	
7	Untersuchungsort bei empirischen Studien	Labor	
		Feld	
8	Anzahl der Untersuchungszeitpunkte bei empirischen Studien	(Quasi-) Experimentell	mit Messwiederholungen
			ohne Messwiederholungen
		Nicht-Experimentell	Querschnitt
			Trend Längsschnitt
9	Anzahl der Untersuchungsobjekte bei empirischen Studien	Gruppenstudie	Stichprobe Vollerhebung
		Einzelfallstudie	

3.3.2 Erkenntnisziel und Gegenstand der Studie

Das Erkenntnisziel dieser Studie ist es nicht, Grundlagenforschung zu betreiben, indem wissenschaftliche Probleme gelöst werden. Es wird ein *Nutzen in der Praxis* der Evaluation des monetären Erfolgs von Personalentwicklungsmaßnahmen angestrebt. Die anwendungswissenschaftliche Studie zielt mittels wissenschaftlicher Methoden und unter Einbeziehung der theoretischen Hintergründe auf die Verbesserung der Evaluationsmethode zur Messung des Erfolgs von Nachwuchsleistungszentren ab. Die Forschung im Rahmen dieser Studie ist *unabhängig* von einem privatwirtschaftlichen oder öffentlich-rechtlichen Mittelgeber. Es handelt sich um eine *Originalstudie mit quantitativem Design*, die zur Lösung der Beantwortung der Forschungsfrage beiträgt.

3.3.3 Datengrundlage der empirischen Studie

Empirische Studien zeichnen sich in der Regel dadurch aus, dass eine eigene Datenerhebung erfolgt und die Analyse auf diesen Daten basiert (Döring & Bortz, 2016). Um neue wissenschaftliche Erkenntnisse zu gewinnen, sind Daten, die eigenständig zu Forschungszwecken erhoben werden, die zentrale Quelle hinsichtlich Forschungen. Qualitative wie quantitative Studien werden als Primäranalyse durchgeführt, d. h., es werden selbst erhobene Daten erstmalig ausgewertet (Döring & Bortz, 2016). Einerseits ist es ein Vorteil dieser Art der Analyse, dass Details des Forschungsdesigns, die Art der Stichprobe, die Datenerhebung und der Datensatz passgenau auf die konkrete Forschungsfrage zugeschnitten werden können (Döring & Bortz, 2016). Andererseits ist zu bedenken, dass jede Form der eigenen Datenerhebung „teuer, zeitaufwändig und personalintensiv“ ist (Behnke, 2009). Planung und Durchführung der Erhebung empirischer Daten machen in Forschungsprojekten einen großen Anteil der kalkulierten Ressourcen (Zeit, Geld und Personal) aus (Tausendpfund, 2018).

Aus forschungsökonomischer Sicht können nur kleine Datensammlungen erzeugt werden (Döring & Bortz, 2016). Um die eigene Forschungsfrage beantworten zu können, ist es deshalb nicht unbedingt sinnvoll oder erforderlich, eine eigene Datenerhebung durchzuführen. Stattdessen kann auf bereits existierende Daten zugegriffen werden. Solche Studien werden als Sekundäranalysen bezeichnet (Helfrich, 2016). Eine zentrale Voraussetzung für die Sekundäranalyse ist, dass relevante Daten verfügbar sind und dass alle mit dem Datenschutz verbundenen Fragen eine angemessene Berücksichtigung finden. Die Auswertung der Daten erfolgt unter einer neuen Forschungsperspektive (Döring & Bortz, 2016). In Sekundäranalysen lassen sich die vorliegenden Daten mit der Erhebung und Generierung neuer Daten verbinden (Heaton, 2008). Für die Wahl von Sekundäranalysen gilt u. a. der Zugang zu Daten, die sonst schwer zu erheben wären, wie z. B. historische Daten im Zeitverlauf. Darüber hinaus sind empirische Untersuchungen in Bezug auf das Datenmaterial immer häufiger mit Zugangsproblemen und steigenden Aufwendungen verbunden, da Betriebe, Unternehmen, Verbände etc. sich seit einiger Zeit mit einer steigenden wissenschaftlichen Datennachfrage konfrontiert sehen und darauf mit einer sinkenden Teilnahmebereitschaft an Befragungen reagieren (Gebel & Rosenbohm, 2017; Pflüger et al., 2017). Eine geringe Anzahl an Untersuchungseinheiten erhöht das Risiko, wegen der hohen Erhebungsdichte zu einer zunehmenden Schließung des Forschungsfeldes zu kommen. Eine Erhebung von Primärdaten ist in diesem Fall nicht effektiv und effizient.

Da im Feld dieser deskriptiven Studie zu den Nachwuchsleistungszentren im Betrachtungszeitraum lediglich eine beschränkte Anzahl von Untersuchungseinheiten ($N=54$) existiert, die darüber hinaus bereits mehrfach Befragungen und Audits unterzogen worden sind, erscheint es sinnvoll, mittels Sekundäranalyse relevante bereits vorhandene Daten zu nutzen. Der Rückgriff auf bereits vorhandene Daten reduziert den *Aufwand* für den Forschungsprozess erheblich, da keine aufwändigen, neuen Erhebungen erforderlich sind (Tausendpfund, 2018). Bei der vorliegenden Studie handelt es sich daher im

Wesentlichen um eine Sekundärdatenanalyse. Es werden Daten aus unternehmensexternen Quellen verwendet, die öffentlich zugänglich sind, von *Transfermarkt.de* regelmäßig veröffentlicht werden oder für diese Studie zu Forschungszwecken von DFB/DFL zur Verfügung gestellt wurden.

Ein genereller Nachteil einer Sekundäranalyse kommt aufgrund der Verwendung oben genannter Daten für diese Studie zum Tragen: Die zur Verfügung stehenden Datensätze enthalten nicht alle relevanten Variablen und Aspekte, die laut Modell zur Berechnung des betriebswirtschaftlichen Erfolgs erforderlich sind (Döring & Bortz, 2016). Die Daten für diese fehlenden, aber wesentlichen erforderlichen Variablen für das zugrundeliegende Berechnungsmodell werden selbst erhoben und anschließend analysiert (Primäranalyse).

Die vorliegende Studie wird somit als *Primär- und Sekundäranalyse* durchgeführt. Untersuchungsobjekte dieser Studie sind die im Betrachtungszeitraum der Jahre 2009 bis 2019 existierenden Nachwuchsleistungszentren des deutschen Profifußballs (N=54).

3.3.4 Erkenntnisinteresse der Studie

Diese empirische *deskriptive* Studie zielt darauf ab, Aussagen zu der Ausprägung des Merkmals *monetärer Erfolg* der Gesamtpopulation der Nachwuchsleistungszentren (N=54) im deutschen Profifußball zu treffen. Sämtliche Populationsmitglieder werden bezüglich ausgewählter betriebswirtschaftlicher Kennzahlen untersucht. Diese Vollerhebung ist möglich, da es sich mit 54 Untersuchungsobjekten um eine überschaubare und gut erreichbare Population handelt (Döring & Bortz, 2016; P. Stein, 2014). Es wird keine reale Versuchssituation hergestellt (nicht-experimentell), sondern auf Basis der im Wesentlichen bereits vorliegenden Daten eine Ex-Post-Facto-Analyse in Form von Korrelations- bzw. Regressionsanalysen vorgenommen (Döring & Bortz, 2016; Helfrich, 2016).

3.3.5 Anzahl der Untersuchungszeitpunkte und -objekte der Studie

Die Studie wird als eine *retrospektive Längsschnittstudie* durchgeführt, in zeitlichem Abstand erfolgt mehrfach eine Erhebung der Daten für den Betrachtungszeitraum 2009 bis 2019, d. h. mit 11 Untersuchungszeitpunkten. Die für diese Jahre generierten Daten werden miteinander verglichen. Die Studie über diesen Zeitverlauf wird ermöglicht durch die Nutzung der öffentlich zugänglichen, bereits vorhandenen Daten für den Betrachtungszeitraum. Die zu erhebenden Primärdaten werden retrospektiv für den relevanten Zeitraum generiert. Paneleffekte, d. h. Veränderungen der Anzahl der Untersuchungsobjekte an der Erhebung, sind ausgeschlossen, da die Daten regelmäßig pro Jahr erhoben werden (P. Stein, 2014). Derselbe Untersuchungsgegenstand – das Nachwuchsleistungszentrum – wird zu unterschiedlichen Zeitpunkten evaluiert, was es ermöglicht, Veränderungen über die Zeit festzustellen und einen (Wirksamkeits-)Zusammenhang der Variablen abzuleiten (Witzel, 2010). Eine Zusammenstellung der Merkmale des Forschungsdesigns zeigt Abbildung 30.

Abbildung 30: Überblick Merkmale des Forschungsdesigns

Forschungs- design	quantitativ	Anwendungs- wissenschaft- lich	Empirisch Original	Primär- und Sekundär- analyse
	Deskriptiv	Feld	Nicht- Experimentell (Längsschnitt)	Vollerhebung

An die Beschreibung des mentalen Modells, der Planung der Evaluation und des Forschungsdesigns schließt sich im Folgenden die Beschreibung der Operationalisierung – der Auswahl von Indikatoren und Datenerhebungsinstrumenten für die Überführung der spezifizierten theoretischen Konzepte in messbare Variablen – an (Döring & Bortz, 2016; Burzan, 2014).

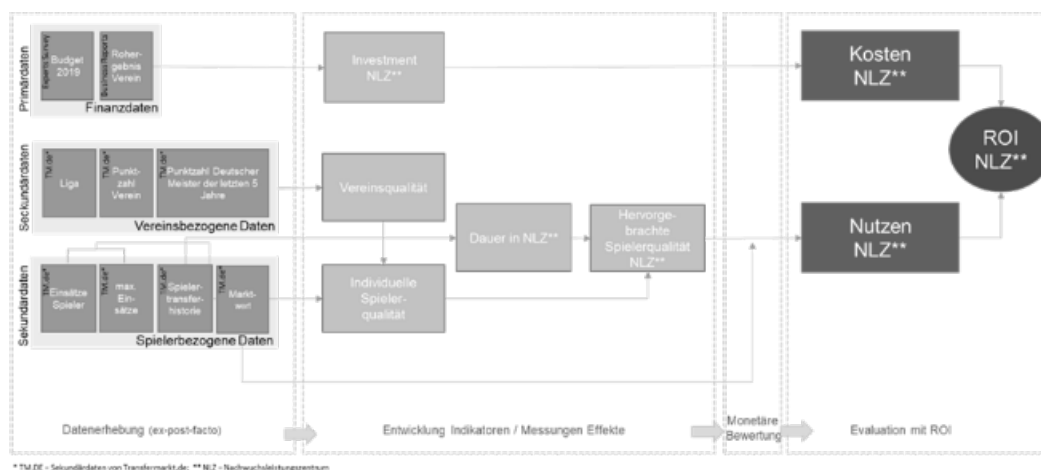
3.4 Operationalisierung

In der vorliegenden quantitativen empirischen Studie werden für alle Nachwuchsleistungszentren (Evaluationsobjekte) der Stichprobe Messwerte erhoben, in Relation gesetzt und mittels statistischer Methoden analysiert, um Rückschlüsse auf die Messung des betriebswirtschaftlichen Erfolgs von Nachwuchsleistungszentren und ihre Investitionsentscheidungen ziehen zu können. Die Messung des Erfolgs über Investitionsrechnungsverfahren kann einerseits z. B. anhand objektiver Erfolgskennzahlen aus Bilanzen bzw. dem Controlling eines Unternehmens erfolgen, andererseits kann Erfolg auch qualitativ gemessen werden (J. Littkemann et al., 2012). Empirische Untersuchungen, z. B. von Götz (2001), zeigen, dass trotz der hohen Investitionssummen in der Branche des Profifußballs, also auch in Nachwuchsleistungszentren, Investitionsentscheidungen häufig nicht anhand von objektiven langfristig und rentabilitätsorientierten Berechnungen, sondern eher intuitiv getroffen werden. Mit der statischen Rentabilitätsrechnung mittels des ROI wird eine Objektivität geschaffen, die auf monetäre Wirkungen fokussiert, indem sie die mit einer Investition verbundenen Zahlungsströme (Ein- und Auszahlungen) als relevant erachtet und es ermöglicht, Investitionsentscheidungen zu Nachwuchsleistungszentren zur Kontrolle zu beurteilen (Götz, 2001). Zielsetzung der Messung des monetären Erfolgs mittels des ROI ist es, die Rentabilität der Investition in ein Nachwuchsleistungszentrum zu bewerten. In dieser Studie wird der ROI als relative Erfolgsgröße von Nachwuchsleistungszentren als zentrale abhängige Variable definiert und es wird die Annahme getroffen, dass die unabhängigen Variablen – *Kosten und Nutzen* – sowie die zu ihrer Berechnung erforderlichen Komponenten messbar sind (zugrundeliegendes mentales Modell: siehe Abbildung 27).

Die *Kosten* der Nachwuchsleistungszentren der Vereine pro Jahr liegen nicht vor. Analysen öffentlich-zugänglicher Daten aus DFL-Wirtschaftsreports, Bilanzen sowie Schätzungen der für die Nachwuchsleistungszentren zur Verfügung stehenden Budgets im Jahre 2019 haben zu der Erkenntnis geführt, dass Kosten hinreichend über das *Investment*

in Nachwuchsleistungszentren operationalisiert werden können. Das Investment wird als *prozentualer Anteil des Rohergebnisses des Vereins* mit einem angeschlossenen Nachwuchsleistungszentrum ermittelt. Die Operationalisierung des *Nutzens* des Ausbildungsprogramms in einem Nachwuchsleistungszentrum erfolgt über die finanzielle Bewertung des Indikators *Hervorgebrachte Spielerqualität NLZ* (=Produktivität) mit Marktwerten. Die Hervorgebrachte Spielerqualität NLZ ist die Summe der Qualitäten aller Spieler, die mindestens eine Halbsaison in dem spezifischen Nachwuchsleistungszentrum ausgebildet worden sind. Die *Individuelle Spielerqualität* (=Relative Qualität des Produkts) wiederum basiert auf der Berechnung eines Vereinsqualitäts-Indikators – der wiederum auf Indikatoren zur Ligaqualität, der Punktzahl des Vereins und der Punktzahl des deutschen Meisters der letzten 5 Jahre basiert –, den maximalen Einsätzen der Spieler und der Anzahl der Einsätze des Spielers. Abbildung 31 zeigt, basierend auf dem vorgestellten mentalen Modell, die für die Evaluation des monetären Erfolgs von Nachwuchsleistungszentren ausgewählten Indikatoren und Datenerhebungsinstrumente zur Überführung der spezifizierten theoretischen Konzepte in messbare Variablen.

Abbildung 31: Überblick Evaluationsprozess und Operationalisierung ROI für Nachwuchsleistungszentren.

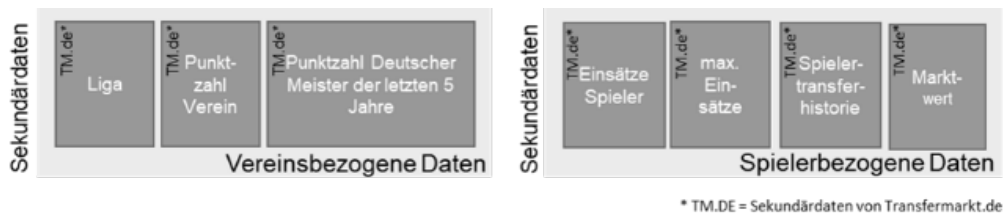


3.5 Durchführung der Datenerhebung

Zwar können mittels theoretischer Argumentation und unter Verwendung bestehender empirischer Ergebnisse logische Erkenntnisse über die Messung des Erfolgs von Nachwuchsleistungszentren gewonnen werden, aber erst die Durchführung einer empirischen Untersuchung erlaubt die Abbildung der Realität und die Generierung eines Erkenntnisfortschritts, aus dem Schlussfolgerungen für die Praxis formuliert werden können (Kromrey, 2002). Eine empirische wissenschaftliche Untersuchung benötigt Daten als Untersuchungsbasis, sie dienen der Überprüfung wissenschaftlicher Aussagen an der Realität (Helfrich, 2016). Die „Daten sind nicht unmittelbar gegeben, sondern müssen erst konstruiert werden. Die Konstruktion besteht in der Datenerhebung und der Datenaufbereitung“ (Helfrich, 2016, S. 152). Basierend auf dem entwickelten Modell der

Berechnung des monetären Erfolgs von Nachwuchsleistungszentren, den dort abgebildeten Zusammenhängen und Wirkungen der einzelnen Komponenten (siehe Abbildung 32), erfolgt zunächst die Auswahl und Erhebung relevanter Daten. Darüber hinaus orientiert sich die Datenerhebung am Untersuchungsziel der vorliegenden Arbeit und an den in der Forschung bereits existierenden Erkenntnissen. In der vorliegenden Studie wird überwiegend auf bereits vorhandene Daten der Internetplattform Transfermarkt.de (www.transfermarkt.de) zurückgegriffen (Sekundärdaten-Analyse). Diese werden im Hinblick auf die Beantwortung der Forschungsfrage untersucht.

Abbildung 32: Überblick Datenherkunft und Kategorisierung



Da nicht alle zu untersuchenden Variablen und Aspekte, die für die vorliegende Forschungsfrage relevant sind, in dem vorgefundenen öffentlich zugänglichen Datensatz enthalten sind, werden die für die Bewertung des monetären Erfolgs von Nachwuchsleistungszentren fehlenden Daten gesondert erhoben und konstruiert (Döring & Bortz, 2016). Der Erfolg von Nachwuchsleistungszentren wird mittels der Berechnung des Return on Investment (ROI) ergebnisorientiert gemessen (Phillips & Schirmer, 2008). In die Messung des ROI gehen *Kosten* und *Nutzen* ein, die sich gemäß des dieser Studie zugrundeliegenden Modells aus dem Konstrukt *Investment* und dem monetär bewerteten Konstrukt *Hervorgebrachte Spielerqualität NLZ* ergeben.

Das Investment in Nachwuchsleistungszentren wird als *Anteil am Rohergebnis Verein mit NLZ* ermittelt. Der für die Population der Nachwuchsleistungszentren anzunehmende Anteil am Rohergebnis Verein wird in einem ersten Schritt auf Basis von Expertenschätzungen des Budgets 2019, im zweiten Schritt für die Jahre des Betrachtungszeitraums ermittelt. Der Nutzen aus der Tätigkeit der Nachwuchsleistungszentren wird über mehrere Schritte unter Berücksichtigung der Vereinsqualität, des Individual-Leistungsindicators Spielerqualität, der Summe aller mittleren Qualitäten der hervorgebrachten Spieler pro NLZ und der monetären Bewertung des Konstrukts Hervorgebrachte Spielerqualität NLZ ermittelt.

Eine Validierung des Indicators *Hervorgebrachte Spielerqualität NLZ* erfolgt anhand von weiteren Sekundärdaten, und zwar den Daten aus bisherigen Evaluationen im Rahmen der Zertifizierungen der Nachwuchsleistungszentren durch *Double PASS*. Double Pass ist ein belgischer Serviceanbieter, der 2004 als Spin-off-Unternehmen der Universität Brüssel (VUB) gegründet wurde. Das Unternehmen berät Vereine, Verbände und Ligen bei der Optimierung ihrer Jugendakademien und ihres Spielerpotentials (Double Pass,

2021). Das *PASS-System* basiert auf den wissenschaftlichen Arbeiten der (späteren) Gründer des Unternehmens zum Qualitätsmanagement im Sport und Messung des Erfolgs (Knop et al., 2004; Bosscher et al., 2006; Hoecke van et al., 2006).

Im Folgenden werden, basierend auf dem mentalen Modell der Berechnung des monetären Erfolgs von Nachwuchsleistungszentren und den dort abgebildeten Zusammenhängen und Wirkungen der einzelnen Komponenten, die in die Konstrukte des Modells eingehenden Daten (Sekundärdaten, Primärdaten) bezüglich ihrer Kategorisierung, Herkunft und Güte näher beschrieben.

3.5.1 Erhebung der Sekundärdaten

Die für die Evaluation des monetären Erfolgs als relevant erachteten öffentlich verfügbaren Daten (siehe Abbildung 32) werden über die Internetplattform *Transfermarkt.de* erhoben. Im Folgenden wird daher zunächst die Internetplattform *Transfermarkt.de* mit ihrer Rolle in der Sportberichterstattung und der Praxis der Fußballbranche erläutert. Des Weiteren werden die verwendeten Daten bzgl. ihrer Relevanz und Qualität beschrieben. Da in der Datenanalyse in den letzten Jahren zunehmend auf die Qualität der Datenerhebung und den Zustand der Daten geachtet wird, wird darüber hinaus erläutert, inwieweit die Qualität und Güte der erhobenen Daten kritischen Überprüfungen in Bezug auf wissenschaftliche Analysen standhalten kann (Mochmann, 2014; RatSWD, 2014).

3.5.1.1 Daten von *Transfermarkt.de*

Für die Entwicklung des Indikators *Hervorgebrachte Spielerqualität NLZ* ist vor dem Hintergrund der Effizienzorientierung des Evaluationsansatzes analog zu Wallebohr und Daumann (2019) eine Datenbasis gewählt worden, zu der ein niedrigschwelliger Zugang besteht: die Internetplattform *Transfermarkt.de*. *Transfermarkt.de* als Datenquelle zeichnet sich dadurch aus, dass die Daten öffentlich leicht zugänglich sind und individuelle Leistungsdaten der in die Betrachtung einbezogenen Spieler verfügbar sind. Die Internetplattform *Transfermarkt.de* betreibt sportjournalistische Berichterstattung und die Pflege umfassender Daten zu Spielern, Vereinen und Ligen im internationalen Fußballgeschäft. *Transfermarkt.de* ist eine deutsche Website zum Thema Sport.

Sie wird betrieben von der *Transfermarkt GmbH & Co. KG* mit Sitz in Hamburg. Im Februar 2021 war die Website, nach dem Onlineauftritt vom Sportmagazin *Kicker*, das größte Internetportal in Deutschland mit dem Schwerpunkt Fußball (agof, 2021). Laut der Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e. V. (IVW) ist *Transfermarkt.de* mit etwa 45,4 Mio. Besuchern einer der am häufigsten besuchten Internetauftritte in der Kategorie Sport in deutscher Sprache (IVW, 2021). Auch Bundesligavereine stützen sich in ihrer offiziellen Berichterstattung auf die von der Plattform publizierten Daten, z. B. TSG 1899 Hoffenheim in ihrem Jahresabschluss (Briel & Lehmann, 2017).

Es werden objektive und intersubjektiv nachprüfbare Statistiken mit persönlichen und sportlichen Daten der Spieler, z. B. Torquoten, Spielminuten, Verletzungen sowie die Transferhistorie, veröffentlicht. Für die Erfassung, Analyse und Vermittlung von detaillierten Sportdaten wird der Großteil der Daten zu Leistungen, Aufstellungen, Auswechslungen, Gelbe Karten (im folgenden „Leistungsdaten“ genannt) etc. von dem Sportdatendienstleister *Opta* geliefert. Das Unternehmen ist seit 20 Jahren am Markt und vertreibt umfassende Leistungsdaten des deutschen Profifußballs zu allen Klubs und Spielern (Optasports, 2020).

Mit Informationen zu Spielern, Vereinen und Wettbewerben sowie als eine der größten Fußball-Communities und eines Forums für alle, die sich zu Fußball mitteilen und austauschen wollen, bietet Transfermarkt.de die weltweit größte frei zugängliche Spielerdatenbank mit mehr als 800.000 Spielern, Vereinen und Trainern (Battis, U., Ingold, A. & Kuhnert, K.; Media Impact, 2019; Media Impact, 2021). Mit der außergewöhnlichen Vorgehensweise der Beteiligung einer User-Community an der Bereitstellung der Informationen, hat sich Transfermarkt.de in Deutschland einen Namen gemacht (Heibel, 2012). Neben den aufgeführten Leistungsdaten gehört die Rubrik *Transfers & Gerüchte* mit Beiträgen (News, Videos) zu möglichen Transfers, mit Beraterangaben und Einblicken von Fans zu den zentralen Angeboten. Daten, die nicht von Opta eingeliefert werden, wie z. B. Trikotnummern, Marktwerte, Positionen, Geburtsdaten usw., werden durch die Community eingebracht sowie ergänzt und korrigiert. Korrekturen und Ergänzungen der Community unterliegen einem durch die Community und sogenannten Experten von Transfermarkt.de aufgestellten Regelwerk, dessen Einhaltung durch diese selbst kontrolliert wird (Transfermarkt.de, 2021b).

Das Portal ist mit den Foren Top-Foren, internationale und nationale Foren sowie Sonderforen in vier Hauptkategorien gegliedert. In den Top-Foren tauschen sich die User insbesondere über Gerüchte des Fußballs bzgl. Spielern, Trainern und Vereinen – auf nationaler wie auch auf internationaler Ebene – aus. Teil der Top-Foren ist die *Marktwertanalyse*, deren Werte lt. Angaben von Media Impact (2019) mit Abstand die meistzitierte Quelle bezüglich Marktwerten von Fußballspielern ist.

Wesentlicher Bestandteil des Modells zur Messung des Erfolgs von Nachwuchsleistungszentren ist der Indikator *Hervorgebrachte Spielerqualität NLZ*. In diese entwickelte Kennzahl gehen spieler- und vereinsbezogene Daten aus Transfermarkt.de ein. Darüber hinaus erfolgt eine monetäre Bewertung der *Hervorgebrachten Spielerqualität NLZ* mittels des Marktwertes von Transfermarkt.de (Lanwehr et al., 2021). Es werden daher zunächst die spieler- und vereinsbezogenen Daten in den Kontext gesetzt und Güte und Verwendbarkeit dieser Daten von Transfermarkt.de für diese Studie erläutert. Im Anschluss daran wird der Prozess der Schätzung des Marktwertes von Transfermarkt.de beschrieben und die Glaubwürdigkeit und Verwendbarkeit des Marktwertes für diese Studie eingeschätzt.

3.5.1.1.1 Spieler- und vereinsbezogene Daten

In die Kategorie *Spielerbezogene Daten* gehen die Einsätze Spieler, die maximalen Einsätze, die Transferhistorie Spieler und der Marktwert ein. In die Kategorie *Vereinsbezogene Daten* gehen die Liga, die Punktzahl des Vereins sowie die Punktzahl des deutschen Meisters der letzten 5 Jahre ein. Spieler, die im Alter von 13 bis 19 Jahren mindestens eine Halbsaison in einem der betrachteten Nachwuchsleistungszentren der deutschen Profifußballligen ausgebildet wurden und die in der Folge mindestens ein Spiel als Profifußballspieler in einer der erhobenen Ligen absolviert haben (n=630), sind in die Untersuchung eingegangen. Erfasst wurden je Spieler die persönlichen Daten (*Name, Vorname, Geburtsdatum*), *Einsätze* in der betrachteten Saison, *Transfers* von einer Mannschaft (im folgenden Verein genannt) zur anderen mit Datum und dem *Marktwert*. Zur eindeutigen Identifikation der spielerbezogenen Daten wurde je Spieler eine eindeutige Spieler-Identifikationsnummer (*SpielerID*) vergeben. Die Anzahl der Einsätze des Spielers pro Saison bedingen laut Berechnungsmodell die Qualität eines Spielers. Isoliert betrachtet lassen sich aber die Kennzahlen eines Spielers (z. B. Torquote und Anzahl Einsätze) nicht sinnvoll interpretieren, diese können nur im Kontext der Vereinsqualität bewertet werden (Lanwehr et al., 2021).

Vereinsbezogene Daten zur Ermittlung der Vereinsqualität werden ebenfalls Transfermarkt.de entnommen. Relevant für die Vereinsqualität sind laut Modell die *Ligazugehörigkeit* des Vereins im Betrachtungsjahr, die erreichte *Punktzahl des Vereins* im Betrachtungsjahr und die *Punktzahl des deutschen Meisters der letzten 5 Jahre* ausgehend vom Betrachtungsjahr. In die Untersuchung sind alle Vereine und Betrachtungsjahre eingegangen, in denen ein Spieler der o. g. Stichprobe mindestens ein Spiel als Profifußballspieler absolviert hat (n=108319); darüber hinaus die für die Ermittlung des Regulativs *Punktzahl des deutschen Meisters der letzten 5 Jahre ausgehend vom Betrachtungsjahr* erforderlichen Vereine, die deutscher Meister geworden sind. Die vereinsbezogenen Daten sind über die vergebenen eindeutigen Vereins-Identifikationsnummern (VereinsID) in der selbstgestellten relationalen Datenbank abgelegt.

3.5.1.1.2 Datenqualität der spieler- und vereinsbezogenen Daten von Transfermarkt.de

Es ist in der Wissenschaft zu verhindern, dass mangelnde Datenqualität zu einer Behinderung des Erkenntnisgewinns führt (Döring & Bortz, 2016). Datenqualität ist definiert als „Mehrdimensionales Maß für die Eignung von Daten, den an ihre Erfassung/Generierung gebundenen Zweck zu erfüllen. Diese Eignung kann sich über die Zeit ändern, wenn sich die Bedürfnisse ändern“ (Würthele, 2003, S. 21). Laut Apel et al. (2015) geht es bei der Betrachtung der Eigenschaften oder Merkmale der Daten immer um die im Kontext stehenden Voraussetzungen oder den Zweck. Eine geringe Qualität von Daten, die in die Berechnung betriebswirtschaftlicher Kennzahlen eingehen, kann zu hohen Folgekosten führen, wenn auf der Basis dieser Kennzahlen falsche unternehmerische Entscheidungen getroffen werden (Döring & Bortz, 2016; Klier, 2008). Auch für den Erfolg von Unternehmen sind Informations- und Datenqualität ein zentraler Bestandteil (Hildebrand et al., 2015).

Transfermarkt.de zählt zu den am häufigsten besuchten Internetseiten Deutschlands und ist eines der größten deutschen Internetportale mit sehr vielen Daten rund um den Fußball (Media Impact, 2021). Für die Transfermarkt GmbH & Co. KG stellt eine hohe Datenqualität eine wertvolle strategische Ressource dar. Die Verbesserung und dauerhafte Sicherung der Datenqualität der bereitgestellten Informationen dient der Erhaltung und Verbesserung der Kundenzufriedenheit und damit einträglichen Werbeeinnahmen, die langfristig den wirtschaftlichen Erfolg ermöglichen (Otto & Österle, 2016).

Die Verwendbarkeit der von Transfermarkt.de veröffentlichten Daten für die vorliegende wissenschaftliche empirische Studie ist abhängig von einem Maß der Qualität der Daten. Transfermarkt.de zeichnet sich als offene Plattform aus, die von interessierten Usern gepflegt wird. Das Erfolgsprinzip des deutschsprachigen Fußballportals ist der Einbezug sogenannter Paten und Datenpfleger, die die frei zugängliche Spielerdatenbank mit weiteren Inhalten füllen, Datenbestände korrigieren, pflegen und Beiträge verfassen. Lorenz et al. (2011) betrachten das Vorgehen nach dem Prinzip der *Weisheit der Vielen* (vgl. dazu Surowiecki, 2005) bei Transfermarkt.de kritisch. Die Autoren geben an, dass sich bereits milde soziale Einflüsse nachweislich auf den Effekt der Weisheit der Vielen auswirken können. Es käme zu einem sozialen Einflusseffekt (social influence effect), der die Meinungsvielfalt verringert, ohne kollektive Fehler zu verbessern. Ebenso käme es zu Bereichsreduzierungseffekten (range reduction effect), die zur Verschiebung der Wahrheit in Randbereiche führen, sodass externen Beobachtern weniger zuverlässiges Fachwissen bereitgestellt würde. Darüber hinaus sei ein Vertrauenseffekt (confidence effect) zu verzeichnen, der trotz mangelnder Genauigkeit zu einem individuellen Glauben an die Schätzungen führt (Lorenz et al., 2011).

Die Datenbank von Transfermarkt.de findet trotz dieses Hinterfragens der Glaubwürdigkeit der Daten mittlerweile in vielen wissenschaftlichen Arbeiten Verwendung. So haben Bryson et al. (2013) anhand von statistischen Analysen auf Basis der Transfermarkt.de-Daten nachgewiesen, dass Fußballspieler, die einen beidfüßig starken Fuß haben, in der Regel mehr Prämien bekommen als Spieler, die nur einen spielstarken Fuß haben. Heuer (2012) gelang es mit Hilfe systemischer quantitativer Analysen der Daten von Transfermarkt.de, ein Modell zur Vorhersage des Ergebnisses von Fußballspielen zu entwickeln. In einer quantitativen Inhaltsanalyse konnte aufgezeigt werden, dass die Eintrittswahrscheinlichkeit von Transfers auf Basis von Fanprognosen höher ist als die Eintrittswahrscheinlichkeit auf Basis der Vorhersagen von Sportjournalisten (Vögele, 2018). Auch Peeters (2018) konnte belegen, dass Fußballergebnisse, die auf Zuschauerbewertungen basieren, genauer sind als diejenigen, die sich auf Standardprädiktoren wie der FIFA-Rangliste stützen. Zur Überprüfung ihrer Annahmen griffen die genannten Autoren auf die Daten von Transfermarkt.de zurück. Diese erfolgreichen wissenschaftlichen Untersuchungen unter Verwendung der Datenquelle Transfermarkt.de lassen auf die Vertrauenswürdigkeit der Transfermarkt.de-Datenbank schließen und bekräftigen deren Validität. Datenqualität ist aber ein „mehrdimensionales, kontextabhängiges Konzept“ (Otto & Österle, 2016, S. 31 nach R. Y. Wang & Strong, 1996), das nicht durch ein einziges Merkmal ausreichend beschrieben werden kann. Die Qualität von Daten

wird in ihrer Gesamtheit über verschiedene Dimensionen beschrieben. Insbesondere sind hier die Vollständigkeit, Fehlerfreiheit, Konsistenz und Aktualität der Daten zu nennen (Heinrich & Klier, 2015; Otto & Österle, 2016).

Über die Beurteilung der zur Glaubwürdigkeit der Daten von Transfermarkt.de auf Basis der bereits vorliegenden Studien hinaus, wird das Maß der Qualität der Daten von Transfermarkt.de und ihre Verwendbarkeit für die vorliegende Untersuchung anhand der Qualitätsdimension der Fehlerfreiheit überprüft (Kasperek, 2020; Strauß, 2020). Fehlerfreiheit wird damit beschrieben, dass Ausprägungen in einer modellierten Realwelt-Entität den Attributwerten in einem Informationssystem entsprechen (Heinrich & Klier, 2015). Batini und Scannapieca (2006) definieren die Fehlerfreiheit als die korrekte Darstellung eines Wertes und die Übereinstimmung eines Wertes mit einem Wert in der Realwelt. Nach Heinrich und Klier (2015) können zwei verschiedene Ansätze zur Überprüfung vorhandener Daten auf Fehlerfreiheit angewandt werden: den *Alles-oder-nichts-Ansatz* und den *Toleranz-Ansatz*. Unter Anwendung des Alles-oder-nichts-Ansatzes wird geprüft, ob Attribute dem realen Sachverhalt entsprechen oder nicht, bei der Anwendung des Toleranz-Ansatzes werden vermeintliche Fehler differenzierter betrachtet, indem zwischen „völlig falsch“ (Würthele, 2003, S. 112) und „teilweise falsch“ (Würthele, 2003, S. 112) unterschieden wird.

Für die Studien ist die Annahme getroffen worden, dass die Daten von Transfermarkt.de vertrauenswürdig sind und in die Evaluation des Erfolgs von Nachwuchsleistungszentren eingehen können. Die Prüfung der Datenqualität anhand eines Abgleichs der für diese Studie relevanten Daten von Transfermarkt.de und der Daten polnischer Fußballstatistik-Webseiten (90minut.pl, <https://www.meczyki.pl/> und <https://transfery.info/>) auf Fehlerfreiheit anhand des „Alles-oder-nichts“-Ansatzes nach Würthele (2003) zeigt für die Datenkategorien *Spielerbezogene Daten* – Leistungsdaten der Spieler (z. B. Auswechslungen, rote Karten, gelb-rote Karten, Elfmeter), personenbezogene Spielerdaten wie Trikotnummern, Position, Geburtsort, Spielfuß sowie Transfers der Spieler bzgl. Zeitpunkten und Ablösesummen – sowie für die Kategorie *Vereinsbezogene Daten* – Vereinsjahresbilanzen mit Platzierung, erreichter Punktzahl und Wettbewerbe – jeweils hohe Übereinstimmungen der überprüften Daten von Transfermarkt.de mit den Daten von den selektierten polnischen Webseiten (Kasperek, 2020). Insgesamt wurden fünf Datensatzkategorien auf deren Fehlerfreiheit in Bezug auf polnische Datenbestände überprüft, es konnten 94,57 % der geprüften Daten bestätigt werden, 5,43 % der Daten von Transfermarkt.de stimmten also nicht mit den Daten der polnischen Fußballwebseiten überein oder wiesen fehlende Informationen auf. Tabelle 3 zeigt die Anzahl der verglichenen Daten sowie die ermittelten fehlerhaften Daten.

Eine vergleichbare Studie zur Fehlerhaftigkeit anhand des „Alles-oder-nichts“-Ansatzes nach Würthele (2003) überprüft die Daten von Transfermarkt.de aus englischen Ligen mit Daten von als geeignet identifizierten fußballbezogenen englischsprachigen Webseiten (www.soccerway.com, www.flashscore.com, www.worldfootball.net, www.sofascore.com und www.soccerpunter.com; Strauß, 2020). Die geringe Fehlerquote dieses Abgleichs zwischen 1,18 % und 2,71 % ($n=2.162$) führt ebenfalls zu dem Ergebnis, dass

die Daten als sehr vertrauenswürdig einzustufen sind (Strauß, 2020). Auch eine Überprüfung von Transfermarkt.de in Bezug auf den Abgleich von Daten der gleichen Kategorien aus der *Major League Soccer* (MLS – höchste Spielklasse im US-amerikanischen und kanadischen Fußball) mit anderen Daten amerikanischer Fußballstatistiken (www.soccerway.com, www.flashscore.com, www.worldfootball.net, www.sofascore.com und www.soccerpunter.com) durch Bialas (2021) zeigt eine sehr geringe Fehlerhaftigkeit der Transfermarkt.de-Daten von 1,87 % (n=38.257; 717 fehlerhafte Daten), während ein Vergleich nur der Spielerleistungsdaten für die spanische La Liga (www.flashscore.com) Fehlerhaftigkeiten (Toleranz-Ansatz) von 0 % (rote Karten, Gegentore bei Torwart, Einsätze, Ein-/Auswechslungen), 0,01 % (Kaderplätze, Spielminuten), 0,11 % (Tore), 0,14 % (Vorlagen), 0,16 % Gelbe Karten), 2,13 % (Gelb-rote Karten) bis maximal 4,76 % (Eigentore) ausweist (Haidl, 2021).

Tabelle 3: Abgleich Transfermarkt.de mit polnischen fußballbezogenen Webseiten: Anzahl überprüfter Daten, ermittelter fehlerhafter Daten und deren Anteile in Prozent sowie die Konfidenzober- und Untergrenzen des Fehlerquotienten bei Annahme der Poisson-Verteilung getrennt nach überprüfter Datensatzkategorie. Eigene Darstellung nach Kasperek (2020).

Datensatzkategorie	Anzahl überprüfter Daten (n)	Anzahl fehlerhafter Daten (n _f)	Anteil fehlerhafter Daten in Prozent (%)	KI _o (%)	KI _u (%)
Leistungsdaten Spieler*	415	17	4,10	6,04	2,15
Vereine Jahresbilanz**	180	8	4,44	7,52	1,36
Spieler*	202	26	12,87	17,82	7,92
Länderspiel Leistungsdaten*	100	0	0,00	0	0
Spieler Transfers*	135	5	3,70	6,95	0,46
Gesamt	1.032	56	5,43	6,85	4,01

*Spielerbezogene Daten; **Vereinsbezogene Daten; o = Obergrenze Konfidenzintervall; u = Untergrenze Konfidenzintervall

Die Qualität von Daten wird nicht nur durch *Fehlerhaftigkeit*, sondern auch durch *Vollständigkeit* beschrieben (Heinrich & Klier, 2015; Otto & Österle, 2016). Eine Prüfung der Transfermarkt.de-Datenbank auf Vollständigkeit nach Hinrichs (2002) sowie Hildebrand et al. (2015) weist eine Vollständigkeitsrate von 94,10 % auf der Datenbankebene aus und dient als Indikation der hohen Güte und Glaubwürdigkeit der Daten von Transfermarkt.de (Strauß, 2020). Tabelle 4 zeigt eine Übersicht der Vollständigkeitsraten der gleichen, auch auf Fehlerhaftigkeit untersuchten Datenkategorien. Eine Vollständigkeitsprüfung der Transfermarkt.de-Daten im Vergleich zu Daten der *Major League Soccer*

(MLS) anderer amerikanischer Fußballstatistiken ergibt, dass Transfermarkt.de mit 99,91 % nahezu vollständige Daten liefert (n=38.257; Bialas, 2021).

Tabelle 4: Übersicht über die Vollständigkeit der für die Studie erstellten Datenbank auf Relationenebene; Eigene Darstellung nach Strauß (2020).

Datensatzkategorie	Leistungsdaten Spieler*	Vereine Jahresbilanz**	Spieler*	Länderspiel Leistungsdaten*	Spieler Transfers*	Gesamt: Datenbank- ebene
Vollständigkeit auf Relationenebene	100,00%	98,51%	85,36%	100,00%	86,63%	94,10%

*Spielerbezogene Daten; **Vereinsbezogene Daten

Aufgrund der Ergebnisse der Datenqualitätsprüfung auf Fehlerhaftigkeit mit einer maximalen Fehlerquote von 6,85 % der Daten im Vergleich zu den polnischen Webseiten (Kasperek (2020) und mit 2,71 % im Vergleich mit den englischen Ligen (Strauß, 2020), sowie mit einer Vollständigkeit von 94,10 % (Strauß, 2020) und 99,91 % (Bialas, 2021) auf Datenbankebene werden die Daten von Transfermarkt.de als verlässliche, valide Quelle für fußballbezogene Daten eingeschätzt und in der wissenschaftlichen Evaluation des Erfolgs von Nachwuchsleistungszentren für das Konstrukt *Hervorgebrachte Spielerqualität NLZ* in dieser Studie verwendet.

Daher wurden die für die Studie relevanten Daten aus Transfermarkt.de erfasst und in einer selbsterstellten relationalen Datenbank (IBM SPSS Statistics, Version 26.0) abgelegt. Die Datenbank enthält Daten von 329 verschiedenen Ligen aus 248 abhängigen und teil-autonomen Staatsgebieten (im Folgenden Länder genannt). Verzeichnet sind insgesamt 120.675 gespielte Halbsaisons von Spielern aus 4.949 Vereinen über einen Zeitraum von 19 Jahren (38 Halbsaisons – eine Fußballsaison besteht aus zwei Halbsaisons, Sommer und Winter).

Der Marktwert ist ebenfalls als ein spielerbezogenes Datum kategorisiert, wird aber aufgrund des besonderen Prozesses des Zustandekommens und der damit anderen Glaubwürdigkeitsdiskussionen im folgenden Kapitel erläutert.

3.5.1.1.3 Marktwert

Das Alleinstellungsmerkmal der Fußballplattform Transfermarkt.de ist neben der Informationsvielfalt zu Vereinshistorien, spielerbezogenen Einsätzen und Leistungsdaten die Angabe und Spekulation von *Marktwerten* einzelner Spieler und für Vereine in Summe. Transfermarkt.de bezeichnet sich als „meistzitierte Quelle bezüglich Marktwerte“ (Media Impact, 2021). Im Gegensatz zu anderen viel genutzten Sportportalen und -redaktionen, wie Kicker (www.kicker.de) oder Sport1 (www.sport1.de), werden bei Transfermarkt.de für die Marktwertanalyse die Schätzungen nicht durch Recherchen von Experten oder Journalisten eingeholt, sondern durch die Einschätzung von vielen engagierten Usern der Plattform. Hinsichtlich der Marktwertanalyse werden zunächst individuelle

Vorschläge der User gesammelt und anschließend zur Prüfung an Experten (erfahrene Mitglieder und/oder Angestellte der Transfermarkt GmbH & Co. KG) weitergeleitet. Diese legen dann auf Grundlage der vorgeschlagenen Marktwerte einen „objektiven“ Marktwert für jeden einzelnen Spieler fest (Prockl & Frick, 2018). Zu jedem Spieler gibt es eine Historie mit entsprechendem Diagramm, das aufzeigt, wie viel ein Spieler in jeder Saison wert war (Transfermarkt.de, 2021a). Um die Datenqualität, der in dieser Art und Weise entstandenen Marktwerte, zu beurteilen, kann auf umfangreiche Forschungen zu den Themen „Weisheit der Vielen“ und „Schwarmwissen“ zurückgriffen werden.

Durch die Anwendung des Prinzips der *Weisheit der Vielen* nach Surowiecki (2005) wird eine Fehleinschätzung einzelner Bewerter durch die Masse an Bewertungen ausgeglichen (Gerhards et al., 2014). Derartige Gruppenentscheidungen stellen vielversprechende Lösungsansätze dar und derart getroffene Einschätzungen sind oftmals verlässlicher als etwa Entscheidungen, die von Einzelnen getroffen werden (Surowiecki, 2005). Das Prinzip der Weisheit der Vielen wird aber von Datenqualitätsforschern kritisiert (Ackermann & Follert, 2018; Charness & Sutter, 2012; Lorenz et al., 2011; Peeters, 2018). Insbesondere vor dem Hintergrund der Verwendung der Transfermarkt.de-Datenbank für wissenschaftliche Zwecke, wird die Qualität der auf Transfermarkt.de veröffentlichten Daten aufgrund der freien Zugänglichkeit und Fehleranfälligkeit kritisch hinterfragt. Obwohl Prockl und Frick (2018) anhand ihrer Untersuchungen Belege für die Aktualität und Glaubwürdigkeit der generierten Werte von Transfermarkt.de präsentieren konnten, geben sie zu bedenken, dass die Veröffentlichung der Daten auf Transfermarkt.de oft aus wissenschaftlicher Sicht zu selektiv sei. Die Validität und die Vorhersagekraft der dort präsentierten Daten werden angezweifelt, aber einige Autoren konnten nachweisen, dass die über Transfermarkt.de ermittelten Marktwerte mit real beobachteten Spielergehältern signifikant zusammenhängen und Ergebnisprognosen auf Basis der dort veröffentlichten Marktwerte möglich sind (z. B. Gerhards et al., 2014; Ackermann & Follert, 2018; Prockl & Frick, 2018).

Die Marktwerte werden zunehmend in der Sportpresse und für Statistikportale wie Statista verwendet. Darüber hinaus werden die veröffentlichten Daten in der Fußballpraxis für die weiteren am Fußballmarkt Beteiligten – Spieler, Unternehmen, Spielerbeobachter – verwendet. Transfermarkt.de liefert mit der Marktwertanalyse einen Ansatz, wie Vereine und Spieler im Kontext von Transferverhandlungen bewertet werden können (Gerhards et al., 2014; Ackermann & Follert, 2018). Spieler nutzen den Marktwert als Argumentation für Vertragsverhandlungen mit einem Verein, Unternehmen nutzen den Marktwert als Kennzahl für die Vermarktungsfähigkeit von Spielern oder Sponsorenverträge (Frick, 2001) und Spielerbeobachter, die als sog. Talentscouts für Vereine neue Spieler identifizieren, nutzen die angebotenen Marktwertstatistiken für eine Analyse des beobachteten Spielers, um seine Eignung für den Auftraggeber zu beurteilen (Peeters, 2018).

Um die Forschungen über die monetäre Messbarkeit des Erfolgs von Nachwuchsleistungszentren im deutschen Profifußball abzusichern, ist zu überprüfen, ob der Marktwert von Transfermarkt.de geeignet ist, in (sport-)ökonomischen Studien verwendet zu

werden, und die Entstehung des Wertes den Qualitätsmaßstäben wissenschaftlicher empirischer Sozialforschung standhält (Döring & Bortz, 2016). Der Marktwert von *Transfermarkt.de* ist Gegenstand einiger wissenschaftlicher Untersuchungen mit jeweils unterschiedlichen Aspekten (hierzu zählen Bryson et al., 2013; Franck & Nüesch, 2012; Gerhards et al., 2014; Herm et al., 2014; Wallebohr & Daumann, 2019). Beispielsweise untersuchten Gerhards et al. (2014) neben der internen Ungleichheit, der kulturellen Diversität und dem Grad der Fluktuation einer Mannschaft, in welchem Maße der Marktwert einer Mannschaft über deren sportlichen Erfolg entscheidet. Der Marktwert selbst wurde durch Herm et al. (2014) auf seine Genauigkeit hin untersucht und es wurde aufgezeigt, welche Attribute für die Marktwertbewertung am wichtigsten sind. Eine Einordnung der Marktwertanalyse in die moderne Bewertungstheorie nahmen Ackermann und Follert (2018) vor.

Die Transfermarkt.de-Community hat sich zur Ermittlung des *Marktwertes* eines Spielers Leitlinien gegeben, an denen sich User und Experten orientieren sollen (Transfermarkt.de, 2013; Ackermann & Follert, 2018). Der Marktwert soll einerseits den objektiv ermittelten Substanzwert eines Spielers darstellen, andererseits aber auch die Zukunftsperspektiven abbilden (Transfermarkt.de, 2017). Zu den Einflussgrößen werden Alter, Leistung, Zukunftsperspektive, real existierende Nachfrage und die bisher gezahlte Ablösesumme genannt. Des Weiteren fließen Prestige und marketingtechnische Aspekte mit ein. Prinzipiell werden die Marktwerte sowohl am Ende als auch in der Mitte einer Saison angepasst. Im Herbst und Frühjahr werden nur bei enormen Formschwankungen Änderungen vorgenommen. Bevor endgültig die Änderung eingetragen wird, werden die Marktwerte von Transfermarkt.de geprüft und anschließend freigegeben (Transfermarkt.de, 2013).

Eine Grundlage zur Schätzung von Marktwerten bieten Vergleiche zu ähnlichen Spielern (z. B. Position, Leistung, Alter) von Ligakonkurrenten oder vergleichbaren Ligen. Ein Spieler wird durch die User und Experten auf Transfermarkt.de aufgrund vorliegender Leitlinien bewertet – unabhängig vom bisherigen Marktwert (Transfermarkt.de, 2013). Es wird nach Spielertypen – wie *junge Talente, gestandene Profis im besten Fußballalter, erfahrene Profis auf das Ende der Karriere zugehend* – unterschieden. Die Unterscheidung dient als ungefähre Einstufung, um den Marktwert des Spielers einzeln und im Gesamtkontext zu schätzen (ebd.).

User und Experten orientieren sich bei der Einschätzung des Marktwertes an der Ablöse, der Vertragslänge beim bisherigen Verein, der voraussichtlichen Vertragslänge beim zukünftigen Verein, dem Wert des Spielers für den Verein sowie dem Gehalt des Spielers und die Chancen auf Gehaltssteigerungen bei Vertragsverlängerung (ebd.). Weitere Einflussfaktoren sind Nationalmannschaftszugehörigkeit und Verletzungen des Spielers (Transfermarkt.de, 2013). Zum Startpunkt einer Karriere als Profifußballer werden je nach Liga, in der der Spieler einen Vertrag hat, in den Leitlinien Annahmen getroffen. Startwerte sind z. B.: für die 4. Liga 25.000 Euro, Liga 3 50.000 Euro und für die 1. und 2. Liga jeweils 100.000 Euro (Transfermarkt.de, 2013). Eine abschließende Bewertung der von den Usern vorgeschlagenen Marktwerte erfolgt durch Experten des

Betreibers der Webseite. Dieser hierarchische Bewertungsprozess ist effektiver als ein durchgängig demokratisches System und beugt Manipulationen etwa durch Spieler, Spielerbeobachter oder Fans eines bestimmten Vereins vor (O. Müller et al., 2017).

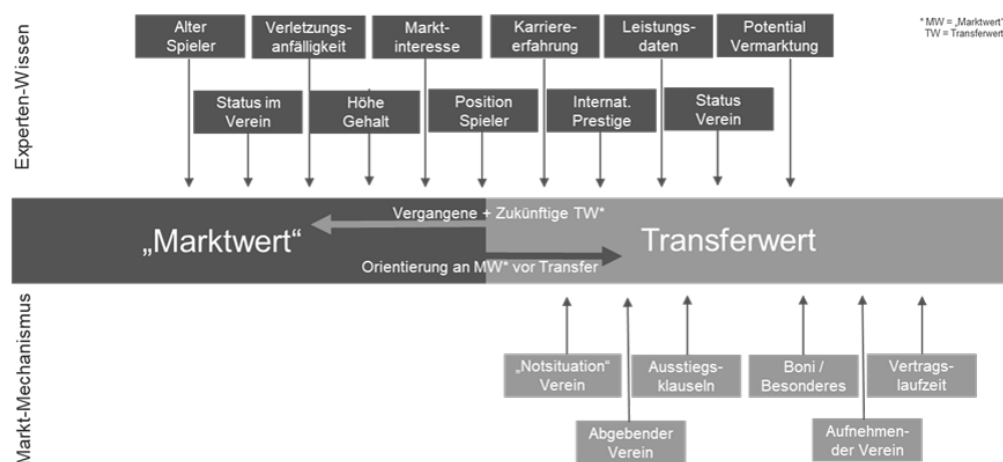
„Grundsätzlich versteht man unter dem Marktwert den Preis eines Gutes zum Zeitpunkt einer Transaktion. Marktwerte konkretisieren sich damit in den Transferwerten, den Geldbeträgen also, den ein Verein für einen Spieler zahlt, wenn er zum Verein wechselt“ (Gerhards et al., 2014, S. 234). Einheitlich angewendete objektive Verfahren zur Ermittlung des Marktwertes von Spielern existieren in Deutschland bisher nicht (Elter, 2011). Der Marktwert von Transfermarkt.de basiert auf einer Sichtweise des sportlichen Erfolgs und der sportlichen Leistung im Fußballkontext (ebd.). Transfermarkt.de ist der Auffassung, dass der Marktwert für Fußballer im wirtschaftswissenschaftlichen Sinne nicht wie der Wert eines Gutes betrachtet werden sollte, da gemäß ihrer Einschätzung der Markt für Profifußballspieler eher geprägt ist durch mangelnde Transparenz, Heterogenität der Spieler und den hohen Ablösen, die bei Transfers von Spielern gezahlt werden (Transfermarkt.de, 2013).

Zusätzlich zum Marktwert werden auf Transfermarkt.de pro Spieler *Transfererlöse* bei Vereinswechseln veröffentlicht. Transfererlöse sind gemäß der Leitlinien von Transfermarkt.de strikt von den Marktwerten zu trennen. Sie unterscheiden sich in vielen Fällen, da situationsbedingte Faktoren wie Finanz- oder Personalnot, Zeitdruck, Wechselwunsch oder zur Verfügung stehende finanzielle Mittel den Preis eines Spielers in beide Richtungen ausschlagen lässt. Der Wert, zu dem bei Wechselwunsch ein Verein bereit ist, einen Spieler zu verkaufen bzw. zu kaufen, ist der *Transferwert*. Der Transferwert wird durch Angebot und Nachfrage auf dem internationalen Spielermarkt durch die beiden Marktteilnehmer – den abgebenden und dem aufnehmenden Verein – bestimmt. Der Transferwert wird insbesondere geprägt durch die restliche Vertragslaufzeit des Spielers, Sonder- und Ausstiegsklauseln und eventuell dringlichem Bedarf für die Kompetenzen des Spielers im Kader des aufnehmenden Vereins. Hinzu kommt der Druck durch enge zeitliche Rahmen, Transfers abzuwickeln (Transferfenster). Marktwert und Transferwert differieren daher häufig.

Einen Überblick über die Einflussfaktoren auf die Höhe des Markt- und des Transferwertes sowie ihre Beziehung untereinander gibt die folgende Abbildung 33.

Zum einen fließt ein Transferwert aus einem vergangenen Transfer und die Abschätzung der Höhe eines zukünftigen Transferwertes in den Marktwert von Transfermarkt.de ein. Zum anderen orientieren sich Spieler und Vereine im Kontext von Transferverhandlungen an dem Marktwert, den die Plattform angibt, was wiederum den Transferwert beeinflusst (Transfermarkt.de, 2017). Eine eindeutige Trennung der Ermittlung der Transfer- und der Marktwerte ist demnach nicht gewährleistet, sie beeinflussen sich gegenseitig.

Abbildung 33: Abgrenzung Einflussfaktoren auf den Marktwert und Transferwert. Eigene Darstellung nach Transfermarkt.de (2013)



Transferwerte werden auf Transfermarkt.de erfasst, wenn bei einem stattgefundenen Transfer für einen Spieler eine Transfersumme gezahlt wird (Transfermarkt.de, 2021a). Somit existiert nicht für jede Saison für jeden Spieler ein Transferwert, der den Wert des Spielers für die betrachtete Saison widerspiegeln könnte. Der Marktwert eines Spielers beruht nicht nur auf den Schätzungen von Usern und Experten, sondern basiert auf verschiedenen Faktoren, die einen objektiven Substanzwert ergeben, mit der sich Spieler bestmöglich vergleichen lassen (Herterich, 2016).

Gerhards et al. (2014) konnten eine Korrelation zwischen der Marktwerteinschätzung und den im Falle eines Transfers tatsächlich erzielten Transfersummen von 0,93 ermitteln. Marktwerte, die auf der Expertenschätzung von Transfermarkt.de (?) beruhen, können damit als reale Marktwerte interpretiert werden (Gerhards et al., 2014). Herm et al. (2014) haben den Einfluss verschiedener Spielermerkmale auf den jeweiligen Marktwert eines Spielers auf Basis der Transfermarkt.de-Daten untersucht. Sie schließen aus ihrer Analyse mittels Regression ($R^2 = 0,90$, $b = 0,95$, $p < 0,01$), dass die Marktwerte mit den tatsächlichen Transfergebühren zusammenhängen und als Prädiktoren dienen können (Herm et al., 2014). Heuer (2012) gelang es, einen Zusammenhang ($r=0,67$) zwischen den Marktwerten von einer Saison und der jeweiligen Tordifferenz nachzuweisen. Torgler et al. (2006) haben in ihrer Studie zur relativen Einkommensposition und Performance, die Marktwerte von Transfermarkt.de mit den Marktwerten des Fußball-Magazins *Kicker Sportmagazin* verglichen. Die Redaktion des *Kicker Sportmagazins* entwickelt zu jedem Saisonbeginn eine Schätzung der Marktwerte der Spieler (ebd.) in den drei oberen Fußballligen(?). Die Daten werden seit mehreren Jahren von fast denselben Fußballexperten systematisch ermittelt und daher als konsistent eingeschätzt. Die durch die Fußball-Community ermittelten Marktwerte von Transfermarkt.de und die Daten der Experten korrelieren mit $r=0,754$ hoch (Torgler et al., 2006).

Der von Transfermarkt.de veröffentlichte Marktwert wird aufgrund der Ergebnisse der publizierten Studien als besonders zuverlässig interpretiert und in dem, dieser Arbeit zugrundeliegenden, Modell zur Messung des Erfolgs von Nachwuchsleistungszentren

zur monetären Bewertung der hervorgebrachten Spielerqualitäten eines Nachwuchsleistungszentrums verwendet. Für die Berechnung des Nutzens kommt es darauf an, dass in den Marktwerten die Spielerqualität eines Spielers auf einen einheitlichen Nenner gebracht wird – d. h., dass diese einen monetären Maßstab ergeben, mit dem man einfach rechnen und prognostizieren kann (Wagner, 2014). Die monetäre Bewertung der hervorgebrachten Spielerqualitäten pro Nachwuchsleistungszentrum kann somit mit dem durchschnittlichen Marktwert pro Spieler pro Betrachtungsjahr erfolgen.

Die für diese Studie relevanten, von Transfermarkt.de extrahierten Daten sind in einer relationalen Datenbank mit verteilten Daten abgelegt (SPSS-Version 26.0.0.0, 64-Bit-Version).

3.5.1.2 Double Pass Scores

Grundlegendes Gütekriterium für eine quantitativ-empirische Untersuchung ist die Konstruktvalidität. Zu prüfen ist dabei, ob die Messwerte theoretisch sinnvoll interpretierbar sind und das interessierende theoretische Konstrukt repräsentieren (Döring & Bortz, 2016). Eine Konstruktvalidität wird als gegeben angesehen, wenn Ergebnisse, die das gleiche begriffliche Konstrukt messen sollen, hoch miteinander korrelieren (Helfrich, 2016). Zur Validierung der entwickelten Bewertung der Nachwuchsleistungszentren mittels der *Hervorgebrachten Spielerqualität NLZ* werden daher die Resultate der bisher zur Zertifizierung der Nachwuchsleistungszentren durchgeführten Evaluationen herangezogen, sog. Double PASS Scores. Die Double PASS Scores wurden vom DFB und der DFL als Excel-Datei zur Verfügung gestellt (D. Feld, Persönliche Mitteilung, 18. April 2021). Im Folgenden wird das Verfahren zur Erhebung der Double PASS Scores und deren Eignung als vergleichbares begriffliches Konstrukt zur Messung des Erfolgs von Nachwuchsleistungszentren erläutert.

Die Vereine mit Nachwuchsleistungszentren waren lt. Statuten des DFB und der DFL – neben der Lizenzierung – verpflichtet, an den regelmäßigen Zertifizierungen der Nachwuchsleistungszentren teilzunehmen. Die Zertifizierung erfolgte von 2007 an in Zyklen durch den Serviceanbieter *Double PASS* mit dem Professional Academy Support System (PASS). Ein Zyklus erstreckt sich über jeweils drei Jahre, für den ersten Zyklus Pass 1.0 über vier Jahre.

Im Zertifizierungsprozess hatten die Vereine für die Messung verschiedener Qualitätsdimensionen hinsichtlich des Zustands des Nachwuchsleistungszentrums Nachweise zu erbringen. Die Datensammlung erfolgte softwarebasiert und war die Basis für das anschließende dokumentarische Audit. Im Anschluss daran evaluierten zwei Experten der Agentur Double PASS die Nachwuchsleistungszentren vor Ort. Gespräche und strukturierte Interviews der Beteiligten (z. B. Geschäftsführung, Vorstand, Sportmanagement, Lizenztrainer*innen, Mitarbeiter*innen, Spieler) und Beobachtungen von Trainingseinheiten und Spielen führten zu einer konsolidierten Bewertung. Die Auditor*innen bewerteten die Ausprägung der Kriterien mittels eines Punktesystems (maximal er-

reichbare Punktzahl: 5000). Darauf aufbauende Scores messen die Qualität der einzelnen Dimensionen und des Gesamtsystems. Die Scores des Gesamtsystems sind als Kennzahl für die Höhe der Qualität des Nachwuchsleistungszentrums durch den DFB und der DFL verwendet worden (Kroemer, 2015).

Im HR-Controlling werden eingesetzte Kennzahlen grob in input-orientiert und output-orientiert unterteilt (Wickel-Kirsch, 2012). Input-orientierte Kennzahlen betrachten dabei z. B. eingegangene finanzielle, personelle und zeitliche Ressourcen. Output-orientierte Kennzahlen messen das Ergebnis von z. B. Bildungsmaßnahmen (Fritz, 2012; Seiber, 2000).

Der Double Pass Score ist den input-orientierten Kennzahlen zuzuordnen. Er misst, in welcher Qualität Kompetenzen im Nachwuchsleistungszentrum zur Talentausbildung entwickelt worden sind bzw. welcher Aufwand betrieben worden ist. Der Double Pass Score ist Maß, das in der Bewertungspraxis zur Auszeichnung der Güte der Nachwuchsleistungszentren mit einer bestimmten Anzahl „Sterne“ eingesetzt wurde und als Kriterium für den Erfolg eines Nachwuchsleistungszentrums zu einem Ranking führt (Kroemer, 2015). Double PASS Scores liegen seit 2007 vollständig für insgesamt 56 Nachwuchsleistungszentren ($n=56$) für vier Zyklen vor: PASS 1.0: $n=41$, Zeitraum 2007 bis 2010; PASS 2.0: $n=44$, Zeitraum 2011 bis 2013; PASS 3.0, $n=53$, Zeitraum 2014 bis 2016; PASS 4.0, $n=54$, Zeitraum 2017 bis 2019 (D. Feld, Persönliche Mitteilung, 18. April 2021). Die visuelle Überprüfung auf Normalverteilung mittels eines Histogramms oder eines QQ-Diagramms erscheint subjektiv, daher wird der Shapiro-Wilk-Test als statistischer Signifikanztest nach objektiveren Maßstäben eingesetzt. Der Shapiro-Wilk-Test bietet eine hohe Teststärke bzw. Testpower, insbesondere bei kleineren Stichproben mit $n < 50$ (Hedderich & Sachs, 2018). Die Double Pass Scores sind gemäß Shapiro-Wilk-Test für den Zyklus PASS 1.0 normalverteilt ($p > 0,05$), für die weiteren Zyklen hingegen nicht ($p < 0,05$). Die deskriptive Statistik zeigt Tabelle 5.

Tabelle 5: Deskriptive Statistik Double Pass Scores nach PASS-Zyklen.

Deskriptive Statistik Double Pass Scores nach PASS Zyklen						
Zyklus im Zeitraum	N	Minimum	Maximum	Mittelwert	Std.- Abweichung	Shapiro-Wilk- Test
Pass 1.0 (2007 - 2010)	44,00	14,32	79,30	54,03	15,26	0,057
Pass 2.0 (2011 - 2013)	47,00	15,45	76,80	51,99	15,95	0,034
Pass 3.0 (2014 - 2016)	56,00	14,10	83,20	61,00	15,17	0,001
Pass 4.0 (2017 - 2019)	57,00	12,87	84,00	61,80	14,20	0,001

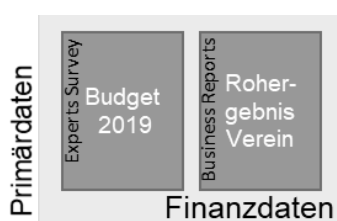
Die Daten des Double-PASS-Scoringmodells werden als Fakten angesehen und als geeignetes Maß eingeschätzt, den entwickelten Indikator auf Glaubwürdigkeit zu testen. Die Validierung der *Hervorgebrachten Spielerqualität NLZ* mit *Double PASS Score NLZ* erfolgt mittels einer bivariaten (einfachen) Korrelationsanalyse (Helfrich, 2016).

3.5.2 Erhebung der Primärdaten

Ziel dieser Studie ist es, anhand von betriebswirtschaftlichen Kennzahlen den Erfolg einzelner Nachwuchsleistungszentren zu ermitteln. Dazu wird in dieser Arbeit der Return on Investment (ROI) berechnet, das Kostenelement des ROI wird über das *Investment in ein Nachwuchsleistungszentrum* operationalisiert. Die Daten zum Investment in das jeweilige Nachwuchsleistungszentrum der betrachteten Vereine sind als Sekundärdaten über öffentlich zugängliche Datenbanken oder Informationssammlungen nicht vollständig und in hinreichender Datenqualität für den gesamten Betrachtungszeitraum verfügbar. Experten des DFB und DFL rieten von einer direkten Befragung der Vereine mit angeschlossenen Nachwuchsleistungszentrum zum *Investment in den Nachwuchsbereich* für den Betrachtungszeitraum ab, da die Return-Rate zu Befragungen im Rahmen von Forschungen erfahrungsgemäß gering sei, Finanzdaten als sensibel eingestuft und nicht herausgegeben würden sowie Investmentdaten in den Nachwuchsbereich gemäß aktueller Definition für zurückliegenden Jahren in den Vereinen eher nicht zur Verfügung ständen (M. Honsel, Persönliche Mitteilung, 7. Februar 2019). Eine Befragung der Vereine wurde daher aus forschungsökonomischen Gründen nicht durchgeführt und wird im Evaluationskonzept nicht implementiert.

Das Investment in ein Nachwuchsleistungszentrum wird unter Berücksichtigung der Erkenntnisse aus der Literatur, relevanter öffentlich zugänglicher Daten und Schätzungen der Experten bei DFB und DFL ermittelt. Für die Ermittlung des Konstrukts *Investment in ein Nachwuchsleistungszentrum* werden daher folgende Finanzdaten erhoben: Rohergebnis Verein und Budget 2019 (siehe Abbildung 34).

Abbildung 34: In Investment Nachwuchsleistungszentrum eingehende Finanzdaten



Im Weiteren wird die Erhebung der Primärdaten im Einzelnen erläutert.

3.5.2.1 Rohergebnis Verein

Laut Lizenzierungsordnung des DFL Deutsche Fußball Liga e.V. (2021, §2) hat jeder Verein für eine Lizenzerteilung den Nachweis der wirtschaftlichen Leistungsfähigkeit zu erbringen. Der Nachweis wird erbracht mit der Übermittlung unter anderem der Konzernbilanz, der Konzerngewinn- und -verlustrechnung für das Geschäftsjahr und einem Wirtschaftsprüferbericht über die Prüfung der erforderlichen Unterlagen (siehe dazu §8 der Lizenzierungsordnung). Die Kennzahl Rohergebnis wird insbesondere in jüngerer

Zeit als relevante finanzielle Kennzahl durch die DFL in den Wirtschaftsreports ausgewiesen (DFL, 2019b, 2020a, 2021a).

Die Erhebung der *Rohergebnisse* erfolgt anhand öffentlich zugänglicher Daten aus Konzernbilanzen, Geschäftsberichten und Gewinn- und Verlustrechnungen der Vereine und/oder bereits über Statista (statista.com) oder andere Portale – wie *fussball-geld.de*, *football-exchange.de*, *spielwerlagerung.de*, *northdata.de* – verdichtete Daten sowie den Wirtschaftsreports der DFL über die Jahre des Betrachtungszeitraum. Daten für die 1. Bundesliga liegen aufgrund der hohen Anzahl öffentlich rechenschaftspflichtiger Vereine in der Regel öffentlich zugänglich vor oder sind über die jeweiligen Mutterkonzerne in Erfahrung zu bringen. Für die 2. Bundesliga ist die Datenerhebung aufgrund der fehlenden Rechenschaftspflicht limitiert und war im Rahmen dieser Arbeit insbesondere für zurückliegende Jahre weniger erfolgreich. Die Erhebung der Ergebniszahlen für die Liga 3 war ebenfalls aus den gleichen Gründen erfolglos, es sei denn, der betreffende Verein war zeitweilig Mitglied der 1. oder 2. Bundesliga. Daher wurde zusätzlich zur Recherche der öffentlich verfügbaren Daten und trotz des Hinweises der DFL und des DFB zur zu erwartenden schlechten Rücklaufquote an 40 Vereine der 2. Bundesliga und Liga 3 eine persönliche E-Mail mit der Bitte um Zusendung der Umsatzdaten geschickt. Die Rücklaufquote hinsichtlich der erfragten Daten lag bei 0 %. Fehlende Werte sind die Folge dieser gewählten Art der Datenerhebung, lediglich öffentlich verfügbare Daten liegen zur Verwendung vor. Bei 54 Nachwuchsleistungszentren und 11 Jahren Betrachtungszeitraum (2009-2019) sind 594 Datenpunkte zu erwarten. 316 Datenpunkte konnten erhoben werden, 278 Datenpunkte fielen aus. Der Datenausfall ist bedingt zufällig (MAR, Missing at Random) in Abhängigkeit von der Liga-Zugehörigkeit und der Zeit. Insbesondere für zwölf Vereine, die in der Saison 2018/19 in der Liga 3 oder niedriger (4. oder 5. Liga) spielten, konnten keine Rohergebnisse erhoben werden, für sechs Vereine der Liga 3 aber insgesamt 24 Datenpunkte. In der 2. Bundesliga sind 83 fehlende Daten zu verzeichnen, davon 56 Datenpunkte bei Vereinen, die bisher nicht in der 1. Bundesliga gespielt und keine wirtschaftlichen Daten in den Jahren 2009 bis 2015 veröffentlicht haben. Drei Vereine starteten ihre Veröffentlichungen ab dem Jahr 2017: Holstein Kiel und Jahn Regensburg, nachdem sie in die 2. Bundesliga aufgestiegen waren, Sandhausen trotz der Zugehörigkeit zur 2. Bundesliga seit der Saison 2012/13 dagegen erst, als die Kennzahl durch die DFL erfragt wurde. 21 Datenpunkte fehlen in der 1. Bundesliga, davon 4 bei RB Leipzig, die 2009 in den Profifußball eingetreten sind und erst ab dem Jahr 2013 nach dem Aufstieg in die Liga 3 Wirtschaftsdaten veröffentlichten. 8 Datenpunkte fehlen in der 1. Bundesliga bei Vereinen, deren wirtschaftliche Ergebnisse in der Bilanz der finanzierenden Konzerne integriert sind und nicht explizit ausgewiesen werden (Bayer Leverkusen, VfL Wolfsburg). Für den 1. FC Nürnberg und Fortuna Düsseldorf sind Datenlücken im Zeitverlauf 2011-2014 zu verzeichnen. Die aufgezeigten Datenlücken bei einer Gesamtpopulation von 54 Nachwuchsleistungszentren sind nicht zu ignorieren und müssen entsprechend behandelt werden.

Daher werden zunächst die zur Verfügung stehenden Optionen für die Füllung der Lücken bei Daten, die Missing at Random sind, geprüft. „Das Ziel aller Ergänzungstechniken ist es, dem Analysten eine vollständige Datenmatrix (oder mehrere) zur Verfügung zu stellen, die mit den üblichen Methoden ausgewertet werden kann“ (Rässler, 2000). Eine Option ist es, die Berechnung nur mit den Vereinen durchzuführen, die im Beobachtungszeitraum keine Datenlücken aufweisen. Diese Fallreduktion würde den dieser Arbeit zugrundeliegenden Datensatz auf 16 Vereine reduzieren. Ein zu hoher Datenverlust für diese Studie, da damit die Repräsentativität der Studie nicht gewährleistet wäre. Es erfolgt aber stattdessen eine Fallreduktion um 12 Vereine der Liga 3 oder tieferer Ligen, für die aus öffentlich zugänglichen Quellen keinerlei Rohergebnisse erhoben werden konnten. Sie werden bei dieser Untersuchung nicht weiter berücksichtigt, damit reduziert sich die Berechnung auf 42 Nachwuchsleistungszentren. Eine weitere Option ist es, die Werte auf individueller Ebene zu ergänzen; liegen entsprechende Daten nicht vor, sollte durch die Forscher*in eine Logik entwickelt werden, die möglichst plausible Schätzungen der fehlenden Werte ermöglicht. (vgl. McKnight & McKnight, 2007).

Als Imputationstechniken für den vorliegenden Datenbestand können unterschiedliche Verfahren wie z. B. Mittelwert/Median basierte Ergänzungen und Hot-Deck-Verfahren (Historische Ergänzung und Regressions-Methoden; zu weiteren möglichen Verfahren siehe Rässler, 2000) angewendet werden. Ihre Eignung und Umsetzbarkeit für die fehlenden Rohergebnisse wurden für diese Studie zunächst überprüft.

In den letzten Jahren wird die *Multiple Imputation* als eine geeignete Regressions-Methode für den Umgang mit fehlenden Daten eingesetzt (Böwing-Schmalenbrock & Jurczok, 2012). Unter der Annahme, dass ein Zusammenhang des Rohergebnisses mit anderen Variablen besteht, deren Erhebung über öffentlich verfügbare Datenquellen möglich ist, wurden die Zusammenhänge der Rohergebnisse pro Jahr mit der Vereinsqualität, der Anzahl Zuschauer, der Ligahöhe, den Platzierungen in der Liga, dem Torverhältnis (geschossene und erhaltene Tore), der Punktzahl, den gewonnenen, verlorenen und unentschiedenen Spielen sowie den erhaltenen Fernsehgeldern analysiert (Analyseergebnisse siehe Anhang). Ein fundiertes Modell, das mit dem höchsten R^2 vollständig plausibel ist, konnte aufgrund von ungeschlossenen und wechselnden Wirkungsrichtungen bei der Hinzunahme von Variablen in das Modell nicht identifiziert werden.

In weiteren Überlegungen wurden fehlende Werte durch den Mittelwert der beobachteten Ausprägungen über alle Vereine pro Jahr (mit und ohne Ausreißer) ersetzt. Da eine große Spannbreite der Ausprägung der Rohergebnisse zwischen den unterschiedlichen Ligen und innerhalb der Ligen zu erkennen war, wurden stattdessen die Mittelwerte pro Liga (Gruppenmittel) eingesetzt. Auch diese Ergänzungen führten zu unplausiblen Fehleinschätzungen.

Das Verfahren der Historischen Ergänzung, unter Berücksichtigung mit vorher und nachher liegenden Daten für einen Verein in der Dimension Verein, sowie einer Trendkorrektur wurde ebenfalls von der Analystin verworfen, da mit diesem Verfahren das Risiko einer erheblichen Unterschätzung einhergehen kann (Rässler, 2000).

Am logischsten erscheint der Analystin die Schätzung der fehlenden Daten auf Basis der jährlichen Steigerungsraten der Rohergebnisse der Vereine, für die in den zwei aufeinanderfolgenden Jahren Daten vorliegen. Die Steigerungsraten werden je Liga getrennt ermittelt. Dieses Vorgehen bildet sowohl die allgemeine wirtschaftliche (Inflation) als auch die wirtschaftliche Entwicklung der Fußballbranche ab. Mit der Imputation liegt zu weiteren Berechnungen und zur Analyse eine vollständige Datenmatrix für 42 Nachwuchsleistungszentren für den Beobachtungszeitraum (11 Jahre) mit 473 Datenpunkten vor.

Die Ermittlung des Investments in ein Nachwuchsleistungszentrum für die Betrachtungsjahre 2009 bis 2019 (11 Jahre) beruht auf der Annahme, dass pro Jahr ein bestimmter prozentualer Anteil am Ergebnis des Vereins in das jeweilige Nachwuchsleistungszentrum investiert wird. Die Ermittlung der Höhe des anzusetzenden Prozentsatzes wird im Folgenden beschrieben und bewertet.

3.5.2.2 Budget 2019

Die Variable *Budget 2019* repräsentiert die Finanzierung des Betriebs des Nachwuchsleistungszentrums in Euro. Allgemein- und Personalkosten für das Nachwuchsleistungszentrum sind enthalten, Kosten für die U23-Mannschaft (wenn vorhanden), Frauenabteilungen, Fußballschule und Breitensport nicht (A. Bonarius, Persönliche Mitteilung, 11. Februar 2019). Die Budget-2019-Daten werden zur Ermittlung eines möglichst realitätsnahen prozentualen Anteils der Investitionen am Umsatz eingesetzt, die Daten für die Saison 2019/2020 sind über eine Expertenbefragung erhoben worden.

Expertenbefragungen stellen eine Variante von Interviews dar. Die Befragten gelten als fachliche Expert*innen, die zu ihrem Fachthema befragt werden, um so ihr strukturelles Fachwissen und/oder Praxis-/Handlungswissen zu erschließen (Döring & Bortz, 2016). Es geht bei Expertenbefragungen darum, fundierte Aussagen zu bestehenden Sachverhalten zu erhalten. Mit dieser Expertenbefragung werden Informationen abgefragt, die aus anderen Quellen als nur schwer zugänglich eingeordnet wurden und lediglich mit erheblich höherem Aufwand hätten erhoben werden können. Bei der Expertenbefragung handelt es sich in der Regel um eine Methode der qualitativen Sozialforschung, deren Gegenstand die Rekonstruktion subjektiver Sichtweisen und die auf das Verstehen von Zusammenhängen ausgerichtet ist. Der Befragung für die vorliegende quantitative Studie liegt ein eher informationsbezogenes Forschungsinteresse an Daten zugrunde, daher werden die Aussagen der Experten als faktische Befunde interpretiert (Helfferich, 2011).

Eine Besonderheit von Experten ist, dass sie in einem bestimmten Feld über besonderes fachliches Wissen verfügen und sich somit für die Befragung vorrangig eignen. Die Auswahl der passenden Gesprächspartner ist dabei abhängig von der jeweiligen Forschungsfrage, die Grundlage und Ausgangspunkt für die Entwicklung der Interviewfragen ist (Helfferich, 2011).

Mit Nachwuchsleistungszentren im deutschen Profifußball wird ein sehr spezifischer Bereich untersucht, der durch Lizenzierungsbestimmungen des DFB und der DFL geprägt ist (siehe dazu: DFL, 2019a). Die im Rahmen der vorliegenden Arbeit befragten Personen (Anzahl=7) sind sämtlich zum Zwecke der Unterstützung der Nachwuchsarbeit im deutschen Fußball bei DFB oder DFL angestellt. Sie stellen die Gesamtpopulation an Experten für Nachwuchsleistungszentren in den Verbänden des deutschen Profifußballs dar. Sie zeichnen sich durch ein spezielles Fachwissen über den Aufbau von und den Umgang mit der Lizenzierung, Zertifizierung und der Koordination von Nachwuchsleistungszentren in Deutschland aus. Nach persönlicher Kontaktaufnahme im Rahmen einer Diskussionsrunde mit DFB und DFL zum Thema „Qualitätsmanagement für Nachwuchsleistungszentren“ erklärten sie ihre Bereitschaft, an der Expertenbefragung teilzunehmen. Für diese Studie wurden die Experten lediglich zu *einem* Sachverhalt, und zwar zur Höhe des eingesetzten Budgets für alle gelisteten Nachwuchsleistungszentren (n=54) für die Saison 2019/2020 in EURO, befragt. Die Befragung fand mittels einer Online-Umfrage über die Plattform SoSci Survey (Näheres zur Plattform unter: [SoSci Survey ▶ professionelle Onlinebefragung Made in Germany](#)) vom 13.02.2019 bis 25.02.2019 statt. Die Online-Umfrage ist mit einer schriftlichen Befragung vergleichbar, allerdings wird die postalische Zustellung des Fragebogens durch das Internet ersetzt (Tausendpfund, 2018). Die Einladung zur Online-Befragung erfolgte nach der ersten persönlichen Kontaktaufnahme in einer Diskussionsrunde Anfang Februar 2019 per E-Mail (M. Honsel, Persönliche Mitteilung, 11. Februar 2019). Mit der Bearbeitung der Befragung durch 5 der 7 Experten für Nachwuchsarbeit beim DFB und der DFL wurde eine hohe Rücklaufquote (brutto wie netto) von 71 % (zur Ermittlung von Rücklaufquoten siehe AAPOR, 2021) und damit eine hohe Repräsentativität der befragten Personen sowie eine hohe generelle Datenqualität erreicht (Hinz, 2000; Schupp & Wolf, 2015). Darüber hinaus haben die antwortenden Experten für alle aufgeführten Nachwuchsleistungszentren (n=54) ihre Einschätzung der Höhe des Budgets für die Saison 2019/2020 abgegeben, es sind keine fehlenden Werte zu verzeichnen. Die Datenbasis ist somit als vollständig anzusehen (Döring & Bortz, 2016). Mit einem Aufwand jedes Befragten von ca. 30 Minuten für die Beantwortung der Frage nach der Höhe des Budgets für die gelisteten Nachwuchsleistungszentren können Messfehler aufgrund von zu schneller Beantwortung oder mangelnder Aufmerksamkeit ausgeschlossen werden (zur Qualitätsbewertung einer Umfrage siehe Weisberg, (2005) .

Die weitere Datenverarbeitung erfolgte über die von SoSci Survey bereitgestellten Tools (Download in Excel-Format). Datenverarbeitungsfehler sind ausgeschlossen, da die Datensätze ohne Kodierung, Bereinigung oder Transformation als Rohwerte zur Datenauswertung zunächst in einer Excel-Datei zusammengeführt und anschließend für die Datenanalyse nach SPSS übertragen wurden (Döring & Bortz, 2016).

Die Reliabilität des Umfrageergebnisses zur Höhe des Budgets für Nachwuchsleistungszentren 2019 wird anhand der Beobachterübereinstimmung (inter-rater-agreement) empirisch überprüft (Döring & Bortz, 2016). Wichtige Voraussetzung für die Zuverlässig-

keit und Aussagekraft der gewonnenen Daten ist die Übereinstimmung zwischen verschiedenen Beurteilern – im Falle dieser Studie die befragten Experten des DFB und der DFL (Wirtz & Kutschmann, 2007). Ist die Beobachterübereinstimmung hoch, ist anzunehmen, dass die zu beobachtenden Fälle messgenau sind (Döring & Bortz, 2016). Für die Berechnung der Beobachterübereinstimmung hinsichtlich kardinalskalierten Daten – wie die Höhe des Budgets für Nachwuchsleistungszentren in 2019 in Euro – wird vielfach die Intraklassenkorrelation (ICC) empfohlen (Shrout & Fleiss, 1979; Döring & Bortz, 2016; Söhnen, 2019). Die ICC wird in der Regel berechnet, wenn mehr als zwei Beobachter vorhanden sind oder/und zwei oder mehr Beobachtungszeitpunkte einbezogen werden sollen. Die ICC misst, ob und wie weit zwei oder mehr Beobachter innerhalb ihrer Einschätzungen hinsichtlich des Untersuchungsobjekts übereinstimmen, wenn sie einer Klasse angehören (Söhnen, 2019). Die ICC schätzt so die Reliabilität eines beliebigen Beobachters ein (Wirtz & Kutschmann, 2007).

Die Berechnung der ICC für die Expertenbefragung erfolgte mit 5 Experten (=unabhängige Variable (UV)) und 54 Untersuchungsobjekten (=Nachwuchsleistungszentren; abhängige Variable (AV)) in SPSS. Ab einem Wert von 0,7 kann lt. Greguras & Robie, 1998; nach Nunally, 1978, von einer hohen Übereinstimmung der Beobachter ausgegangen werden. Andere Autoren schlagen für eine Kategorisierung in eine schlechte, durchschnittliche, gute und sehr gute Bewertung der Reliabilität Wertintervall-Bereiche vor. Für Fleiss (2011) gilt eine ICC ab 0,60 als gut, größer gleich 0,75 als sehr gut, während eine neuere Einschätzung von Koo und Li (2016) erst bei größer gleich 0,9 von einer sehr guten Reliabilität ausgeht. Mit einer ICC von 0,932 und einem Konfidenzintervall von [0,892, 0,959] kann demnach die Zuverlässigkeit der Experten dieser Befragung als sehr gut eingeschätzt werden. Die Ergebnisse der Befragung werden damit als sehr reliabel interpretiert und für die Ermittlung des prozentualen Anteils des Investments am Umsatz des Vereins mit Nachwuchsleistungszentrum verwendet (siehe zur Ermittlung Investment Kapitel 3.6.1).

3.5.3 Güte der Datenerhebung

Als Zwischenfazit beziehungsweise methodische Konsequenz für die Untersuchung ist festzuhalten, dass eine Überprüfung der Daten anhand der drei Gütekriterien – Objektivität, Reliabilität und Validität – nicht für alle Komponenten des vorliegenden Modells en détail gleich sinnvoll war. Die Überprüfung der *Objektivität* im Rahmen der Datenerhebung bezieht sich auf die Durchführung, die Auswertung und die Interpretation der Erhebungen, auf eine statistische Prüfung wird meist verzichtet (Döring & Bortz, 2016; Krebs & Menold, 2014). Die Objektivität der Daten, die für diese Studie erhoben werden, wird durch die Dokumentation der Recherchemethoden und Quellenangaben, Offenlegung der Herleitung der verwendeten Indikatoren sowie der vollständigen Sicherung der Daten gewährleistet. Die Reliabilität (Zuverlässigkeit) und Validität (Gültigkeit) werden empirisch anhand von Kennwerten bzw. Koeffizienten geschätzt. Reliabilität und Validität setzen Objektivität voraus, eine hohe Reliabilität wird für eine hohe Vali-

dität vorausgesetzt (Döring & Bortz, 2016). Die verwendeten Sekundärdaten der öffentlich zugänglichen Quelle Transfermarkt.de sind im Hinblick auf ihre *Reliabilität* anhand veröffentlichter Studien zu diesem Thema überprüft worden. Auf Basis der Analyseergebnisse dieser Studien – Korrelations- und Regressionsanalysen – wird die Zuverlässigkeit der Daten von Transfermarkt.de als sehr hoch eingeschätzt und diese werden für weitere Datenanalysen in dieser Arbeit verwendet. Die Reliabilität der per Expertenbefragung erhobenen Primärdaten Budget 2019 wurde anhand der ausgewiesenen Intra-klassenkorrelation ($ICC = 0,932$) als sehr hoch interpretiert.

Eine Übersicht über die erhobenen Daten, die Instrumente und Konstrukte sowie die Beurteilung der Gütekriterien zeigt die Tabelle 6.

Die statistische Software wurde in Abhängigkeit vom Untersuchungsziel und der empirischen Prüfbarkeit von Aussagen gewählt. Zum Einsatz kommt SPSS 26.0, da hier ein umfangreiches Angebot an Features für die deskriptive Statistik, inferenzstatistische Prüfkriterien und Tests zur Verfügung steht. Die Methoden sind zur Gewinnung der angestrebten Ergebnisse in der vorliegenden Studie unerlässlich und dienen zur Überprüfung der Reliabilität und der Validität der Daten.

3.6 Ermittlung der Datenkonstrukte

3.6.1 Investment

Dem Datenkonstrukt *Investment in ein Nachwuchsleistungszentrum* liegt die Annahme zugrunde, dass die Höhe über einen prozentualen Anteil am Rohergebnis dargestellt werden kann. Für die Ermittlung des Konstrukts Investment in ein Nachwuchsleistungszentrum wurden daher die Finanzdaten *Rohergebnis Verein* und *Budget 2019* erhoben. Das Ergebnis der Überprüfung der Güte der eingehenden Rohwerte erlaubt ihre Verwendbarkeit für die Ermittlung des Konstrukts. Im Folgenden wird zunächst der Prozess zur Ermittlung der Höhe des prozentualen Anteils am Rohergebnis eines Vereins beschrieben und im Anschluss daran die Verlässlichkeit des Konstrukts untersucht.

3.6.1.1 Ermittlung Prozentualer Anteil Investment in ein NLZ am Rohergebnis

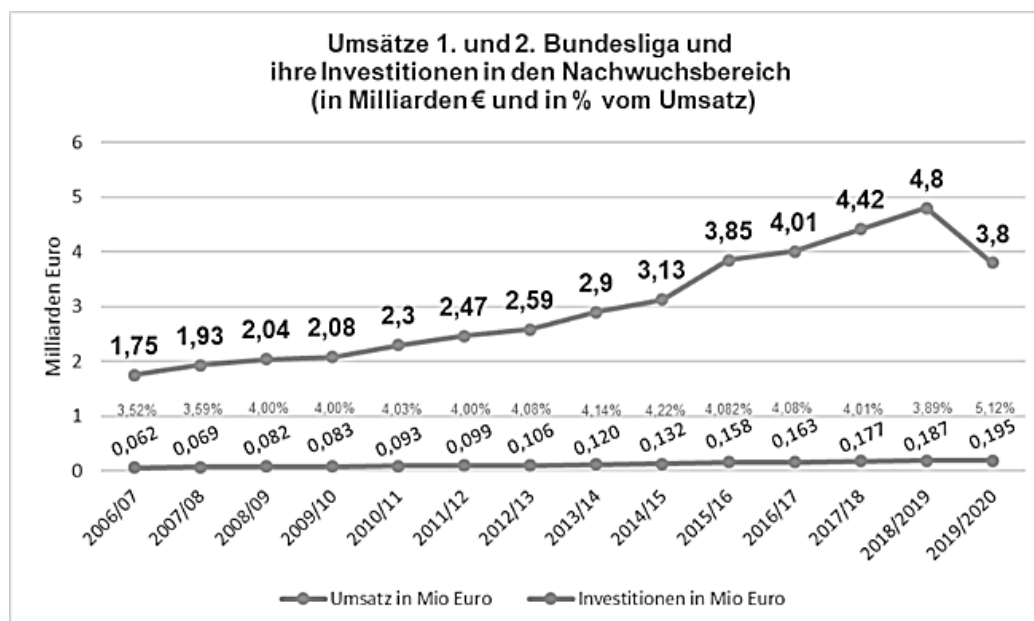
Die DFL weist in ihren *Wirtschaftsreports* regelmäßig Umsatzzahlen und Investitionen in Nachwuchsleistungszentren in Euro insgesamt für die Vereine der 1. und 2. Bundesliga pro Saison aus (DFL, 2020a). Die Abbildung 35 zeigt die publizierten Daten *Umsatz* und Investitionen sowie die prozentualen Anteile der Investitionen am Umsatz pro Saison.

Die gezeigte Entwicklung der Umsätze und der Investitionen in den Nachwuchsbereich über die Saisons 2003/04 bis 2019/20 beziehen sich auf die Gesamtheit der Vereine der 1. und 2. Bundesliga, die Nachwuchsleistungszentren der Liga 3 sind nicht enthalten. Durch den Ausweis der signifikanten Korrelation zwischen Umsatz und Investition nach Pearson ($r(13) = 0,96, p = 0,000$) wird oben getroffene Annahme gestützt.

Tabelle 6: Übersicht zur Datenerhebung für die vorliegende Studie

Datengrundlage	Datenerhebung, -herkunft	Daten	Prüfung Gütekriterien			
			Ansatz	Ergebnis	Autor	
Sekundärdaten	Transfermarkt.de	Liga	Fehlerfreiheit "Alles-oder-Nichts-Ansatz",	max. Fehlerquote von 6,85% (vs. Daten Polen), 2,71% (vs. Daten Premier League) sowie 1,87% (vs. Daten MLS)	Kasperek (2020) Strauß (2020) Bialas (2021)	
		Punktzahl des Vereins				
		Punktzahl des Meisters der letzten 5 Jahre	Vollständigkeitsprüfung	94,1% (Premier League), 99,91% (MLS)	Strauß (2020) Bialas 2021	
		Einsätze Spieler				
		max Einsätze				
		Transfers Spieler	Korrelation Marktwert und erzielte Transfersummen	$r = 0,93$	Gerhards et al. (2014)	
		Einfluss Spielermerkmale auf Marktwert	$R^2 = 0,90, b = 0,95, p < 0,01$	Herrn et al. (2014)		
	Marktwert	Korrelation Marktwerte in der Saison und Tordifferenz	$r = 0,67$	Heuer (2012)		
		Korrelation Marktwerte <i>transfermarkt.de</i> und Marktwerte <i>Kicker Sportmagazin</i>	$r = 0,75$	Torgler (2006)		
	DFB / DFL	Double Pass Scores	Fakten; keine Prüfung auf Gütekriterien; Einsatz der Daten für Validitätsprüfung mit "Hervorgebrachter Spielerqualität NLZ"			
Primärdaten	Geschäftsberichte u.ä.	Rohergebnisse	Imputation fehlender Daten erforderlich	für N=54 und 11 Jahre Beobachtungszeitraum: 316 Datenpunkte erhoben, 278 Datenpunkte missing von max 594 Datenpunkte. Fallreduktion um 12 Vereine: 132 Datenpunkte. Imputation von 140 Datenpunkten. Datenmatrix vollständig 462 Datenpunkte für n=42.		
				Honsel (in dieser Arbeit)		
	Expertenbefragung	Budget 2019	Rücklaufquote		71%	
			Vollständigkeit		100%	
			ICC	0,932		
	Instrumente / Konstrukte	Vereinsqualität	Vereinsqualitäten im Vergleich zur Tabellensituation folgender Ligen: im deutschen Profifußball 1. Bundesliga, 2. Bundesliga, 3., 4. und 5. Liga in Deutschland, La Liga, 2. und 3. Spanische Liga, Premier League, 2. Liga England über 16 Saisons	„Die Vereine unterscheiden sich über die erreichte Punktzahl sinnvoll innerhalb einer Liga (Mittelwert (M) VQ= 3,99; Standardabweichung (SD) VQ= 0,89)“.		Lanwehr et al. (2021)
				Spielerqualität	Zusammenhang individuelle Spielerqualität mit den Marktwerten der Spieler	
		Zusammenhang Mittelwerte Individuelle Spielerqualität mit den Mittelwert Marktwert der Spieler	Pearson $r=0,7662$			
		Hervorgebrachte Spielerqualität NLZ	Zusammenhang mit Double Pass Scores	Spearman's $\rho = 0,600, p < 0,001$, bei N=171 Wertepaaren; Kendall-Tau-b und -c Werte von $\rho = 0,424, p < 0,001$, bei N=171 (b und c-Werte sind identisch)		Honsel (in dieser Arbeit)
	Investment in NLZ	Zusammenhang Budgetschätzungen für 2019 mit Rohergebnis 2018/19	Pearson $r=0,898$ Spearman $r = 0,894$			

Abbildung 35: Umsatzentwicklung und Investitionen in den Nachwuchsbereich 1. und 2. Bundesliga in Mrd. Euro und prozentualer Anteil der Investitionen am Umsatz. Eigene Darstellung nach DFL (2020a).



Die Analyse der prozentualen Anteile des Investments am Umsatz der 1. und 2. Bundesliga pro Saison zeigt, dass das *Investment in Nachwuchsleistungszentren* in Euro über alle 13 Saisons durchschnittlich **4,10 %** des Umsatzes beträgt ($M = 4,10 \%$, $SD = 0,003$). Der prozentuale Anteil von 5,12 % für die Saison 2019/20 scheint keine überproportionale Erhöhung der Investition in den Nachwuchsbereich zu sein, sondern ist begründet in der Beibehaltung der Höhe der geplanten Investitionssumme trotz coronabedingtem Umsatzeinbruch um 1,0 Mrd. Euro. Wird dieser Wert als Ausreißer interpretiert und geht nicht in die Berechnung des durchschnittlichen Prozentsatzes ein, ergibt sich ein prozentualer Anteil von **4,01 %** ($M = 4,01$, $SD = 0,002$).

3.6.2.2 Verlässlichkeit des Konstrukts Investment in Nachwuchsleistungszentren

Um sicherzustellen, dass in der Untersuchung verlässliche Konstrukte verwendet werden, sind neben der Objektivität die Reliabilität und Validität der verwendeten Konstrukte zu überprüfen (Weiber & Mühlhaus, 2014). Die Überprüfung der *Objektivität* bezieht sich auf die Durchführung, die Auswertung und die Interpretation der angewandten Methode. Die Methode ist als objektiv einzustufen, wenn das Vorgehen und das erhaltene Ergebnis intersubjektiv nachvollziehbar sind. Die Ermittlung des Grads der Übereinstimmung zwischen verschiedenen Forschern basiert häufig auf Korrelationen zwischen den Ergebnissen der einzelnen Beurteiler (Helfrich, 2016). Die Feststellung der Reliabilität und Validität gehört zu den komplexesten und aufwändigsten Verfahren. Bewertet wird die Güte der im Modell abgebildeten Komponenten beziehungsweise Konstrukte, die im Falle der Erhebung quantitativer Daten eine Reihe von Tests mit Hilfe komplexer statistischer Auswertungsverfahren erfordern (Schnell et al., 2011).

Mit *Reliabilität* wird die Zuverlässigkeit oder Verlässlichkeit eines Instruments bezeichnet und damit das Ausmaß, in dem wiederholte Messungen eines Objekts mit ein und demselben Instrument die gleichen Werte liefern (Döring & Bortz, 2016; Helfrich, 2016). Mit der *Validität* wird die Gültigkeit eines Messinstruments angegeben. Es ist dann valide, wenn mit ihm das erfasst wird, was erfasst werden soll (Helfrich, 2016). Zur Beurteilung der Validität der Messinstrumente werden verschiedene (Validitäts-)Aspekte herangezogen:

Durch die *Inhaltsvalidität* wird geprüft, ob das Messinstrument sachlich und logisch geeignet ist, das in Frage stehende Konstrukt darzustellen (Helfrich, 2016). Es gilt, möglichst alle Aussagen des theoretischen Begriffs in die Operationalisierung einzubeziehen. Es kommt darauf an, dass die zu messende Eigenschaft inhaltlich-semantic und/oder empirisch durch die Indikatoren repräsentiert wird (Kromrey, 2002). Die *Kriteriumsvalidität* liegt vor, sobald die Messergebnisse eines Faktors mit denen eines anderen relevanten Merkmals (Außenkriterium) empirisch übereinstimmen, d. h., korrelieren (Helfrich, 2016). Mit der *Konstruktvalidität* wird überprüft, inwiefern die durch einen Test erhobenen Variablen auch tatsächlich das übergeordnete, zugrunde gelegte Konstrukt und nicht irgendein anderes Konstrukt messen. Die *Konvergenzvalidität* fordert ausreichend hohe Zusammenhänge zwischen den jeweiligen Indikatoren, die dem gleichen Konstrukt zugeordnet sind. Ebenso müssen Faktoren, die zu derselben Dimension verdichtet werden, hohe Korrelationen aufzeigen (Hornburg & Giering, 1996).

Um den vorliegenden Sachverhalt erklären zu können, ist zunächst zu prüfen, ob es einen Zusammenhang zwischen den verwendeten Merkmalen gibt (Helfrich, 2016). Mittels einer Zusammenhangsanalyse (Korrelationsanalyse) wird festgestellt, ob und in welchem Grad eine lineare Beziehung zwischen den Merkmalen X (Prädiktor) und Y (Kriterium) besteht (Döring & Bortz, 2016; Helfrich, 2016). Das Vorzeichen des berechneten Korrelationskoeffizienten bestimmt die Richtung des Zusammenhangs (positiv oder negativ), die Höhe des Korrelationskoeffizienten stellt die Stärke des Zusammenhangs dar (schwach, stark). Das deskriptiv-statistische Ergebnis der Berechnung des Korrelationskoeffizienten ist anschließend einem Signifikanztest zu unterziehen, um die Hypothesen zu überprüfen (Döring & Bortz, 2016). Die Korrelationsart, mit der die Hypothesenüberprüfung vorgenommen wird, wird bestimmt durch die Anzahl der Variablen und des Skalenniveaus der betrachteten Variablen (Döring & Bortz, 2016; Helfrich, 2016). Der Zusammenhang zwischen zwei Variablen wird durch eine bivariate Korrelation repräsentiert, der Korrelationskoeffizient kann einen Wert zwischen -1 und +1 annehmen, wobei -1 den perfekten negativen Zusammenhang und die +1 den stärksten positiven Zusammenhang repräsentieren. Bei Werten um Null ist kein systematischer Zusammenhang zu verzeichnen (Döring & Bortz, 2016). Die Kombination der Skalenniveaus der zwei zu untersuchenden Variablen bestimmen das einzusetzende Assoziationsmaß (Döring & Bortz, 2016 nach Eid et al., 2010, S. 539). Zusammenhangshypothesen für zwei kardinalskalierte Variablen werden in der Regel mittels der Pearson-Korrelation getestet. Voraussetzungen für den Pearson-Korrelationstest sind ein linearer Zusammenhang zwischen den Variablen, eine bivariate Normalverteilung der Variablen,

zwei metrische Skalen und keine Ausreißer bei den Merkmalsausprägungen (Bortz & Schuster, 2016) Die lineare Beziehung zwischen den Variablen kann grafisch mittels eines Streudiagramms oder über die Schiefe der Statistik bestimmt werden (Hanna & Dempster, 2017). Normalität wird bei einem Stichprobenumfang von n größer als 30 angenommen, bei einer Stichprobengröße von unter 30 ist eine Überprüfung der Normalität erforderlich (Bortz & Schuster, 2016) Üblicherweise erfolgt die Prüfung auf Normalität mit Hilfe des Kolmogorov-Smirnov-Anpassungstests und des Shapiro-Wilke-Tests (Bühl, 2019).

Gepprüft werden die Daten *Budget 2019 und Rohergebnis Verein 2018/19* mit einer Stichprobe aus Vereinen der 1. und 2. Bundesliga sowie der Liga 3 ($n=42$). Der vorliegende Datensatz wurde vor der Berechnung eines Korrelationskoeffizienten zunächst auf die Voraussetzungen für den Pearson-Korrelationstest auf Linearität, fehlende Werte, Ausreißer und dessen Verteilungsart inspiziert. Anhand der grafischen Analyse mittels Streudiagramm konnte keine Linearität festgestellt werden, die Datenpaare waren vollständig, ein Ausreißer auf Seiten des Rohergebnis Verein 2018/19 wurde eliminiert ($n=41$). Eine Prüfung der Datensätze auf Normalität erfolgte mittels des Kolmogorov-Smirnov-Anpassungstests. Sowohl der Datensatz *Budget 2019* und *Rohergebnis Verein mit NLZ 2018/19* sind nicht hinreichend normalverteilt (*Budget 2019*: M: 4.130.952,38; SD: 3.324.373,26; $p=3,1255E-7$; *Rohergebnis Vereine 2018/19*: M: 94.938.073,17; SD: 98.006.776,54; $p=0,002$). Bestätigt wird das Fehlen der Normalverteilung durch Interpretation der Normalverteilungsdiagramme (Q-Q-Diagramm) und der trendbereinigten Normalverteilungsdiagramme (trendbereinigtes Q-Q-Diagramm). Es wird empfohlen, bei nicht-normalverteilten Daten die Korrelation mithilfe von Bootstrapping (mit zurücklegen) den Pearson-Korrelationskoeffizienten zu berechnen oder auf nicht-parametrische Verfahren wie den Spearman-Korrelationskoeffizienten zurückzugreifen (Bortz & Lienert, 2008). Der Pearson-Korrelationskoeffizient von $r_p = 0,898$ mit einer Signifikanz von $p_p = 0,000$ und der Spearman-Korrelationskoeffizient von $r_{sp} = 0,894$ und $p_{sp} = 0,000$ können als ein sehr starker Zusammenhang zwischen *Budget 2019* und *Rohergebnis Verein mit NLZ 2018/19* interpretiert werden. Bei Betrachtung der unteren Grenzwerte der Konfidenzintervalle für beide berechneten Korrelationswerte (Pearson: $r = [0,826, 0,955]$; Spearman: $r = [0,778, 0,944]$) ist festzustellen, dass sie in das Intervall $> 0,7 < 1$ für einen sehr starken Zusammenhang fallen und von einer hohen Validität der erhobenen Daten zeugen. Die Daten *Budget 2019* und *Rohergebnis Verein mit NLZ 2019* können daher in die Generierung des Investments in Nachwuchsleistungszentren eingehen.

Der nächste Schritt zur Generierung des Indikators ist die Ermittlung und Beschreibung des Investments in NLZ 2019 als prozentualen Anteil des Budgets 2019 am Rohergebnis Verein mit NLZ 2019. Die Interpretation und Beurteilung der Daten erfolgt über die Maße der zentralen Tendenz (Lagemaß) und die Maße der Variabilität (Streumaß, (Bortz & Schuster, 2016; Hanna & Dempster, 2017). Das gebräuchlichste Maß zur Darstellung der zentralen Tendenz ist der Mittelwert. Der Mittelwert als guter Schätzer des

Zentrums einer Verteilung ist sensitiv gegenüber untypischen Beobachtungen, sogenannten Ausreißern oder Extremwerten (Bortz & Schuster, 2016). Darüber hinaus gilt der Median als weiteres häufig verwendetes Lagemaß. Vermutet oder identifiziert man untypische Beobachtungen in den Daten, sollte der Median als Lagemaß verwendet werden, da der Median im Vergleich zum Mittelwert nicht oder nur wenig von Ausreißern beeinflusst wird (Bortz & Schuster, 2016). Eine Erklärung der Unterschiedlichkeiten des untersuchten Merkmals wird mittels verschiedener Variabilitätsmaße gegeben. Das gebräuchlichste Maß für die Streuung ist die Standardabweichung, sie informiert darüber, wie stark die Werte des untersuchten Merkmals vom Durchschnittswert abweichen (Hanna & Dempster, 2017). Mit einer Überprüfung des aufgetretenen Unterschieds der Mittelwerte kann objektiv unterschieden werden, ob ein aufgetretener Unterschied der Mittelwerte zufällig zustande gekommen ist oder nicht (Bühl, 2019). In Abhängigkeit von der Art und Anzahl der zu vergleichenden Stichproben, der Verteilung und des Skalenniveaus der Daten werden unterschiedliche Tests eingesetzt.

Die Lage- und Streumaße für das Verhältnis von *Budget 2019 am Rohergebnis Verein mit NLZ 2018/19* wurden in SPSS ermittelt. Es wurden drei untypische Beobachtungen identifiziert und aus dem Datensatz entfernt, weitere Berechnungen erfolgten auf Basis des Datensatzes Exklusive Ausreißer ($n=38$). Die Prüfung des Datensatzes mittels des Kolmogorov-Smirnov-Anpassungstest zeigt, dass die Daten normalverteilt sind ($p=0,095$). Der Mittelwert liegt bei $M = 0,0500$ (5,0 % Anteil des Budgets 2019 am Rohergebnis Verein mit NLZ) bei einer Standardabweichung von $SD = 0,0176$. Mit 95 % Sicherheit liegt der Mittelwert innerhalb des Konfidenzintervalls von $[0,0442; 0,0558]$. Eine Übersicht der Lage- und Streumaße des untersuchten Datensatzes zeigt die Tabelle 7.

Die Prüfgröße des Einstichproben-t-Tests weist mit ($T = 3,460$; $df = 37$; $p=0,014$, Signifikanzniveau von 0,01) aus, sodass der Mittelwert des Datensatzes *Prozentualer Anteil Budget 2019 an Rohergebnis Verein mit NLZ* ($M = 0,0500$; $SD = 0,0176$) signifikant unterschiedlich ist zum Mittelwert des Datensatzes *Investment in NLZ als Anteil des Umsatzes der 1. und 2. Bundesliga* ($M = 0,0410$; $SD = 0,002$). Die Nullhypothese – kein Unterschied in den Mittelwerten – wird nicht bestätigt.

Es gibt einen signifikanten Unterschied im Vergleich der bisher ermittelten Mittelwerte. Um aber einen möglichst realitätsnahen prozentualen Anteil des Budgets am Rohergebnis Verein mit NLZ zu schätzen, erfolgt zunächst eine tiefergehende Analyse des prozentualen Anteils anhand der Liga-Zugehörigkeit der Vereine 2018/19. Es wird mittels einfaktorieller Varianzanalyse untersucht, ob sich die Mittelwerte in den drei betrachteten Ligen (1. und 2. Bundesliga und Liga 3) signifikant unterscheiden. Dazu werden in einem ersten Analyseschritt unterschiedliche Varianzen ausgewiesen (1. Bundesliga: $S^2=1,210$, 2. Bundesliga: $S^2=1,710$ und Liga 3: $S^2=0,974$), sodass im zweiten Analyseschritt der Test zur Ermittlung der Signifikanz anhand des Welch-Tests (bekannt als ein modifizierter t-Test, bei dem die Varianz der Stichproben-Mittelwertdifferenz aus den Stichprobenvarianzen ermittelt wird, Bortz & Schuster, 2010) erfolgt. Das Ergebnis

weist mit $p=0,001$ hohe signifikante Unterschiede der Mittelwerte aus. In der 1. Bundesliga liegt der Mittelwert bei 3,69 %, in der 2. Bundesliga bei 5,75 % und in Liga 3 bei 7,92 %.

Tabelle 7: Deskriptive Statistik Datensatz Prozentualer Anteil Budget 2019 und Rohergebnis Verein mit NLZ 2018/19

Prozentualer Anteil Budget 2019 an Rohergebnis Verein mit NLZ 2018/19			
Deskriptive Statistik			Standard Fehler
Mittelwert		0,0500	0,0029
95% Konfidenzintervall des Mittelwerts	Untergrenze	0,0442	
	Obergrenze	0,0558	
5% getrimmtes Mittel		0,0497	
Median		0,0488	
Varianz		3,1078	
Standard Abweichung		0,0176	
Minimum		0,0176	
Maximum		0,0881	
Spannweite		0,0705	
Interquartilbereich		0,0190	
Schiefe		0,4269	0,3828
Kurtosis		-0,2376	0,7497

Der H-Test nach Kruskal und Wallis, eine Ausweitung des U-Tests von Mann und Whitney bei mehr als 2 unabhängigen Stichproben, weist ein höchst signifikantes Ergebnis aus ($H = 20,787$; $df=2$; $p < 0,001$) (Bühl, 2019). Die Nullhypothese, alle Ligen seien bezüglich des Mittelwerts gleich, ist daher auch nach Kruskal und Wallis zu verwerfen.

Durch den paarweisen Vergleich der Ligen untereinander zeigen sich signifikante Unterschiede zwischen der 1. und 2. Bundesliga ($p = 0,000$) und der 1. Bundesliga und der Liga 3 ($p = 0,001$). Der Unterschied der Mittelwerte zwischen 2. Bundesliga und Liga 3 ist nicht signifikant ($p = 0,461$).

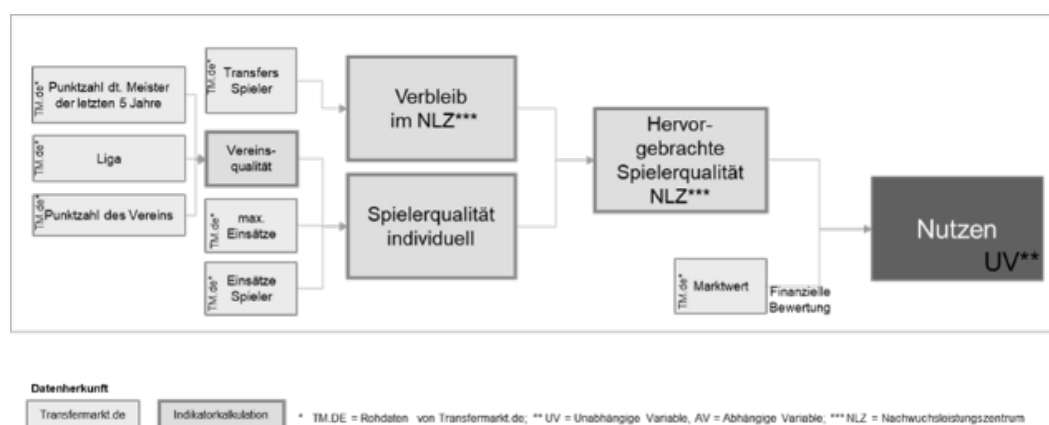
Die Gruppierungen mit signifikanten Unterschieden bei paarweisem Vergleich weisen starke zentrale Tendenzunterschiede aus. Nach Cohen (1988) liegen die Effektgrenzen bei 0,1-0,3 (schwach), 0,3-0,5 (mittel) und größer 0,5 (stark). Die Effektstärke r für den

Vergleich der 1. und 2. Bundesliga liegt bei $r=0,6425$, für die 1. Bundesliga und Liga 3 bei $r=0,7768$. Die Unterschiede der Mittelwerte der drei betrachteten Ligen sind sehr stark, sodass der Mittelwert über alle Vereine in der Stichprobe ($M = 0,500$) nicht für die Ermittlung der Investition für alle Vereine verwendet werden kann. Unter der Annahme, dass sich in der Art der Beziehung der Variablen über die Jahre nichts geändert hat, kann der Indikator – *Investment in NLZ* – für den Bereich der empirischen Beobachtungen hinaus geschätzt werden (Kromrey, 2002). Das *Investment in Nachwuchsleistungszentren* wird für diese Studie daher mit den ermittelten ligenspezifischen prozentualen Anteilen am Rohergebnis des Vereins wie folgt zum Ansatz gebracht: 1. Bundesliga: 3,69 %, 2. Bundesliga: 5,75 %, Liga 3: 7,92 %, vom *Rohergebnis Verein mit NLZ* pro Jahr ermittelt.

3.6.2 Nutzen

Ziel dieser Studie ist es, anhand der betriebswirtschaftlichen Kennzahl *ROI* den Erfolg einzelner Nachwuchsleistungszentren zu ermitteln. Wesentliches Element der Kalkulation des *ROI* ist, neben den Kosten, das Nutzelement des *ROI*. Der Nutzen wird für diese Studie über die monetäre Bewertung des Konstrukts *Hervorgebrachte Spielerqualität NLZ* operationalisiert. In diesem Kapitel erfolgt eine Beschreibung und Erläuterung der Entwicklung der Formeln der zum Konstrukt führenden Elemente – *Vereinsqualität*, *Verbleib im Nachwuchsleistungszentrum*, *Individuelle Spielerqualität* – und des Konstrukts *Hervorgebrachte Spielerqualität NLZ* insgesamt. Abbildung 36 gibt einen Überblick über die Zusammensetzung des Konstrukts sowie die Herkunft der zugrundeliegenden Daten.

Abbildung 36: Überblick Konstrukt „Nutzen für ROI für NLZ“.



Die Entwicklung des Konstrukts *Nutzen NLZ* erfolgt in mehreren Schritten: Zunächst wird ein Indikator *Vereinsqualität (VQ)* entwickelt, der die Qualität der Vereine abbildet und ein Ranking der Vereine über alle deutschen Ligen hinweg abbildet. Im Anschluss daran erfolgt die Entwicklung des Indikators *Individuelle Spielerqualität (SQ_S)*, der über die Variablen *Vereinsqualität*, *Anzahl Einsätze des Spielers* sowie des Korrektivs maximale

Einsätze (= maximale Anzahl Einsätze eines Spielers über alle Wettbewerbe hinweg, an denen ein Verein teilnimmt) die Qualität eines aus dem Nachwuchsleistungszentrum hervorgebrachten Spielers misst. Der Erfolg des Nachwuchsleistungszentrums wird mittels des Konstrukts *Hervorgebrachte Spielerqualität NLZ* gemessen. Die Summe der anteiligen Ausbildung über alle betrachteten Saisons (tagesgenau) – dargestellt über *Verbleib im Nachwuchsleistungszentrum (AU)* – aller Spieler im NLZ, die später Profifußballspieler geworden sind, gewichtet mit der *Individuellen Spielerqualität* der einzelnen Spieler ergibt die *Hervorgebrachte Spielerqualität NLZ (SQ_{NLZ})*. Eine Bewertung der *Hervorgebrachten Spielerqualität NLZ* mit mittleren Marktwerten der Spieler pro Spielerqualitäts-Punkt (Marktwert-Faktor) pro Jahr ergibt den Nutzen in Geldeinheiten.

Im Folgenden werden die Schritte zur Berechnung des Nutzens – Vereinsqualität, Individuelle Spielerqualität, hervorgebrachte Spielerqualität, monetäre Bewertung – detaillierter beschrieben.

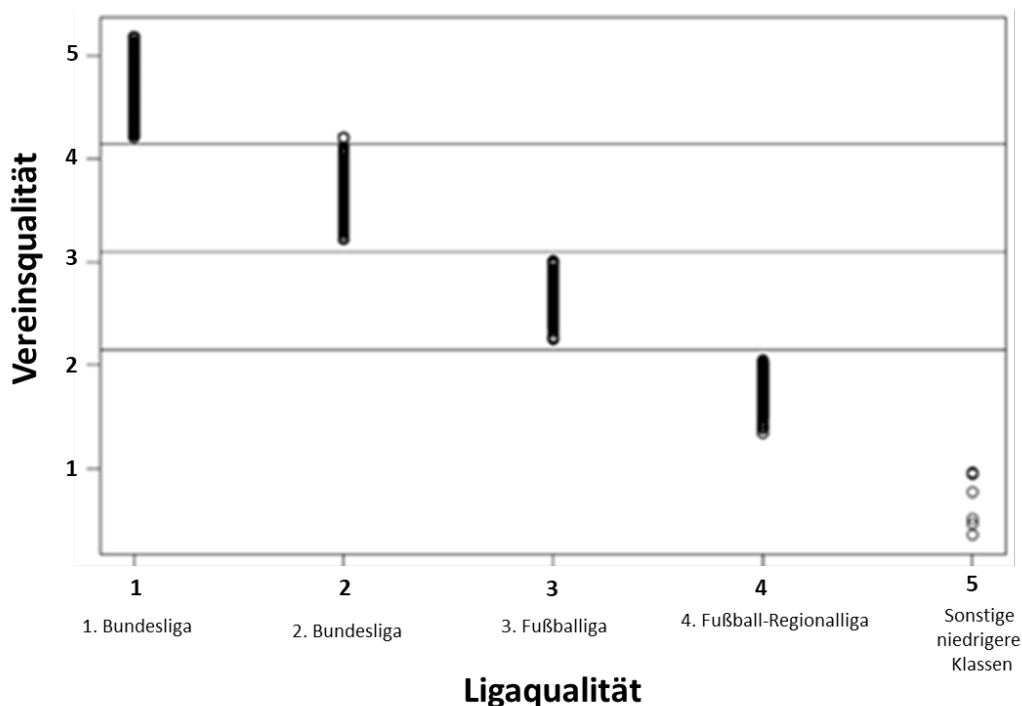
3.6.2.1 Vereinsqualität

Der Indikator *Vereinsqualität (VQ)* misst die Qualität der Vereine und erlaubt ein Ranking der Vereine über alle deutschen Ligen. Abbildung 37 zeigt die jeweilige Liga in Relation zur Vereinsqualität der betrachteten Vereine. Der Plot zeigt, dass die Vereinsqualität in hochrangigen Ligen (z. B. 1. Bundesliga) am höchsten ist und die Qualität der besten Vereine der zweiten Liga ca. der der schlechtesten Vereine der ersten Liga entspricht (Lanwehr et al., 2021). „Die Vereine unterscheiden sich über die erreichte Punktzahl sinnvoll innerhalb einer Liga (Mittelwert (M) $VQ=3,99$; Standardabweichung (SD) $VQ=0,89$)“ (Lanwehr et al., 2021, S. 30).

Zur Schätzung der Vereinsqualität (VQ) wird die Punktzahl des Vereins (PZ_i) zu den durchschnittlich erreichten Punkten der deutschen Meister über 5 Jahre in Relation gesetzt. Um eine Wertigkeit der Ligen, in der ein Verein in der betrachteten Saison spielt, abzubilden, erfolgt die Berücksichtigung der Ligaqualität. Die *Ligaqualität (LQ)* kann die Ausprägungen 1 bis 5 annehmen. Sie repräsentiert das Ranking der Ligen im deutschen Profifußball: die 1. Bundesliga ist mit dem Wert 1, die 2. Bundesliga mit dem Wert 2, die Liga 3 mit dem Wert 3, die Regionalligen mit dem Wert 4 und alle anderen mit dem Wert 5 versehen.

Die Ligaqualität stellt ein Korrektiv dar, um die Relation der Punktzahl des Vereins zu den durchschnittlichen Punkten des deutschen Meisters über die letzten 5 Jahre ($dPZDM5$) einzuschätzen. Die erreichte Punktzahl des Vereins pro Saison in der Liga bildet das Maß des Erfolgs. Die durchschnittliche Punktzahl des deutschen Meisters über die letzten 5 Jahre beträgt zum Beispiel für die Saisons 13/14 bis 17/18 84,6. Die Berechnung der Vereinsqualität ist mit der Formel 2 dargestellt.

Abbildung 37: Vereinsqualitäten im Vergleich zur Tabellensituation: 1. Bundesliga, 2. Bundesliga, 3., 4. und 5. Liga in Deutschland, La Liga, 2. und 3. Spanische Liga, Premier League, 2. Liga England über 16 Saisons. Nach Lanwehr et al. (2021).



Formel 2: Berechnung Vereinsqualität – VQ. Eigene Darstellung nach Lanwehr et al. (2021, S. 31).

$$VQ_v = (5 - LQ) + \frac{PZv}{dPZDM5}$$

VQ_v = Vereinsqualität, LQ = Ligaqualität (Ausprägungen 1-5), PZv = Punktzahl des Vereins, $dPZDM5$ = Ø Punkte Deutscher Meister über 5 Jahre.

3.6.2.2 Individuelle Spielerqualität

Der Indikator *Individuelle Spielerqualität* misst die Qualität eines Spielers. Um die von einem Nachwuchsleistungszentrum hervorgebrachten Spielerqualitäten zu ermitteln, wird von zwei Annahmen ausgegangen: Erstens haben gute Spieler eine höhere Anzahl an Einsätzen als weniger gute Spieler und zweitens spielen gute Spieler häufiger bei Vereinen mit höherer Qualität. Die Variablen *Vereinsqualität* (VQ_v) und *Anzahl Gespielte Spiele in der Saison* (GS_s) gehen daher in die Kalkulation des Indikators *Individuelle Spielerqualität* (SQ_s) ein. Um auch Spielersituationen abzubilden, in der sehr gute Spieler bei Vereinen mit hoher Vereinsqualität aufgrund hoher Konkurrenz nicht häufig oder gar nicht zum Einsatz kommen, wird die Vereinsqualität quadriert und ist damit entscheidender als die Anzahl der Einsätze, die ein Spieler hat. Da die Anzahl der Einsätze pro Spieler sehr hohe Zahlen annehmen kann, erfolgt eine Normalisierung dieser Werte mit der Anzahl

der maximalen Einsätze eines Spielers in der Saison. Für den Fall, dass ein Spieler keinen Einsatz hatte (d. h., der Bruch würde den Wert Null annehmen), wird über die Addition des Verhältnisses der Anzahl der Vereine, bei der der Spieler gespielt hat, verhindert, dass die Spielerqualität mit Null ausgewiesen wird. Spieler, die keinen Einsatz hatten, erhalten so mindestens die Spielerqualität der quadratischen Vereinsqualität (Lanwehr et al., 2021).

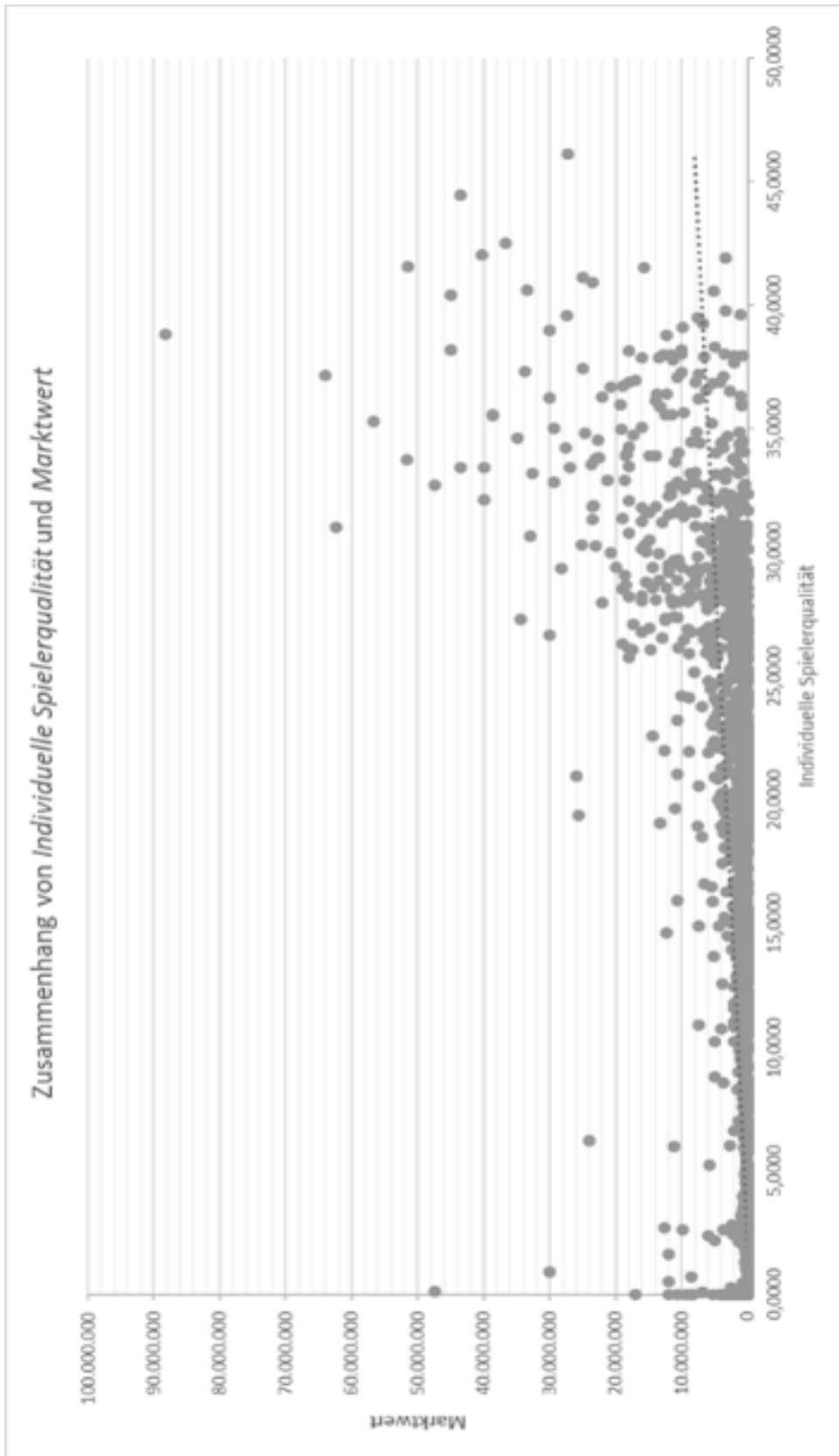
Formel 3: Berechnung Individuelle Spielerqualität. Eigene Darstellung, Weiterentwicklung nach Lanwehr et al. (2021, S. 31).

$$SQ_s = \sum_{\text{Alle Vereine}} VQ_s^2 \times \left(\frac{GS_s}{MAX(GS)} + \frac{1}{AnzV} \right)$$

SQ_s = Individuelle Spielerqualität, VQ = Vereinsqualität, GS_s = Gespielte Spiele Spieler, $Max (GS)$ = Maximal gespielte Spiele in der Saison, $AnzV$ = Anzahl Vereine

Eine Prüfung der Validität wurde anhand des Zusammenhangs der *individuellen Spielerqualitäten* mit den *Marktwerten* der Spieler (Datenherkunft: transfermarkt.de) vorgenommen. Es besteht ein starker Zusammenhang zwischen den beiden Variablen (Spearman's $\rho = 0,676$, $p < 0,001$, bei $N=2.627$ Wertepaaren). Dies bescheinigt dem Konstrukt *Individuelle Spielerqualität* eine hohe Validität: Es misst das, was gemessen werden sollte (Interpretation der Höhe des Korrelationskoeffizienten lt. Cohen, 1988). Mit hoher Messgenauigkeit gehen die ermittelten Spielerqualitäten in die Berechnung des im Folgenden erläuterten Konstrukts *Hervorgebrachte Spielerqualität Nachwuchsleistungszentrum* ein. Die grafische Darstellung des Zusammenhangs zeigt Abbildung 38.

Abbildung 38: Streudiagramm von Individuelle Spielerqualität und Marktwerte der Spieler.



3.6.2.3 Hervorgebrachte Spielerqualität Nachwuchsleistungszentrum

Das Konstrukt *Hervorgebrachte Spielerqualität Nachwuchsleistungszentrum* (*Spielerqualität NLZ*) repräsentiert laut dem zugrundeliegenden mentalen Modell den Nutzen, den ein Nachwuchsleistungszentrum mit der Ausbildung von Profifußballspielern erzeugt hat. In Abhängigkeit der Summe der Tage in Ausbildung aller Spieler im Nachwuchsleistungszentrum über alle betrachteten Saisons hinweg, gewichtet mit der durchschnittlichen individuellen Spielerqualität, ergibt sich die Spielerqualität NLZ.

Um die *Individuelle Spielerqualität* auf die Nachwuchsleistungszentren beziehen zu können, sind diverse Schritte erforderlich. In einem ersten Schritt erfolgt die Ermittlung der Variable *Verbleib im NLZ* ($AU_{i,NLZ,j}$). Sie stellt dar, wie viel Zeit für die Ausbildung eines

Spielers bei einem Verein aufgewendet worden ist. Die im Nachwuchsleistungszentrum verbrachte Zeit wird auf Basis der Informationen des Transfers der Spieler – Datum abgebender Verein, Datum aufnehmender Verein in der Alterszeitspanne des Spielers von 13 bis 19 Jahren – tagesgenau berechnet. Die Berücksichtigung der Tagesgenauigkeit stellt eine Präzisierung der Berechnung der hervorgebrachten Spielerqualität gegenüber der ursprünglich durch Lanwehr et al. (2021) publizierten Formel dar (nur Halbsaison orientiert). Im zweiten Schritt wird die mittlere Individuelle Spielerqualität berechnet, da sich die Variable pro Saison je Profifußballspieler verändert (vgl. dazu Formel 3 im vorherigen Abschnitt). Dies führt zur Aggregation der ermittelten Individuellen Spielerqualitäten auf einen Wert – die durchschnittliche Spielerqualität (\overline{SQ}_s) bis Stichjahr in Tagen. Darauf folgt im dritten Schritt die Multiplikation des Anteils an der Ausbildung eines Spielers mit der durchschnittlichen individuellen Spielerqualität. Da die Ausbildungslänge der Spieler variiert, wird das Produkt aus Verbleib im Nachwuchsleistungszentrum (Tage) und durchschnittlicher individueller Spielerqualität mit der Gesamtsumme an Tagen des Bestands des Nachwuchsleistungszentrums relativiert. Im abschließenden Schritt werden alle Spieler eines Nachwuchsleistungszentrums aggregiert.

Um zu prüfen, ob die Bewertung des Erfolgs von Nachwuchsleistungszentren mittels des entwickelten Indikators *Hervorgebrachte Spielerqualität NLZ* eine valide Evaluationsmethode darstellt, wird eine Validierung mithilfe der *Double Pass Scores* vorgenommen.

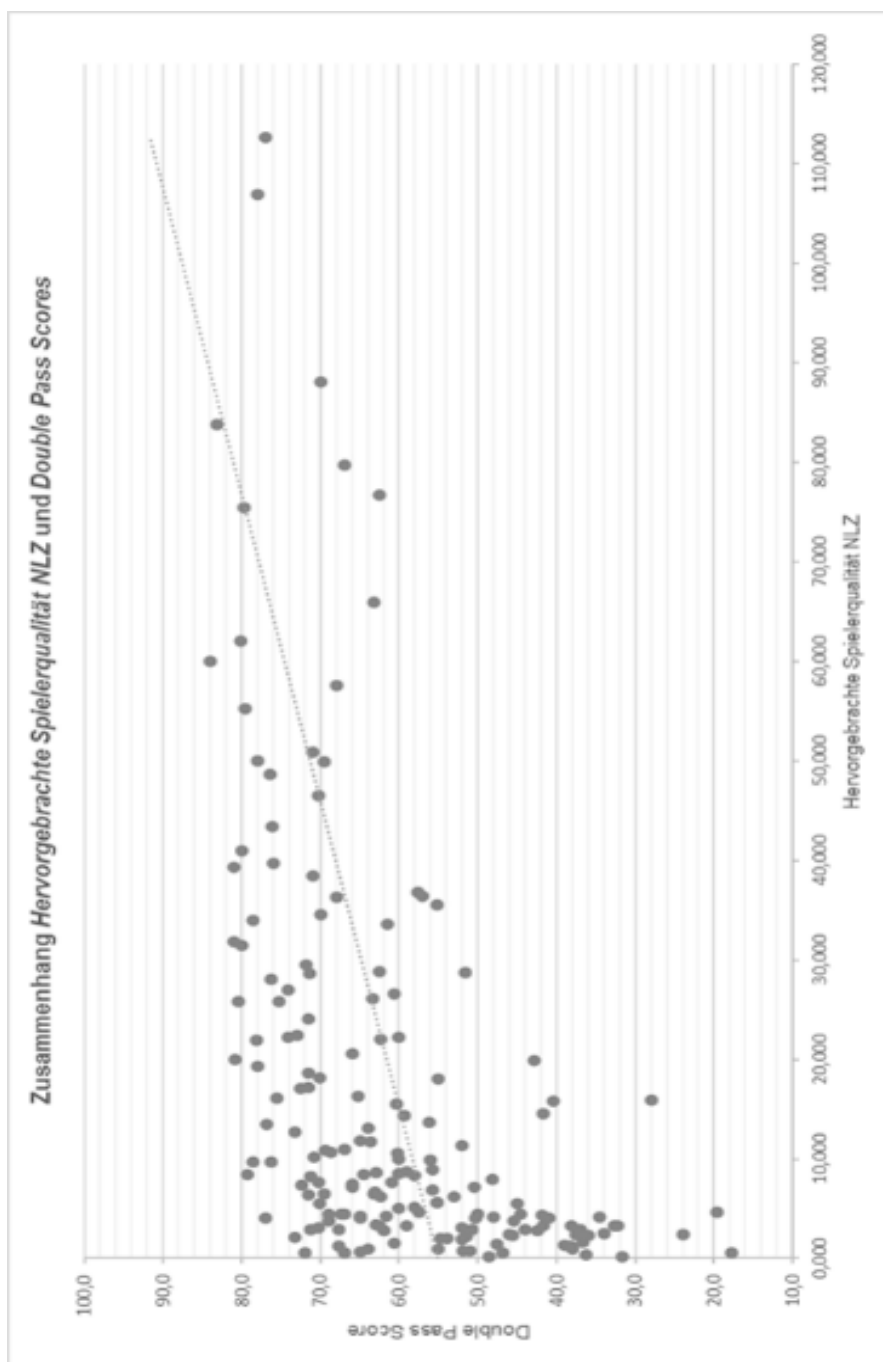
Formel 4: Berechnung Hervorgebrachte Spielerqualität NLZ. Eigene Weiterentwicklung nach Lanwehr et al. (2021).

$$SQ_{NLZ} = \sum_{\text{Alle Spieler}} \sum_{\text{Alle Saisons}} \frac{AU_{i,NLZ,j} \times \overline{SQ}_s \text{ bis Stichjahr in Tagen}}{(\text{Jahr}_{\text{Aktuell}} - \text{Jahr}_{\text{Beginn NLZ}}) \text{ in Tagen}}$$

SQ_{NLZ} = Spielerqualität NLZ; $AU_{i,NLZ,j} = 1$, wenn Spieler i im NLZ in Saison j ausgebildet wurde;
 \overline{SQ}_s = durchschnittliche Individuelle Spielerqualität

Die Validierung der *Hervorgebrachten Spielerqualität NLZ* mit *Double Pass Score NLZ* erfolgt mittels einer bivariaten (einfachen) Korrelationsanalyse (Helfrich, 2016). Für diese Analyse werden die zur Verfügung stehenden *PASS-Zyklen 1.0 bis 4.0* sowie die Mittelwerte der im jeweiligen Zeitraum der *PASS-Zyklen Hervorgebrachten Spielerqualität NLZ* in einen Zusammenhang gebracht. Abbildung 39 zeigt in der grafischen Darstellung des Zusammenhangs unabhängig vom *Pass-Zyklus* in einem Streudiagramm zunächst ein indifferentes Bild mit Ausreißern.

Abbildung 39: Zusammenhang Hervorgebrachte Spielerqualität NLZ und Double Pass Scores.



Es wurde daher eine bivariate Korrelationsanalyse durchgeführt. Sowohl die *Hervorgebrachte Spielerqualität NLZ* (SQ_NLZ) als auch die *Double Pass Scores* sind nicht normalverteilt und weisen Ausreißer auf, sodass als Korrelationskoeffizienten sowohl der Spearman-Rho als auch der Kendall-Tau-b betrachtet werden (Bühl, 2019). Mit einem Spearmans $\rho = 0,600$, $p < 0,001$, bei N=171 Wertepaaren wird lt. (Cohen, 1988) der Zusammenhang als stark interpretiert, während Kendall-Tau-b und -c Werte von $\rho = 0,424$, $p < 0,001$, bei N=171 (b- und c-Werte sind identisch) lediglich als mittlerer mit Tendenz zu einem starken Zusammenhang zu bewerten ist.

Betrachtet man die Spearman- und Kendall-Tau-Korrelationskoeffizienten pro PASS-Zyklus (siehe Tabelle 8), fällt auf, dass mit einer breiteren Datenbasis (Anzahl Werte) die Kendall-Tau-Korrelationskoeffizienten steigen, während der Spearman-Rho zunächst auch steigt, der Zusammenhang der Datenreihen für den PASS-4.0-Zyklus aber etwas sinkt. Die Korrelationen der PASS-Zyklen mit den Mittelwerten der Hervorgebrachten Spielerqualität in Analogie zu den Zeiträumen der PASS-Zyklen sind ab PASS 2.0 mit einem Spearmans $\rho > 0,50$ als stark zu interpretieren, während die Höhe der Kendall-Tau-b-Koeffizienten als mittlere Stärke zu bewerten ist.

Die Ermittlung des erarbeiteten Nutzens von Nachwuchsleistungszentren mittels des entwickelten Indikators *Hervorgebrachte Spielerqualität NLZ* wird aufgrund der Interpretation der Korrelationskoeffizienten von der Autorin als ein valides Messinstrument eingestuft und geht in die Kalkulation eines ROI für ein Nachwuchsleistungszentrum ein.

3.6.2.4 Monetäre Bewertung der Hervorgebrachten Spielerqualität

Die monetäre Bewertung der *Hervorgebrachten Spielerqualität* der Nachwuchsleistungszentren erfolgt mittels der Marktwerte der Spieler. Die Betrachtung der Lagemaße *Mittelwert* und *Median* der individuellen Spielerqualitäten und der Marktwerte (hier getrimmtes Mittel, $\alpha = 5\%$) zeigt, dass sich die Lagemaße der Individuellen Spielerqualität über den Betrachtungszeitraum stabil parallel entwickeln, die Lagemaße des Marktwertes aber eine große Diskrepanz aufweisen (siehe Abbildung 40). Die hohen Steigerungsraten des getrimmten Mittelwerts der Marktwerte über die Jahre spiegelt sowohl die wirtschaftliche Entwicklung des Transfermarktes der Spieler wider als auch die Erhöhung der Spannbreite der Spielermarktwerte. Das getrimmte arithmetische Mittel ist dem arithmetischen Mittel vorgezogen worden, um auf Ausreißer und Extremwerte der Höhe der Marktwerte zu reagieren, um ein resistenteres Lagemaß zur Bewertung der hervorgebrachten Spielerqualität einzusetzen. 5 % der Randdaten am Anfang und am Ende der Skala sind nicht in die Berechnung des Mittelwerts eingegangen (Fahrmeir et al., 2004).

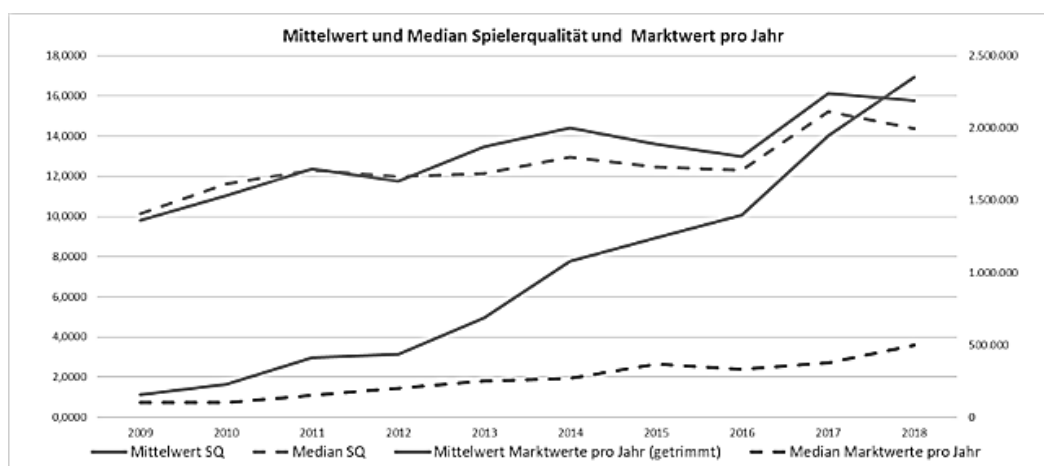
Tabelle 8: Deskriptive Statistik PASS-Zyklen und hervorgebrachte Spielerqualität (SQ_NLZ).

Zyklus / Zeitraum	MW	SD	n	Kolmogorov-Smirnov
PASS 1.0	55,8	14,530	39	$p = .200$
PASS 2.0	53,9	15,151	42	$p = .112$
PASS 3.0	63,8	13,019	49	$p = .200$
PASS 4.0	63,9	12,689	50	$p = .021$
SQ_NLZ 1.0	4,383	3,332	40	$p = .009$
SQ_NLZ 2.0	11,298	12,049	49	$p = .000$
SQ_NLZ 3.0	17,805	20,818	53	$p = .000$
SQ_NLZ 4.0	23,808	27,454	55	$p = .000$

Spearman-Rho	SQ_NLZ 1.0	SQ_NLZ 2.0	SQ_NLZ 3.0	SQ_NLZ 4.0
PASS 1.0	0,451			
PASS 2.0		0,571		
PASS 3.0			0,662	
PASS 4.0				0,638

Kendall-Tau-b	SQ_NLZ 1.0	SQ_NLZ 2.0	SQ_NLZ 3.0	SQ_NLZ 4.0
PASS 1.0	0,325			
PASS 2.0		0,391		
PASS 3.0			0,451	
PASS 4.0				0,465

Abbildung 40: Entwicklung Mittelwert und Median Spielerqualität und Marktwert pro Jahr



Die Relation der Mittelwerte der Marktwerte pro Jahr für die Spieler, die in einem der betrachteten Nachwuchsleistungszentren ausgebildet wurden, zu einem mittleren Spielerqualitätspunkt ergibt den Marktwertfaktor pro Jahr (siehe Tabelle 9).

Tabelle 9: Marktwertfaktor zur Bewertung der hervorgebrachten Spielerqualität.

Jahr	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Marktwertfaktor	15.792	20.613	33.236	37.094	50.869	74.749	91.173	107.806	120.690	148.950	136.514

3.6.3 Bewertung des Erfolgs von Nachwuchsleistungszentren mittels ROI

Die Bewertung des Erfolgs von Nachwuchsleistungszentrum mittels ROI basiert auf der Ermittlung der gemäß der Evaluation von Personalentwicklungsmaßnahmen überwiegend benutzten Formel (vgl. dazu Formel 1 und Kapitel 2.2.5.2 in dieser Arbeit). Der ROI wird wie in der Praxis üblich pro Jahr ermittelt (Phillips & Schirmer, 2008). Der Nettonutzen des Nachwuchsleistungszentrums ergibt sich durch die Subtraktion des Konstrukts *Investment* (vgl. Kapitel 3.6.1) vom Konstrukt *Nutzen* (siehe Kapitel 3.6.2). Dividiert durch die Kosten Nachwuchsleistungszentrum multipliziert mit 100 ergibt sich der Prozentsatz der Investition, den man nach der Maßnahme zurückerhalten wird.

Formel 5: ROI für Nachwuchsleistungszentren (NLZ)

$$ROI_{NLZ} (\%) = \frac{(SQ_{NLZ} \times \overline{MW}) - \text{Kosten}_{NLZ}}{\text{Kosten}_{NLZ}} \times 100$$

SQ_{NLZ} =Spielerqualität NLZ; MW =durchschnittlicher Marktwert Saison

Damit liegt ein konkreter ROI-Wert vor, den verantwortliche Manager*innen der Vereine interpretieren können, um zurückblickend die Effektivität der Ausbildung im Nachwuchsleistungszentrum zu beurteilen. Mit dem ROI_{NLZ} liegt eine Kennzahl vor, die den finanziellen Return jedes investierten Euros in ein Nachwuchsleistungszentrum repräsentiert.

3.6.4 Evaluation der Evaluation der Nachwuchsleistungszentren (Metaevaluation)

Ziel der vorliegenden Arbeit ist die Entwicklung eines Evaluationsmodells für die Messung des monetären Erfolgs von Nachwuchsleistungszentren, das outcome-orientiert ist, Resultate im Zeitverlauf über 11 Jahre zeigt, öffentlich zugängliche objektive Daten nutzt, Individual-Leistungsfaktoren einsetzt, mit geringem Aufwand anzuwenden ist und betriebswirtschaftliche Kennzahlen zur Bewertung des Erfolgs nutzt. Ziel dieser Evaluation ist es, nachvollziehbare und gültige Ergebnisse auszuweisen. Um bewerten zu können, ob und wie oben angeführte Ziele und die seit einigen Jahren existierenden Standards für Evaluationen (siehe dazu Deutsche Gesellschaft für Evaluation (DeGEval, 2017c)) erreicht werden, wird für die entwickelte Evaluationsmethode eine Metaevaluation bezüglich Einhaltung und Umsetzung der Evaluationsstandards gemäß DeGEval durchgeführt (Döring & Bortz, 2016). Definiert wird eine *Metaevaluation* im Allgemeinen als die Bewertung einer Evaluation. Stufflebeam (2011) definiert sie als ein Verfahren zur Beschreibung einer Evaluationstätigkeit und ihrer Beurteilung anhand einer Reihe von Kriterien, die eine gute Evaluation ausmachen. Metaevaluationen werden von der DeGEval als wichtig erachtet, da damit eine Erhöhung der Glaubwürdigkeit der Evaluation, aber auch der Aktivitäten der Evaluator*innen insgesamt einhergeht. Dar-

über hinaus soll über eine Aggregation und Systematisierung von Evaluationen die Erzeugung generalisierbares Wissen erzielt werden (DeGEval, 2017b). Die Metaevaluation der Evaluation der Nachwuchsleistungszentren wird im Rahmen einer Selbstevaluation durch die Evaluierende retrospektiv anhand der aktuellen Checkliste der DeGEval (DeGEval, 2017a) durchgeführt und in geeigneter Form dokumentiert, archiviert und über die Publikation dieser Arbeit zugänglich gemacht. Eine Analyse der in dieser Evaluationsmethode aufgefundenen Stärken und Schwächen (siehe Limitationen) führt dazu, dass eine Durchführung in der Zukunft – unabhängig von der Evaluatorin der vorliegenden Studie – abgesichert angeleitet werden kann (Thomas Widmer, 1996).

3.7 Zwischenfazit Untersuchungsrahmen

Um die hohe Qualität einer wissenschaftlichen Studie sicherzustellen, sollte mittels einer Selbstreflexion im methodischen Teil der Arbeit angegeben werden, wie genau vorgegangen wird sowie ob und wie die vier Kriterien wissenschaftlicher Qualität (Inhaltliche Relevanz, Methodische Strenge, Ethische Strenge und Präsentationsqualität) erfüllt werden (Döring & Bortz, 2016). Die vorliegende Studie zeichnet sich dadurch aus, dass sie mit wissenschaftlicher Methodik und hohen Ansprüchen an die Güte der Daten und der Konstrukte das relevante Forschungsproblem bearbeitet (Döring & Bortz, 2016). Die Ergebnisse der Studie sind eindeutig und unmittelbar für die weitere Forschung und Praxis anwendbar. Durch die Ergebnisse der vorliegenden Evaluation lassen sich konkrete Handlungsempfehlungen für die Lizenzierung und Zertifizierung von Nachwuchsleistungszentren ableiten. Der praktische Nutzen liegt darin, dass die Befunde helfen, die Praxisprobleme bei der Bewertung des Erfolgs von Nachwuchsleistungszentren zu lösen und die etablierte Evaluationsmethode abzulösen. Die Entwicklung einer validen Messmethode des monetären Erfolgs von Nachwuchsleistungszentren trifft den Informationsbedarf der beteiligten Akteure (Vereine mit NLZ) und der beteiligten Experten für Nachwuchsförderung im deutschen Profifußball bei DFB und DFL. Angewandte Methodologien und Methoden folgen stringent dem gewählten Forschungsdesign einer quantitativen, wissenschaftlich unabhängigen, empirischen, deskriptiven, nicht experimentellen Längsschnitt-Originalstudie mit Vollerhebung. Mit der angewandten Sekundäranalyse werden bereits erhobene objektive Daten verwendet. Diese Daten sind unabhängig von subjektiven Selbsteinschätzungen, die bei der bisher etablierten Evaluationsmethode seitens der Verantwortlichen im Nachwuchsleistungszentrum und den eingesetzten Auditor*innen die Basis bilden. Da die Primärdaten ebenfalls auf der Basis objektiver öffentlich zugänglicher Daten generiert wurden, wurde auf eine Aufklärung der Nachwuchsleistungszentren über die Untersuchung verzichtet.

Die Gütekriterien werden für alle Phasen des Forschungsprojektes – „also die Operationalisierung der begrifflichen Konstrukte, die Hypothesenbildung, die Wahl der Forschungsstrategie, die Datenerhebung und -aufbereitung, die Datenauswertung und die Ergebnisinterpretation“ (Helfrich, 2016, S. 95) – berücksichtigt. Die Absicherung der verwendeten Daten bezüglich der drei Gütekriterien – Objektivität (Sachlichkeit), Reliabilität (Zuverlässigkeit), Validität (Gültigkeit) – wird in den Kapiteln Datenerhebung und Datenauswertung erläutert (Helfrich, 2016). In der vorliegenden Arbeit werden das Forschungsprojekt vollständig und strukturiert dargestellt sowie die Ergebnisse lesbar und anschaulich (auch mittels) Grafiken dokumentiert (Döring & Bortz, 2016).

Die Entwicklung des Evaluationsmodells selbst wird mittels einer retrospektiven Metaevaluation nach den Standards der DeGEval bezüglich ihrer Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit untersucht und damit dokumentiert. Die Analyse der aufgefundenen Stärken und Schwächen führt dazu, dass Interessenten in Zukunft die Evaluation unabhängig von der Evaluatorin der vorliegenden Studie nach Anleitung mit geringem Aufwand eigenständig durchführen können. (Thomas Widmer, 1996).

Im Anschluss an die Beschreibung des Ablaufs dieser empirischen Untersuchung (Ort und Zeit der Durchführung, Art und Umfang der Stichprobe, verwendete Daten und deren Aufbereitung und Güte) und des Zwischenfazit werden die Ergebnisse der statistischen Analysen mit Angabe ihrer relevanten Kenngrößen sowie deren Bewertung in Bezug auf die Forschungsfrage aufgeführt.

4 Ergebnisse und Diskussion der Evaluation

Nachdem die Erhebung und Aufbereitung der in dieser Studie verwendeten Daten in Kapitel 3 beschrieben worden sind, werden die gewonnenen Daten und Kennzahlen im Folgenden zusammengefasst und in Hinblick auf ihren Geltungsbereich beurteilt. Dies geschieht mit Hilfe statistischer Auswertungsverfahren. Da es sich bei dieser Studie um eine Vollerhebung handelt, werden die Ergebnisse mit Hilfe der Verfahren der deskriptiven Statistik beschrieben. Inferenzstatistische Methoden kommen nicht zum Einsatz (Helfrich, 2016).

4.1 Ergebnisse – Deskriptive Aussagen

In die Untersuchung eingegangen sind 108.319 Spieler, für die öffentlich verfügbare Informationen aus dem Profifußball vorlagen. Für diese Spieler sind Leistungsdaten aus insgesamt über 1 Million Einsätzen in 329 Ligen und 4949 Vereinen analysiert worden. 630 Profifußballspieler, die in deutschen Nachwuchsleistungszentren (N=54) ausgebildet wurden, sind in die Untersuchung eingegangen. Der Return on Investment konnte für 39 Nachwuchsleistungszentren berechnet werden. Um einen Überblick über die vorliegenden Daten zu erhalten und um diese zu ordnen und zusammenzufassen, werden die Daten und Ergebnisse mittels der üblichen Maße der deskriptiven Statistik im Folgenden beschrieben.

4.1.1 Kosten

Die Kosten sind mit dem ligaspezifischen prozentualen Anteil am Rohergebnis des Vereins pro Jahr berechnet worden (siehe dazu Kapitel 3.6.1.1). Kosten im Beobachtungszeitraum können für 39 Nachwuchsleistungszentren (n=39) aus der vorliegenden Grundgesamtheit (N=54) berechnet werden. Tabelle 10 zeigt die ermittelten Kosten der Nachwuchsleistungszentren pro Jahr des Betrachtungszeitraums in Millionen Euro nach Ligazugehörigkeit der Vereine in der Saison 2018/19.

Insgesamt sind über den Betrachtungszeitraum (2009-2019) 1,1 Mrd. Euro in die Nachwuchsleistungszentren investiert worden. Betrachtet man die 36 Vereine der 1. und 2. Bundesliga, werden durchschnittlich 74 % der Investitionssumme durch Vereine der 1. Bundesliga gestellt, während 26 % in der 2. Bundesliga, trotz des höheren Prozentsatzes als Anteil der Investitionen am Rohergebnis, investiert wird.

Tabelle 10: Kosten Nachwuchsleistungszentren in Mio. EURO pro Liga in 2018/19.

Kosten Nachwuchsleistungszentren in Mio Euro*												
Liga in 2018/2019	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Summe
1. Liga	60,3	63,2	69,0	76,8	87,2	97,2	110,4	114,7	132,8	138,7	126,4	1.076,7
2. Liga	28,1	28,9	28,9	27,8	25,9	30,6	32,8	36,8	39,9	39,7	35,7	355,0
3. Liga	7,9	10,8	11,6	10,5	9,4	7,8	8,7	8,8	7,9	5,9	7,1	96,4
Summe	96,3	102,9	109,5	115,1	122,5	135,6	151,8	160,3	180,6	184,4	169,2	1.528,1

*Stichprobe 39 Vereine

Die TOP-6-Vereine investierten über die betrachteten 11 Jahre insgesamt eine Summe von 667 Mio. Euro, wobei hier zwei Vereine auch aus den TOP 6 herausstechen, Bayern München mit 214 Mio. Euro und BVB Dortmund mit 129 Mio. Euro. Unter diesen 5 Vereinen ist ein Verein, der in der Saison 2018/19 in der 2. Bundesliga spielt, aber bis dahin lange Jahre in der 1. Bundesliga verbracht hat (Hamburger SV). Die anderen drei Vereine der TOP 6 – FC Schalke 04, Bayer Leverkusen und der VfL Wolfsburg – haben mit 95, 77 und 73 Mio. Euro deutlich weniger investiert. In Summe tragen diese 6 Vereine 44 % der gesamten Investitionen in die Nachwuchsleistungszentren. Von den folgenden 18 Vereinen investieren 7 im Bereich von 40 bis 61 Mio. Euro, 5 im Bereich von 30 bis 35 Mio. Euro und 6 im Bereich von 21 bis 26 Mio. Euro. Aufsummiert zu 661 Mio. Euro repräsentieren diese 18 Vereine einen ähnlich hohen Anteil von 43 % der gesamten Investitionen in den Nachwuchsbereich. Darauf folgen 11 Vereine mit Investitionssummen von 10-19 Mio. Euro (11 %). Sie stellen eine Gesamtinvestitionssumme von 180 Mio. Euro dar. Die Low 6, Vereine aus der 2. Bundesliga und Liga 3, vereinigen 51 Mio. Euro (2 %) der Kosten (siehe dazu Abbildung 41).

Abbildung 41: Cluster-Kosten NLZ in Mio. Euro.

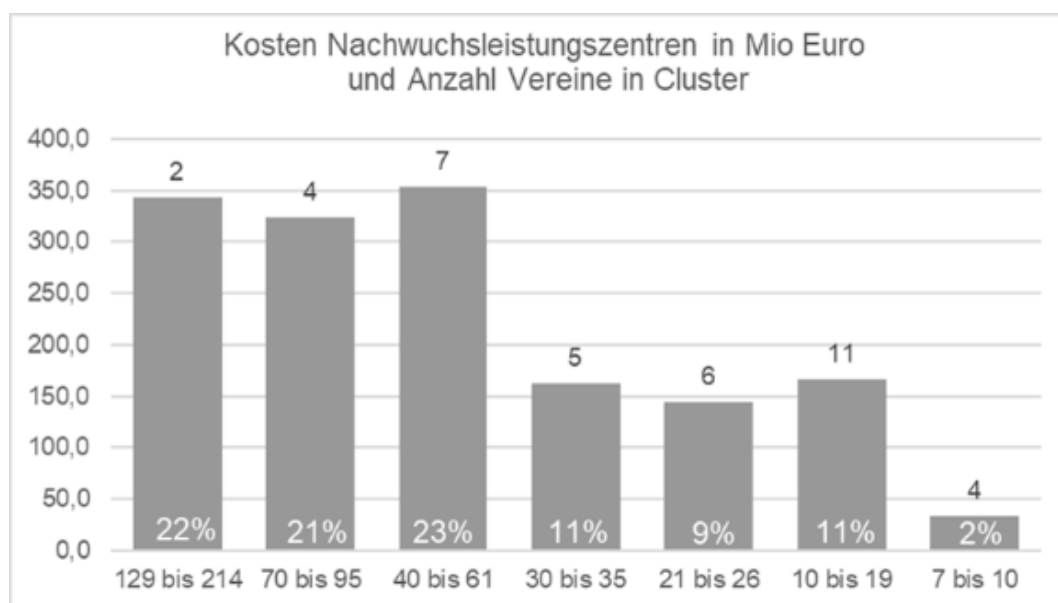


Tabelle 11 zeigt die Entwicklung der Kosten der beobachteten 42 Nachwuchsleistungszentren für die Jahre 2009-2019.

Tabelle 11: Kosten Nachwuchsleistungszentren sortiert nach Ligazugehörigkeit in 2018/2019.

Kosten Nachwuchsleistungszentren in Euro (sortiert nach Gesamtsumme)															
Liga in 2018/2019	Verein	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Summe		
1	FC Bayern München	12.922.380	12.117.960	13.778.460	15.970.320	19.509.090	19.324.530	23.128.920	23.634.450	24.258.060	26.339.626	23.423.161	214.406.897		
1	BVB Dortmund	3.722.804	5.033.234	7.055.760	10.052.888	9.990.122	10.219.381	13.885.470	14.981.400	19.182.133	17.564.289	17.479.050	129.166.531		
1	Schalke 04	6.931.665	8.271.135	7.039.413	7.632.027	7.944.570	9.767.430	9.782.190	8.858.583	12.486.443	9.710.862	6.254.661	94.678.979		
2	Hamburger SV	8.538.175	8.303.000	7.674.525	7.590.000	6.290.000	7.365.570	7.072.500	7.020.750	7.682.000	7.256.788	5.457.958	79.251.445		
1	Bayer Leverkusen	3.515.213	3.690.000	5.090.900	5.586.300	6.438.867	7.250.787	8.712.090	7.520.220	9.136.661	9.564.000	10.011.745	76.516.784		
1	VfL Wolfsburg	5.512.860	5.391.090	3.690.000	5.678.577	6.545.227	7.370.559	8.856.000	7.011.000	8.074.015	7.144.541	7.782.210	73.056.079		
2	FC Köln	3.754.750	3.990.500	4.835.750	3.484.500	4.163.000	5.117.500	6.152.500	7.417.500	9.366.175	6.131.570	6.598.010	61.011.755		
1	Bor Mönchengladbach	2.424.330	2.619.900	4.512.870	3.372.660	4.789.620	5.926.140	7.265.610	6.616.170	6.035.622	7.396.568	5.710.201	56.669.692		
1	RB Leipzig	831.235	872.566	905.301	993.397	1.145.007	2.932.443	4.376.709	8.017.632	9.049.688	10.133.183	11.994.603	51.251.765		
1	VfB Stuttgart	4.236.120	4.317.300	4.029.480	4.236.120	3.974.130	4.630.950	4.889.250	4.151.250	5.425.555	6.124.330	3.405.390	49.419.875		
1	Eintracht Frankfurt	2.468.610	2.642.040	2.247.210	3.191.850	3.653.100	4.033.170	4.084.830	5.195.520	6.261.008	6.715.800	5.387.621	45.880.759		
1	Werder Bremen	4.495.158	4.204.386	3.274.875	3.127.644	3.368.970	3.800.700	3.985.200	4.557.150	4.205.825	5.982.084	4.373.277	44.985.270		
1	Hoffenheim	2.003.670	2.704.770	3.136.500	2.619.900	2.435.400	2.638.350	4.723.200	4.095.900	5.973.557	5.968.944	8.410.100	44.710.291		
1	Hertha BSC Berlin	2.911.410	1.682.640	2.424.330	1.413.270	2.619.900	2.808.090	3.018.420	4.143.870	4.948.696	5.194.856	4.130.402	35.295.883		
1	Mainz 05	1.188.180	1.479.690	2.520.270	2.523.960	2.830.230	2.778.570	3.867.120	4.059.000	4.185.456	5.331.091	4.248.186	35.011.753		
3	FC Kaiserslautern	2.707.541	3.754.080	4.034.448	3.618.501	3.202.553	2.679.059	2.976.732	3.038.904	2.851.200	1.322.640	1.190.801	31.376.458		
1	Hannover 96	2.214.000	2.952.000	2.874.510	2.771.190	2.749.050	3.154.950	1.667.880	3.188.160	3.477.677	3.801.401	2.038.910	30.889.728		
2	FC Ingolstadt	2.593.843	2.674.278	2.677.685	2.572.973	2.407.525	2.908.930	3.122.250	3.961.750	3.106.453	3.182.683	1.079.275	30.287.645		
1	SC Freiburg	1.143.900	1.354.230	1.808.100	1.845.000	2.583.000	2.904.030	1.815.480	2.339.460	3.555.869	3.368.158	3.143.880	25.881.107		
1	Augsburg	745.380	656.820	1.287.810	1.468.620	1.752.750	2.191.159	3.553.470	3.498.120	3.392.512	3.524.762	3.673.026	25.744.429		
2	St. Pauli	1.873.891	1.932.000	1.742.250	1.730.750	1.759.500	1.904.975	2.340.250	2.725.500	2.549.205	2.856.830	2.728.548	24.143.698		
1	FC Nürnberg	2.066.945	2.169.720	2.248.176	2.447.547	2.771.625	3.081.983	1.704.780	1.354.230	1.685.629	2.621.155	1.985.036	24.136.825		
2	Darmstadt	1.982.403	2.043.877	2.046.481	1.966.453	1.840.005	2.223.215	2.386.250	2.748.500	1.975.873	1.759.385	1.785.203	22.757.645		
2	Union Berlin	952.463	981.999	983.250	1.173.000	1.568.025	1.512.250	1.795.725	2.213.750	2.670.473	3.453.278	4.118.840	21.423.052		
1	Fortuna Düsseldorf	975.774	1.024.293	1.062.720	1.886.328	2.136.096	2.375.289	1.070.100	1.466.775	1.427.329	2.588.609	2.965.838	18.979.150		
2	VfL Bochum	1.526.169	1.573.495	1.575.500	1.581.250	1.564.000	1.587.000	1.851.500	1.903.250	1.888.760	1.942.063	1.888.128	18.881.114		
3	Karlsruher SC	1.702.800	2.177.496	2.353.245	2.133.303	1.913.362	1.600.601	1.778.445	1.805.760	1.242.569	399.960	1.736.777	18.844.317		
2	Dynamo Dresden	1.113.992	1.148.537	1.150.000	1.265.000	1.236.250	1.092.500	1.178.750	1.725.000	1.773.415	2.011.810	1.858.745	15.553.999		
2	Greuther Fürth	885.500	885.500	1.541.000	1.782.500	1.506.500	1.518.000	1.414.500	1.466.250	1.663.763	1.530.985	1.286.160	15.480.668		
2	SC Paderborn	1.527.899	1.575.279	1.577.286	1.515.605	1.418.148	1.713.500	1.380.000	402.500	715.185	1.534.388	1.853.858	15.213.646		
3	Hansa Rostock	1.104.026	1.530.763	1.654.313	1.499.696	1.345.079	1.125.210	1.250.233	1.269.435	1.148.400	1.386.000	1.377.367	14.690.523		
3	1860 München	1.006.819	1.395.983	1.508.654	1.367.651	1.226.648	1.026.138	1.140.153	1.157.664	1.047.286	831.600	1.061.280	12.769.876		
2	MSV Duisburg	1.178.750	1.610.000	948.750	883.778	826.948	999.174	1.072.446	1.195.492	1.362.750	1.362.750	1.218.971	12.659.809		
2	Holstein Kiel	913.502	941.829	943.029	906.152	847.884	1.024.469	1.089.596	1.225.758	1.397.250	1.771.748	1.561.183	12.632.400		
3	Wehen Wiesbaden	811.041	1.124.530	1.215.292	1.101.708	988.123	826.603	918.447	932.553	843.638	1.076.407	969.112	10.807.455		
2	Arminia Bielefeld	228.368	235.450	235.750	327.750	546.250	575.000	730.250	868.250	1.584.413	2.594.170	1.751.910	9.677.561		
2	Jahn Regensburg	622.534	641.839	642.657	617.526	577.817	698.157	749.355	835.331	952.200	1.127.748	1.396.790	8.861.954		
3	VfL Osnabrück	565.557	784.160	847.451	768.245	689.040	554.400	609.840	601.920	776.774	899.158	809.531	7.908.075		
2	FC Magdeburg	358.266	369.375	369.846	355.383	332.531	401.786	431.250	1.053.619	1.201.028	1.230.500	1.100.674	7.204.258		
Gesamtsumme		96.257.922	102.857.745	109.543.847	115.059.321	122.485.881	135.642.727	151.837.692	160.258.328	180.562.542	184.366.726	169.246.415	1.528.119.147		

4.1.2 Nutzen

Im Folgenden werden die Ergebnisse der Datenkonstrukte, die zur Berechnung des Nutzens entwickelt wurden, beschrieben und interpretiert. Diese sind Vereinsqualität, Individuelle Spielerqualität, Hervorgebrachte Spielerqualität des Nachwuchsleistungszentrums sowie die monetäre Bewertung der Hervorgebrachten Spielerqualität.

4.1.2.1 Vereinsqualität

In Abhängigkeit von der Liga-Höhe, der Punktzahl des Vereins und der durchschnittlichen Punktzahl des Deutschen Meisters der letzten 5 Jahre (ausgehend vom Betrachtungsjahr) wird die Vereinsqualität berechnet. Das Konstrukt *Vereinsqualität* ist wesentlicher Bestandteil der Berechnung des Konstrukts *Individuelle Spielerqualität*. Die Bildung des Konstrukts *Vereinsqualität* und die Validierung wurden in Kapitel 3.6.2 dargestellt.

Bewertet wurden im Betrachtungszeitraum (Jahre 1999-2020) insgesamt 1.240 von 6.459 Vereinen über die relevanten Ligen in Deutschland. Die Mittelwerte der jeweiligen Ligen zeigen deutliche Abstände zwischen den Ligen, sie repräsentieren die Höhe der Qualität der Ligen und damit der Vereine, die in diesen Ligen spielen (siehe dazu Tabelle 12).

Tabelle 12: Deskriptive Statistik Vereinsqualität per Liga in Deutschland.

Liga	Anzahl Vereine	Datenpunkte*	Min	Max	Mittelwert	STD
1. Bundesliga	37	394	4,545	4,694	4,546	0,044
2. Bundesliga	62	380	3,536	3,687	3,589	0,049
Liga 3	104	565	2,587	2,731	2,608	0,038
Liga 4	397	2.248	1,372	1,724	1,550	0,073
Liga 5	640	2.872	0,000	0,686	0,504	0,161
Summe / MW	1.240	6.459	2,408	2,704	2,559	0,073

Die Distanzen zwischen den Ligen bleiben über den gesamten Betrachtungsraum relativ konstant (siehe dazu Abbildung 42). Die Vereine, die in der jeweiligen Liga spielen, haben eine entsprechend hohe Qualität, sodass davon ausgegangen werden kann, dass die Vereinsqualität als ein stabiles Maß in das Konstrukt *Individuelle Spielerqualität* eingehen kann und ein bestimmtes Leistungsniveau der Spieler repräsentiert.

Abbildung 43 zeigt die Mittelwerte der Vereinsqualität der Vereine der 1. Bundesliga und ihre Distanz zum Mittelwert der 1. Bundesliga über die Jahre 1999-2020. In der 1. Bundesliga – der höchsten Spielklasse in Deutschland – spielten in dieser Zeit (1999-2020) 37 Vereine bei einer Ligastärke von 18 teilnehmenden Vereinen pro Jahr. Davon gehörten 6 Vereine jedes Jahr der 1. Liga an (22 Jahre), 3 Vereine verbrachten 19 Jahre in der 1. Liga und 10 Vereine waren nur 1 oder 2 Jahre beteiligt.

Die Anzahl der Jahre der Zugehörigkeit zur 1. Bundesliga korreliert stark mit der Höhe des Mittelwerts des Vereins (Spearman's $\rho = 0,749$, $p < 0,001$; siehe Abbildung 44). 16 Vereine (43 % von insgesamt 37 Vereinen) liegen mit ihrem vereinseigenen Mittelwert über dem Mittelwert der gesamten 1. Bundesliga. 11 Vereine davon sind zwischen 17 und 21 Jahren Teil der 1. Bundesliga gewesen.

Abbildung 42: Mittelwerte Vereinsqualität per Liga in Deutschland im Zeitverlauf

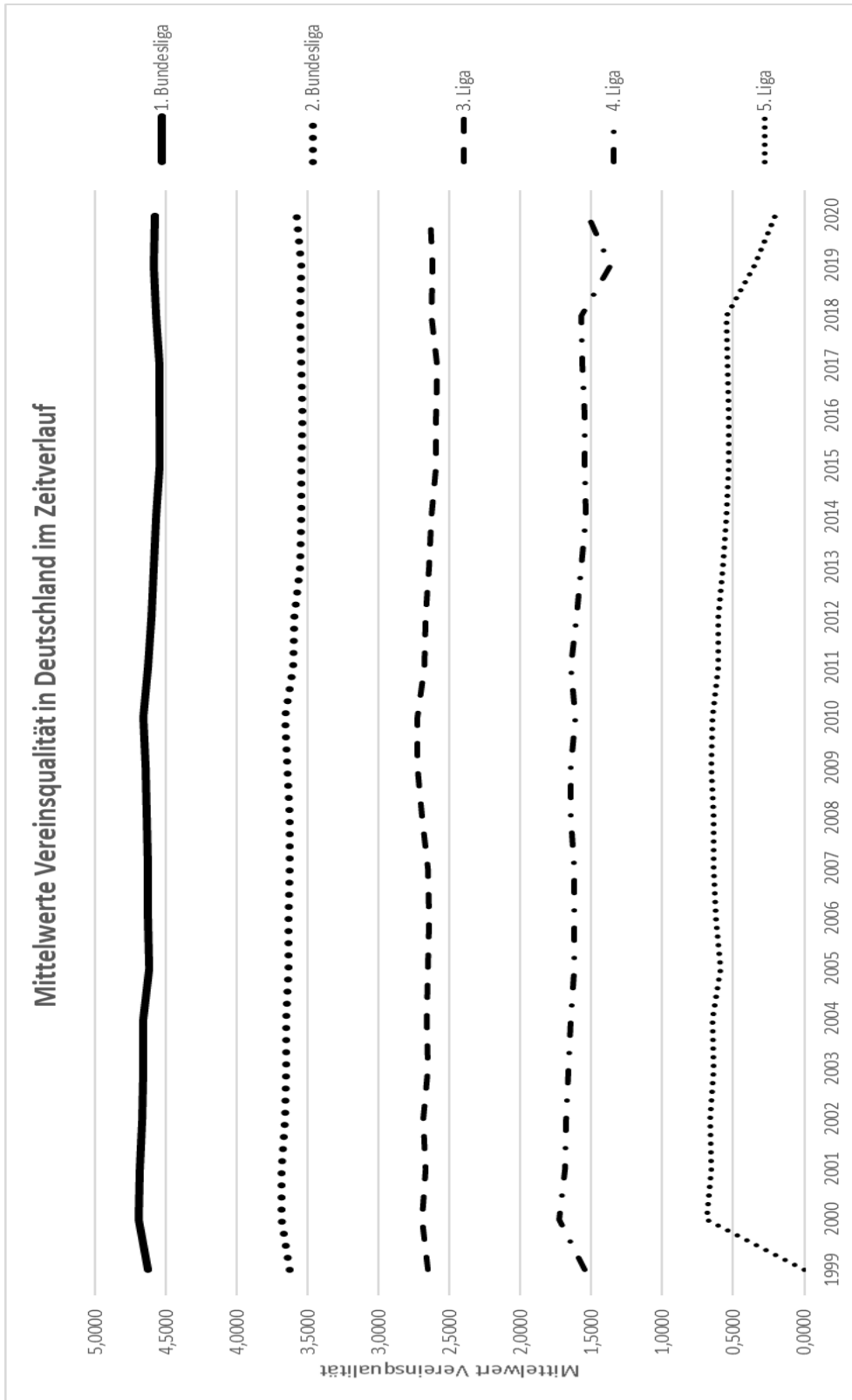


Abbildung 43: Mittelwert Vereinsqualität pro Verein in der 1. Bundesliga im Betrachtungszeitraum (1999-2020)

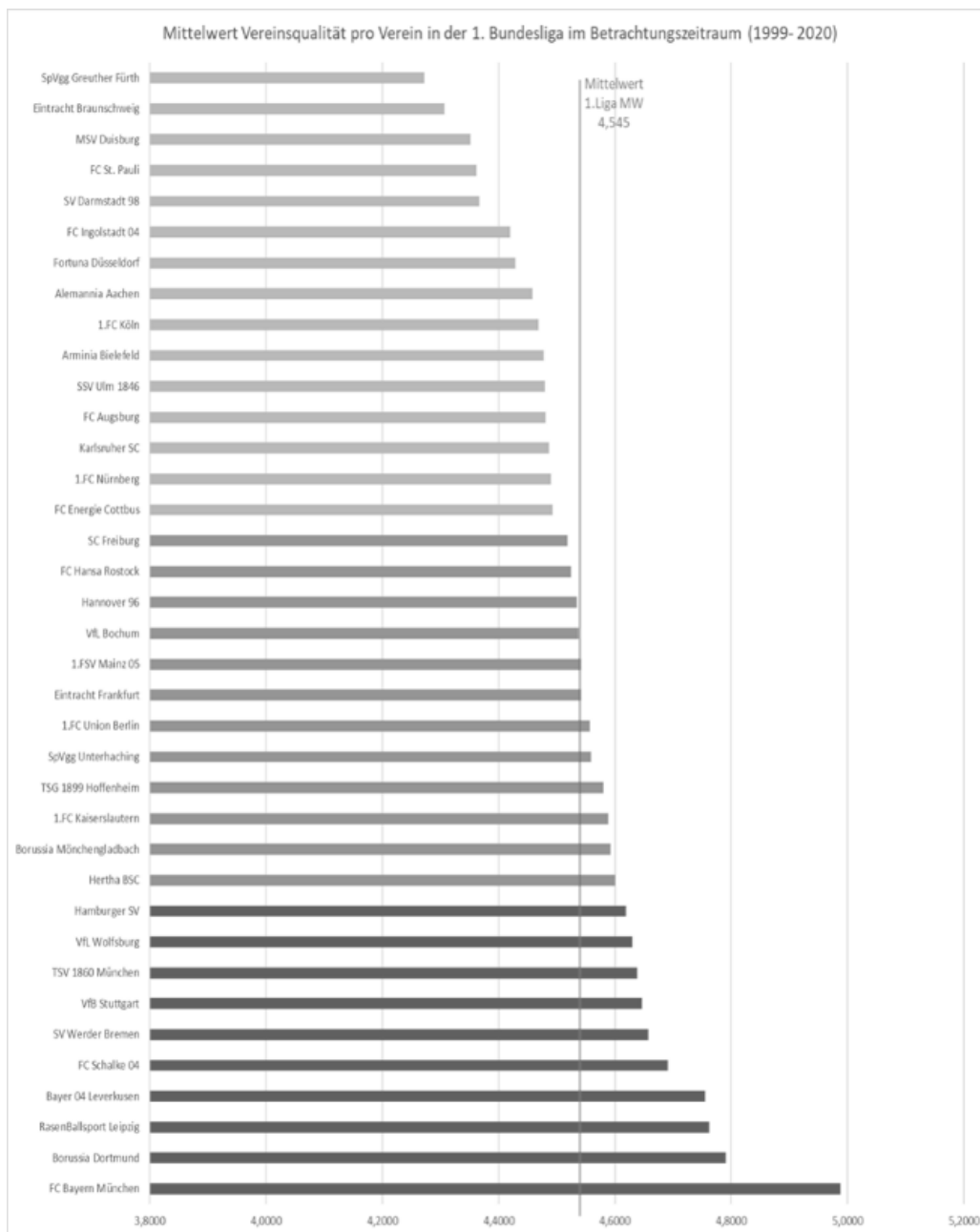
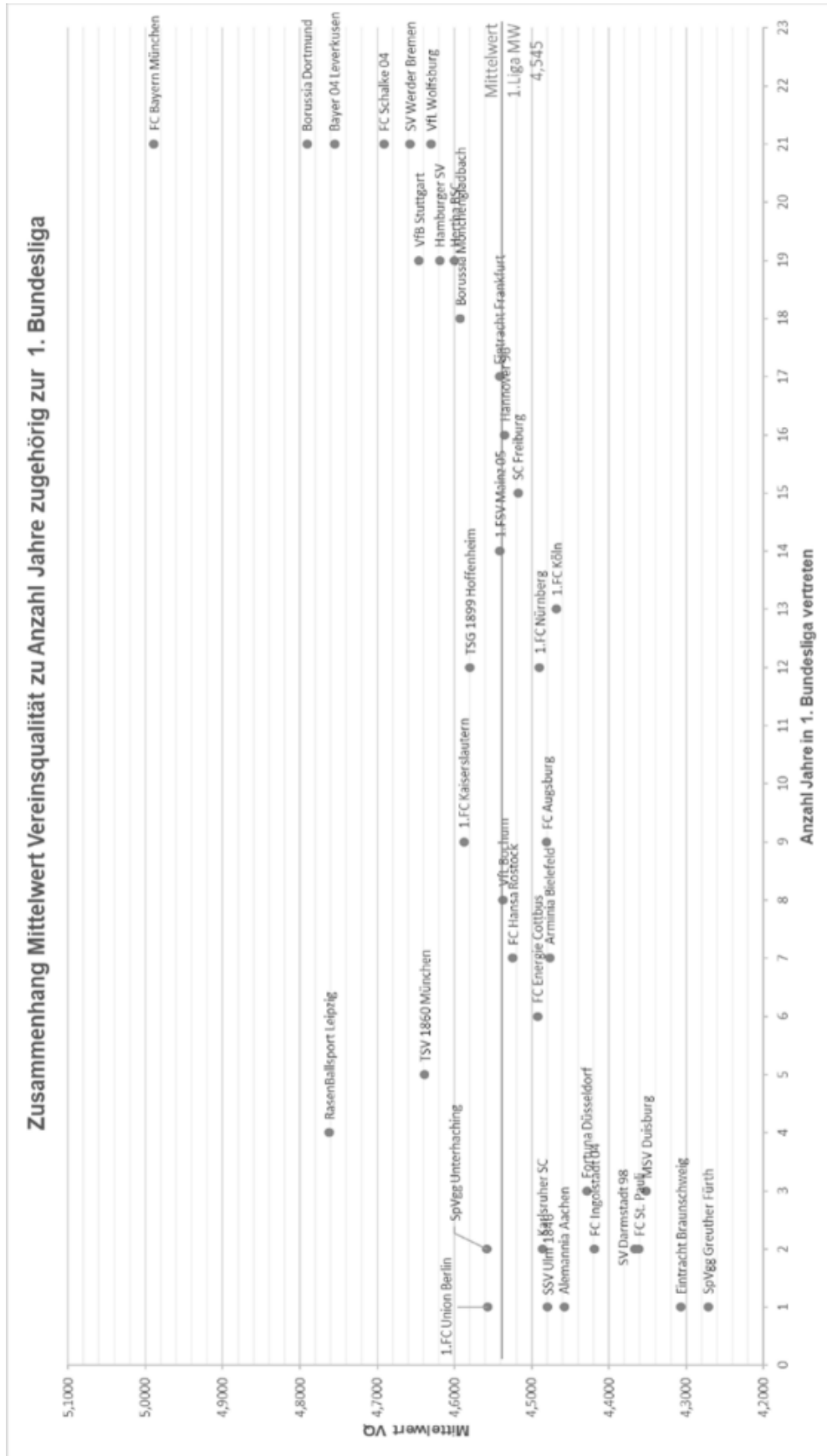


Abbildung 44: Zusammenhang Mittelwert Vereinsqualität zu Anzahl Jahre zugehörig zur 1. Bundesliga



In der 2. Bundesliga – der zweithöchsten Spielklasse in Deutschland – spielten im Zeitraum 1999-2020 62 Vereine, bei ebenfalls 18 Vereinen in dieser Profiligen pro Jahr. Lediglich ein Verein gehörte im Betrachtungszeitraum 21 Jahre der 2. Bundesliga an und nur 1 Jahr der 1. Liga. Im Vergleich zur Dauerhaftigkeit der Zugehörigkeit von Vereinen zur 1. Bundesliga fällt auf, dass viel mehr Vereine nur wenige Jahre – 12 Vereine für 3-4 Jahre und 18 Vereine lediglich 1-2 Jahre – an den Wettbewerben der 2. Bundesliga teilnahmen. Es besteht nur eine mittlere Korrelation zwischen den Jahren der Zugehörigkeit zur 2. Liga und den Mittelwerten (Spearman's $\rho = 0,341$, $p < 0,007$; vgl. dazu Abbildung 45). 29 Vereine liegen mit ihrem vereinseigenen Mittelwert über dem Mittelwert der gesamten 2. Bundesliga (47 % von insgesamt 62 Vereinen, vgl. dazu Abbildung 46).

4.1.2.2 Individuelle Spielerqualität

Der Indikator *Individuelle Spielerqualität* misst die Qualität eines Spielers anhand der Qualität des Vereins, in dem der Spieler in der jeweiligen Saison gespielt hat und anhand der Anzahl der gespielten Spiele in Relation zur maximalen Anzahl gespielter Spiele eines Spielers in der Saison. Innerhalb des Betrachtungszeitraums sind pro Saison mehr und mehr Spieler in die Analyse eingeflossen. Die Datenbasis wurde breiter, da die Dokumentation und Verfolgung der Spielerleistungsdaten über Transfermarkt.de im Zuge der Digitalisierung und der Entwicklung der Fußballbranche zunahm. Seit 2016 ist die Anzahl der Spieler, die in die Analyse eingehen, weitgehend stabil (siehe Abbildung 47). Auch die Mittelwerte der Individuellen Spielerqualität entwickelten sich mit zunehmender Professionalisierung des Fußballsports und des Ligenbetriebs in den Big-5-Ländern. Die Entwicklung des Mittelwerts – positive und negative Steigungen – wurde insbesondere beeinflusst durch die Qualität der Vereine, in denen die Spieler in den Jahren ihre Leistungen abgeliefert haben.

Abbildung 45: Mittelwert Vereinsqualität pro Verein in der 2. Bundesliga im Betrachtungszeitraum (1999-2020)

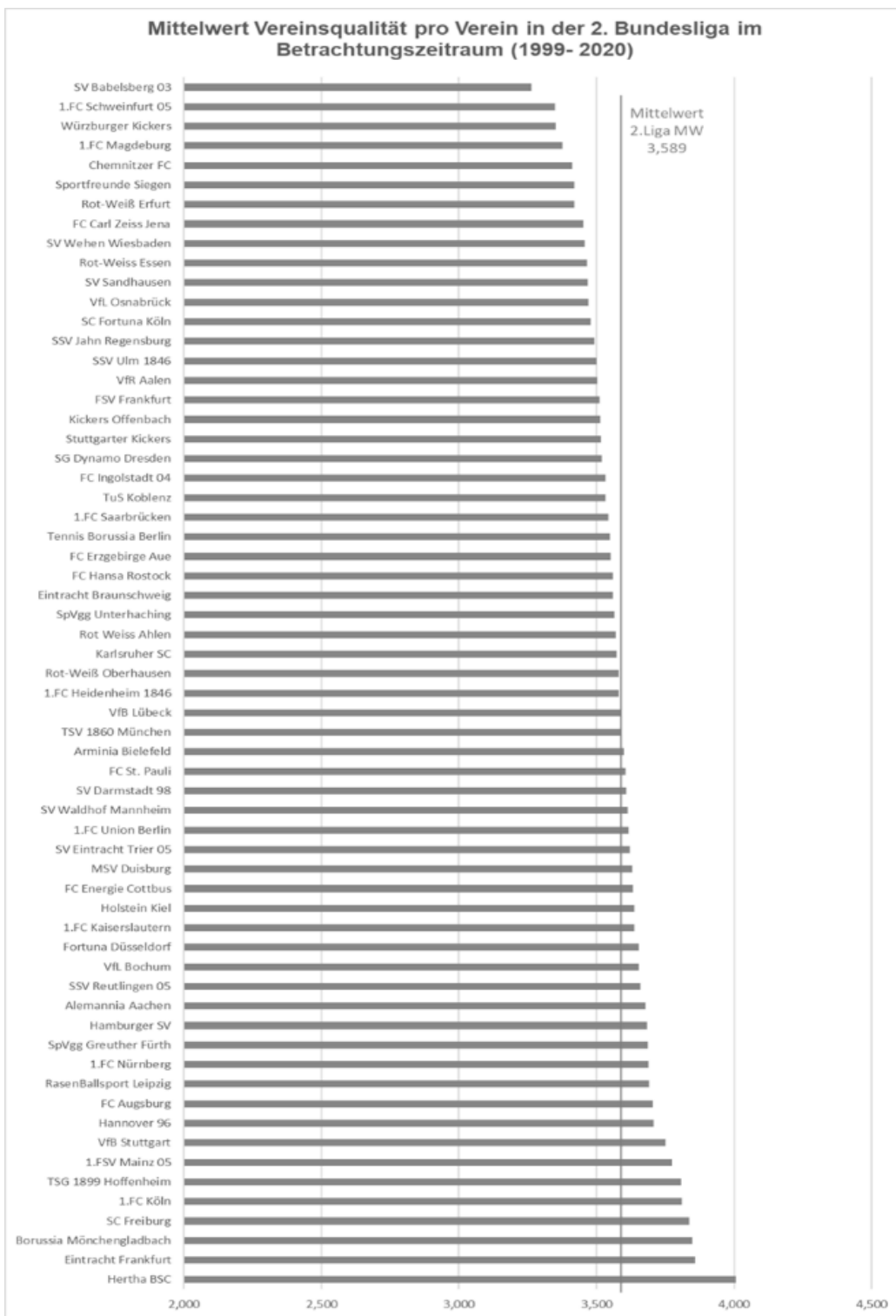


Abbildung 46: Zusammenhang Mittelwert Vereinsqualität zu Anzahl Jahre zugehörig zur 2. Bundesliga

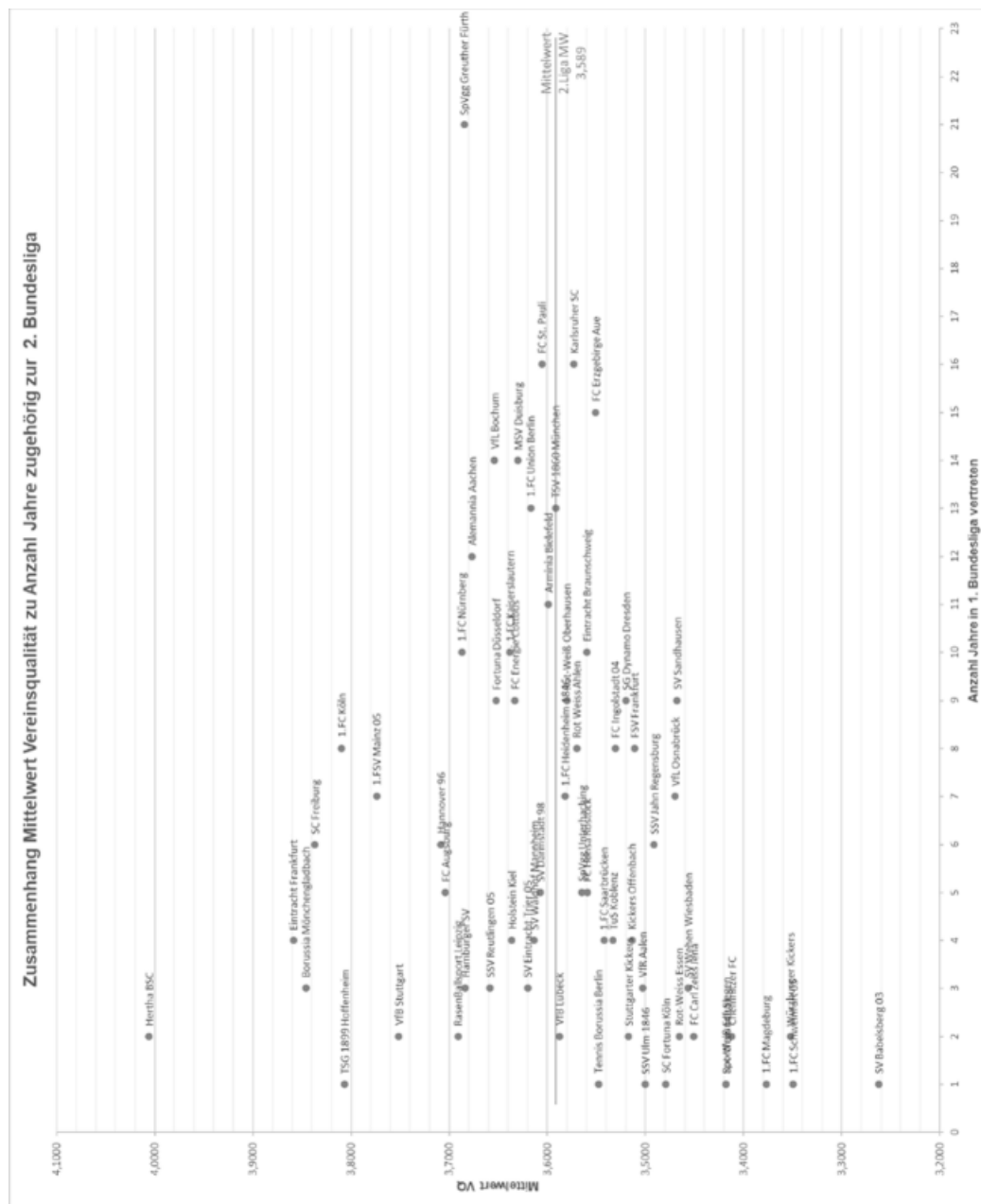
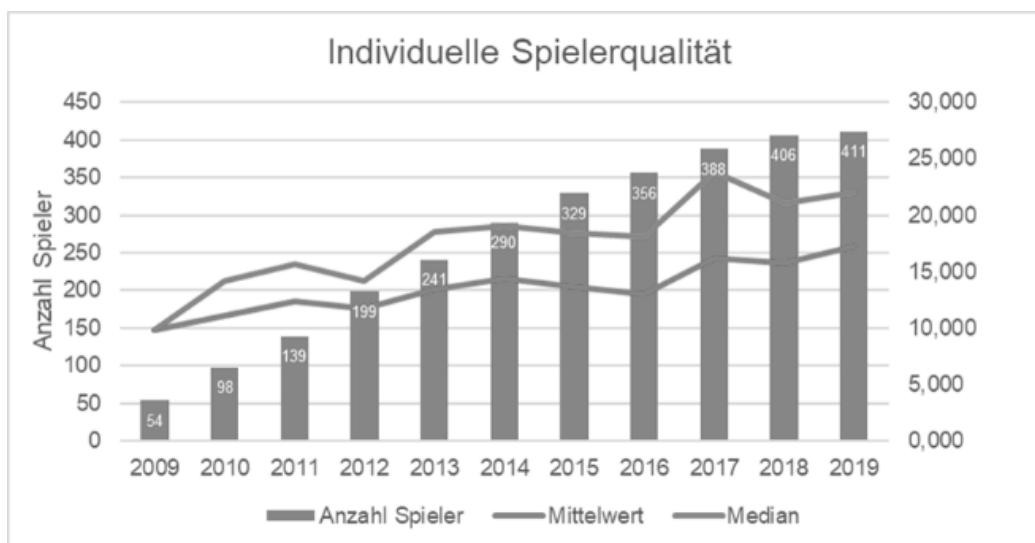


Abbildung 47: Anzahl Spieler, Median und Mittelwert Individueller Spielerqualitäten im Betrachtungszeitraum



Die Standardabweichung der Individuellen Spielerqualität liegt im Mittel über die Jahre 2009-2019 bei 8,680 (siehe Tabelle 13).

Tabelle 13: Deskriptive Statistik Individuelle Spielerqualität

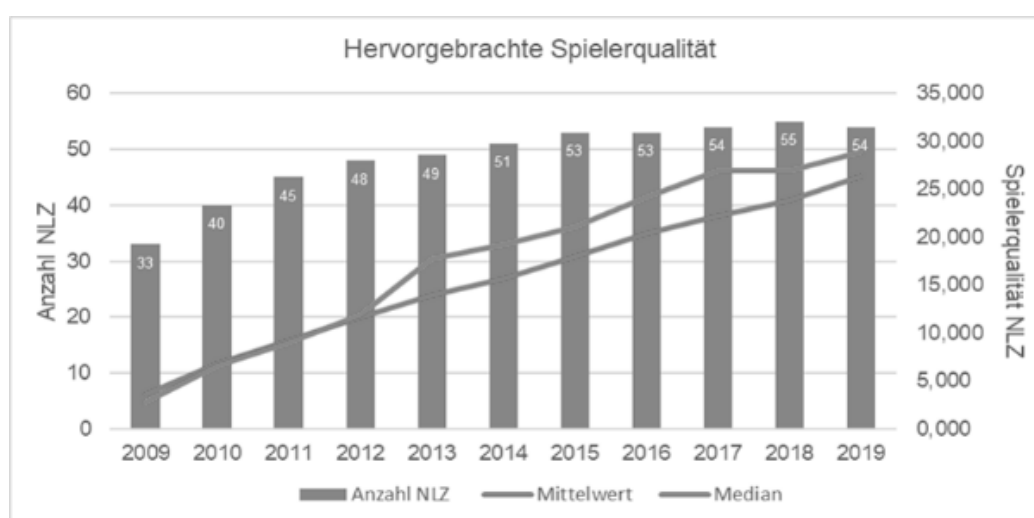
Jahr	Anzahl Spieler	Min	Max	Mittelwert	STD	Median
2009	54	1,364	25,168	9,813	6,526	9,833
2010	98	0,444	44,718	11,055	7,701	14,164
2011	139	0,146	43,786	12,372	8,319	15,710
2012	199	0,276	40,298	11,764	8,055	14,134
2013	241	0,262	46,293	13,473	9,182	18,474
2014	290	0,216	45,496	14,413	9,676	19,084
2015	329	0,189	36,149	13,600	8,800	18,411
2016	356	0,293	34,992	12,976	8,258	18,113
2017	388	0,143	47,413	16,143	9,895	23,819
2018	406	0,316	44,082	15,782	9,353	21,046
2019	411	0,110	44,655	17,220	9,715	22,041
MW	265	0,342	41,186	13,510	8,680	17,712

Die Individuelle Spielerqualität pro Jahr geht in die Berechnung der *Hervorgebrachten Spielerqualität* eines Nachwuchsleistungszentrums ein.

4.1.2.3 Hervorgebrachte Spielerqualität

Die *Hervorgebrachte Spielerqualität* (Spielerqualität NLZ) repräsentiert den Nutzen, den ein Nachwuchsleistungszentrum mit der Ausbildung von Profifußballspielern erzeugt hat. In Abhängigkeit der Summe der Ausbildungstage aller Spieler im Nachwuchsleistungszentrum über alle betrachteten Saisons hinweg, gewichtet mit der durchschnittlichen *Individuellen Spielerqualität*, wird die Spielerqualität NLZ berechnet. Abbildung 48 zeigt den Verlauf der Anzahl Nachwuchsleistungszentren sowie den Median und Mittelwert der Hervorgebrachten Spielerqualitäten im Betrachtungszeitraum.

Abbildung 48: Anzahl Nachwuchsleistungszentren, Median und Mittelwert Hervorgebrachte Spielerqualitäten in Betrachtungszeitraum



Während zu Beginn des Betrachtungszeitraums in den ersten 5 Jahren die Anzahl der Nachwuchsleistungszentren stetig stieg, stabilisierte sich ihre Anzahl seit 2014 bei über 50 Nachwuchsleistungszentren, die in die Beobachtung einfließen konnten. Durch die vermehrte Anzahl ausgebildeter Spieler, deren Zeit im Profifußball und des Anstiegs der individuellen mittleren Spielerqualität (siehe Kapitel 4.1.2.2) steigen auch Mediane und Mittelwerte über die Jahre stetig an.

Tabelle 14 zeigt die deskriptive Statistik des Indikators *Hervorgebrachte Spielerqualität* im Laufe der Jahre 2009-2019. Eingegangen in diese Statistik, sind die Nachwuchsleistungszentren, die im Betrachtungszeitraum eine Hervorgebrachte Spielerqualität als Indikator ihrer Beteiligung an der Ausbildung eines Profifußballspielers ausweisen.

Tabelle 14: Deskriptive Statistik Hervorgebrachte Spielerqualität

Jahr	Anzahl NLZ	Min	Max	Mittelwert	STD	Median
2009	33	0,203	16,518	3,620	3,507	2,743
2010	40	0,224	23,637	6,817	5,892	6,517
2011	45	0,207	42,351	9,197	9,211	9,007
2012	48	0,117	51,054	11,608	12,688	11,851
2013	49	0,117	58,414	13,935	14,773	17,768
2014	51	0,117	72,765	15,639	16,929	19,209
2015	53	0,139	93,745	18,015	21,499	21,107
2016	53	0,139	100,798	20,321	24,430	24,217
2017	54	0,139	104,182	22,162	25,485	26,935
2018	55	0,139	115,108	23,829	27,280	26,900
2019	54	0,399	124,595	26,298	30,132	28,902
MW	49	0,176	73,015	15,586	17,439	17,741

Tabelle 15 zeigt ein Ranking der Nachwuchsleistungszentren, die im Betrachtungszeitraum eine Hervorgebrachte Spielerqualität als Indikator ihrer Beteiligung an der Ausbildung eines Profifußballspielers ausweisen.

Die Top-5-Nachwuchsleistungszentren nach der Summe der Hervorgebrachten Spielerqualität – FC Schalke 04, Bayer 04 Leverkusen, VfB Stuttgart, Borussia Dortmund und Eintracht Frankfurt – bleiben in ihren Rankings nach Mittelwert, der Summe der Hervorgebrachten Spielerqualität in den Pass-Zyklen 2.0, 3.0 und 4.0 ebenfalls immer unter den Top-5. Auch die letzten 15 Nachwuchsleistungszentren dieses Rankings weisen über den Beobachtungszeitraum eine hohe Stabilität in der Platzierung nach *Hervorgebrachter Spielerqualität* auf.

Hervorzuheben sind hier RasenBallSport Leipzig und der SC Paderborn 07, die im Pass-2.0-Zyklus mit den Rängen 50 (RB Leipzig) und 42 (Paderborn) eine hohe Abweichung von ihrem Rang nach der Summe der Hervorgebrachten Spielerqualität 23 (RB Leipzig) und 28 (Paderborn) liegen. D. h., die Nachwuchsleistungszentren haben zum Zeitpunkt des PASS-2.0-Zyklus im Vergleich zur Gesamtsumme recht wenig Spielerqualität hervorgebracht. Das spiegelt beim SC Paderborn 07 auch die Bewertung mit dem Double Pass Score wider, der ebenfalls relativ gering ausfällt, d. h., das Nachwuchsleistungszentrum bekam auch eine relativ niedrige Qualität nach den Maßstäben des DFB/der DFL bescheinigt.

Tabelle 15: Ranking NLZ nach Summe Hervorgebrachter Spielerqualität über Betrachtungszeitraum, Mittelwert per Jahr und Summen per Pass-Zyklen.

Verein	Ranking Hervorgebrachte Spielerqualität				
	Summe	Mittel-Wert	Pass2.0 Summe	Pass3.0 summe	Pass4.0 summe
FC Schalke 04	1	1	2	1	2
Bayer 04 Leverkusen	2	2	1	2	1
VfB Stuttgart	3	3	3	3	5
Borussia Dortmund	4	4	4	5	3
Eintracht Frankfurt	5	5	5	4	4
VfL Wolfsburg	6	6	7	6	7
FC Bayern München	7	7	11	7	6
Hertha BSC	8	8	6	9	8
TSG 1899 Hoffenheim	9	9	12	8	9
Hamburger SV	10	10	9	11	11
SV Werder Bremen	11	11	18	10	10
Hannover 96	12	13	17	12	15
1.FC Köln	13	14	14	21	13
FC Augsburg	14	12	15	14	14
1.FC Nürnberg	15	15	13	13	16
1.FSV Mainz 05	16	16	21	16	12
SC Freiburg	17	17	10	15	21
1.FC Kaiserslautern	18	18	8	18	20
TSV 1860 München	19	19	16	17	22
Karlsruher SC	20	20	19	19	17
Borussia Mönchengladbach	21	21	20	20	19
VfL Bochum	22	23	22	22	23
RasenBallSport Leipzig	23	22	50	35	18
FSV Frankfurt	24	24	31	23	24
FC Hansa Rostock	25	25	23	26	33
Holstein Kiel	26	26	32	25	25
FC Energie Cottbus	27	27	24	27	28
SC Paderborn 07	28	29	42	24	26
SG Dynamo Dresden	29	28	34	28	30
MSV Duisburg	30	34	25	30	37
1.FC Union Berlin	31	30	29	36	29
Rot-Weiss Essen	32	35	26	31	36
Arminia Bielefeld	33	31	39	40	27
SpVgg Greuther Fürth	34	37	28	38	35
Fortuna Düsseldorf	35	38	36	34	32
SV Wacker Burghausen	36	32	27	29	39
FC St. Pauli	37	39	35	37	34
VfL Osnabrück	38	33	40	32	31
1.FC Saarbrücken	39	36	30	33	38
SV Darmstadt 98	40	40	33	39	40
Rot-Weiß Erfurt	41	42	38	41	41
SV Wehen Wiesbaden	42	41	37	43	43
Stuttgarter Kickers	43	45	44	45	42
Alemannia Aachen	44	43	41	42	46
Rot-Weiß Oberhausen	45	44	46	44	44
SpVgg Unterhaching	46	47	43	46	47
Kickers Offenbach	47	46	48	47	45
SV Eintracht Trier 05	48	48	51	48	48
FC Carl Zeiss Jena	49	52	45	49	52
SSV Jahn Regensburg	50	49	52	50	49
FC Ingolstadt 04	51	53	49	51	50
Eintracht Braunschweig	52	54	47	52	54
1.FC Magdeburg	53	50	53	54	51
VfR Aalen	54	51	54	55	53
TuS Koblenz	55	55	55	53	55

Die Arbeit des Nachwuchsleistungszentrums des SC Paderborn entwickelte sich daraufhin erfolgreich weiter und konnte einen Sprung im Ranking der Hervorgebrachten Spielerqualität vom Pass-2.0-Zyklus zum 3.0-Zyklus von Rang 42 auf Rang 24 erarbeiten. RasenBallsport Leipzig begann erst im Jahre 2009 mit seinen Aktivitäten im deutschen Profifußball und konnte im Pass-2.0-Zyklus nur eine geringe Hervorgebrachte Spielerqualität aufgrund der eigenen Nachwuchsarbeit auf sich verbuchen. Der Erfolg der Arbeit des Nachwuchsleistungszentrums von RasenBallsport Leipzig nach Hervorgebrachter Spielerqualität kann über die Entwicklung hin zu Rang 35 im Pass-3.0-Zyklus und zu Rang 18 im Pass-4.0-Zyklus abgelesen werden. Arminia Bielefeld weist ebenfalls eine auffällige Entwicklung der Rangplätze auf: von Rang 39 im Pass-2.0-Zyklus zu Rang 40 im Zyklus 3.0 und dann zu Rang 27 – 13 Plätze besser als in den vorherigen 3-6 Jahren. Das Nachwuchsleistungszentrum von Arminia Bielefeld hat eine größere Anzahl an Spielern mit höherer Spielerqualität während des Betrachtungszeitraums in den Profifußball entlassen können.

Zu beobachten sind aber auch Vereine, die im Laufe des Betrachtungszeitraum nachgelassen haben und hinsichtlich des Erfolgskriteriums hohe Rückstufungen im Ranking der Nachwuchsleistungszentren hinnehmen mussten. Dies sind zum Beispiel der 1. FC Kaiserslautern, die Spielvereinigung Greuther Fürth und der SV Wacker Burghausen, die 10 Plätze verloren haben. Der 1. FC Kaiserslautern hat zwar 15 Spieler im Nachwuchsleistungszentrum zu Profifußballspielern ausgebildet, zu Beginn des Beobachtungszeitraums 3 mit höherer Spielerqualität, aber dann Spieler mit relativ geringer Spielerqualität hervorgebracht. Der SV Wacker Burghausen hat über die Jahre lediglich drei Spieler für den Lizenzbetrieb ausgebildet, darunter sind zwei Spieler mit höherer Spielerqualität hervorgebracht worden.

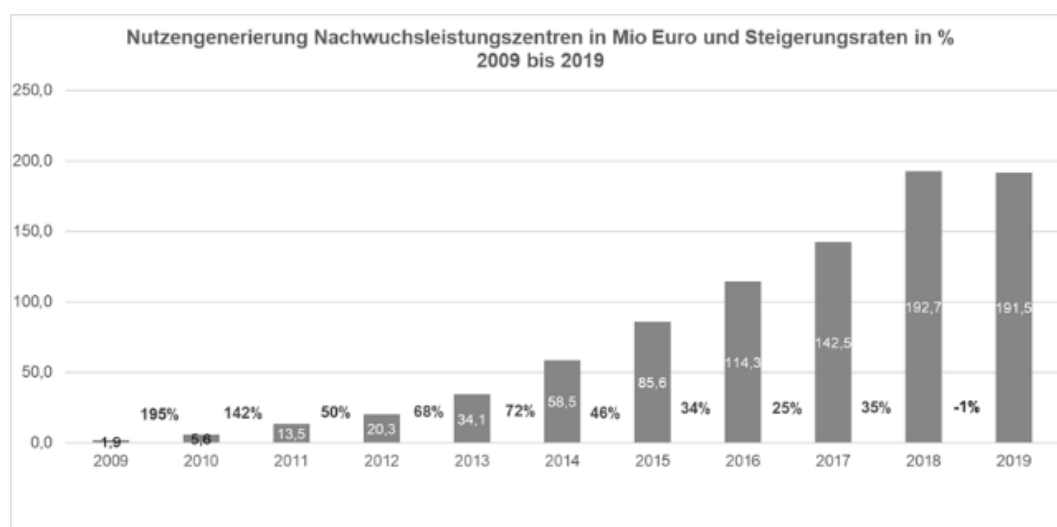
Der Indikator *Hervorgebrachte Spielerqualität* konnte auf insgesamt 55 Vereine angewandt werden, da aus diesen Vereinen Spieler hervorgegangen sind, die den Auswahlkriterien entsprechen. Davon sind 49 Vereine, die zur Grundgesamtheit der Nachwuchsleistungszentren (N=54) gezählt werden. 6 Vereine fallen somit aus der Tabelle 15 (siehe blau markierte Vereine) heraus und werden im weiteren Verlauf der Arbeit nicht weiter betrachtet. Das sind: Wacker Burghausen (Platz 36), Saarbrücken (39), Alemannia Aachen (44), Trier (48), VfR Aalen (54), TUS Koblenz (55). Für 5 Vereine, die Nachwuchsleistungszentren betreiben, konnten keine den Kriterien genügenden Profifußballspieler identifiziert werden. Das heißt, für die Nachwuchsleistungszentren 1.FC Heidenheim, Chemnitzer FC, Erzgebirge Aue, SV Elversberg und SV Sandhausen konnte keine Hervorgebrachte Spielerqualität berechnet werden.

4.1.2.4 Nutzen – Monetär bewertete Hervorgebrachte Spielerqualität

Der Nutzen konnte für den Beobachtungszeitraum für 55 Vereine durch die Multiplikation der Hervorgebrachten Spielerqualität mit dem Marktwertfaktor pro Jahr ermittelt werden. Die weiteren Auswertungen fokussieren auf die 49 Vereine, die auch als Nachwuchsleistungszentrum gelistet sind. Das Ranking der Vereine verändert sich dadurch ab Platz 36. Eine Veränderung des Rankings durch die monetäre Bewertung erfolgt

nicht, da mit denselben Marktwertfaktoren relativiert wurde. Durch diese 49 Vereine wurde über den Beobachtungszeitraum mit der Ausbildung von Profifußballspielern insgesamt 860 Mio. Euro Nutzen generiert. Seit 2009 ist ein stetiges Wachstum zu verzeichnen, 2019 ein leichter Rückgang von 1 %. Das Jahr 2019 war durch einen niedrigeren Marktwertfaktor geprägt. Die ersten 3 Jahre des Beobachtungszeitraum steigerte sich der Nutzen durchschnittlich um 169 % (2009-2011), die folgenden 3 Jahre (2012-2014) dagegen lagen die Steigerungsraten im höheren 2-stelligen Bereich (63 %). Im Anschluss daran (2015-2018) stabilisierte sich die Wachstumsrate bei durchschnittlich 35 % (siehe Abbildung 49).

Abbildung 49: Nutzengenerierung: 49 Nachwuchsleistungszentren 2009-2019



Den generierten Nutzen für die untersuchten Nachwuchsleistungszentren in Euro zeigt Tabelle 16.

Tabelle 16: Nutzen Nachwuchsleistungszentren in Euro.

Verein	Nutzen Nachwuchsleistungszentrum in Euro (sortiert nach Ranking Summe Hervorgebrachter Spielerqualität)											Summe 2009 bis 2019
	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	
FC Schalke 04	138.213	487.227	1.097.835	1.779.520	2.971.462	5.439.136	8.251.376	10.866.646	12.573.710	15.402.624	15.433.501	74.441.250
Bayer 04 Leverkusen	57.009	262.189	1.407.609	1.893.784	2.863.410	4.677.180	8.547.079	10.230.680	11.852.101	17.145.318	17.009.033	75.945.392
VfB Stuttgart	70.146	298.379	845.822	1.299.477	2.467.133	3.464.096	6.208.788	7.730.279	8.731.816	11.370.315	10.579.981	53.056.232
Borussia Dortmund	260.846	452.289	894.044	1.421.576	2.093.897	3.211.775	4.248.666	6.792.987	8.789.803	11.894.071	11.779.636	51.839.590
Eintracht Frankfurt	17.941	341.508	642.683	1.467.566	2.120.000	3.709.270	4.468.112	7.963.398	8.349.815	10.788.717	12.053.399	51.922.510
VfL Wolfsburg	42.707	285.013	577.098	787.410	1.395.954	2.754.383	4.339.344	6.604.274	6.729.470	8.534.123	9.142.931	41.192.706
FC Bayern München	113.151	201.111	419.010	651.066	1.208.859	2.514.831	3.542.910	4.606.828	7.526.537	9.924.798	9.384.479	40.093.559
Hertha BSC	70.751	282.374	544.668	885.990	1.339.555	2.333.924	2.586.147	4.562.468	5.951.572	7.571.952	8.941.256	35.070.656
TSG 1899 Hoffenheim	103.653	224.169	450.088	563.410	1.153.355	2.119.389	3.385.023	4.114.209	5.424.079	7.690.100	7.265.633	32.493.108
Hamburger SV	167.937	296.403	513.213	573.539	1.264.142	2.090.770	2.888.325	3.721.141	4.980.232	5.948.751	5.678.487	28.122.939
SV Werder Bremen	39.888	134.339	259.491	439.600	1.055.388	1.964.393	2.853.271	4.082.866	4.642.795	6.250.764	6.782.656	28.505.351
Hannover 96	22.120	163.168	342.002	438.991	1.084.357	1.712.305	3.081.731	3.175.553	4.176.690	5.232.260	5.062.361	24.782.738
LFC Köln	130.413	227.229	470.668	607.359	903.831	1.435.857	1.569.699	2.299.595	3.626.776	5.149.941	7.272.065	23.693.435
LFC Augsburg	77.317	421.844	510.223	1.082.964	1.726.312	2.525.518	3.104.401	4.169.787	6.290.240	4.563.748	24.472.354	
LFC Nürnberg	22.669	152.488	362.359	646.705	1.036.753	1.987.715	2.661.559	3.063.378	3.354.681	4.462.108	4.174.742	21.976.956
LFSV Mainz 05	12.014	132.671	299.355	376.865	676.161	1.277.533	2.130.759	2.866.452	4.431.729	6.337.488	5.435.990	24.002.065
SC Freiburg	12.014	70.783	286.411	818.459	1.209.799	1.900.609	2.359.744	2.793.890	3.154.233	4.006.825	3.395.682	19.948.448
LFC Kaiserslautern	43.322	56.547	655.498	737.599	1.011.524	1.532.799	2.065.136	2.442.806	3.259.829	3.839.862	3.477.655	18.113.577
TSV 1860 München	55.454	221.250	390.510	599.497	936.735	1.577.743	1.924.406	2.610.793	3.004.552	3.724.414	3.045.209	30.090.562
Karlsruher SC	57.736	95.088	296.411	498.591	667.130	1.347.593	1.701.559	2.698.976	3.353.934	4.370.239	3.995.259	19.082.817
Borussia Mönchengladbach	76.729	166.935	312.513	414.932	615.515	1.390.104	1.871.807	2.232.933	3.011.967	4.005.410	3.945.505	18.044.351
VfL Bochum	51.975	67.842	168.114	390.175	538.432	791.194	1.200.619	1.522.962	1.747.051	2.759.691	2.510.958	11.729.014
Rasenballsport Leipzig		95.198	104.536	171.311	234.932	327.977	470.247	827.143	1.109.010	4.991.925	5.893.228	13.324.350
FSV Frankfurt		76.839	136.487	220.073	341.615	468.482	688.406	893.682	1.608.245	1.564.327	2.310.958	10.252.224
FC Hansa Rostock		36.955	48.237	77.778	86.804	379.449	379.449	757.801	896.042	1.111.497	1.155.284	6.788.785
Holstein Kiel		40.198	136.888	254.334	381.199	560.149	683.225	1.034.943	1.156.612	1.444.699	2.017.634	8.689.674
FC Energie Cottbus		4.874	6.362	10.258	11.448	208.525	615.763	1.070.979	1.073.341	1.402.108	1.730.649	7.758.502
SC Paderborn 07		3.206	88.991	146.018	227.721	312.290	458.892	559.720	662.130	741.264	1.036.909	5.187.478
SG Dynamo Dresden		12.121	15.821	134.388	191.791	263.017	411.089	389.078	460.056	1.369.307	1.144.137	5.816.937
MSV Duisburg		12.121	15.821	134.388	191.791	263.017	411.089	389.078	460.056	1.369.307	1.144.137	5.816.937
LFC Union Berlin		12.121	15.821	134.388	191.791	263.017	411.089	389.078	460.056	1.369.307	1.144.137	5.816.937
Rot-Weiss Essen		12.121	15.821	134.388	191.791	263.017	411.089	389.078	460.056	1.369.307	1.144.137	5.816.937
Arminia Bielefeld		24.194	55.072	146.240	163.212	223.824	254.071	401.162	474.343	940.856	1.012.059	927.563
SpVgg Greuther Fürth		4.585	36.105	85.525	101.749	146.521	351.097	428.240	589.237	917.141	1.237.151	1.255.560
Fortuna Düsseldorf		49.122	68.950	111.176	133.273	84.857	289.936	365.812	467.086	523.910	1.371.026	1.256.561
FC St. Pauli		49.122	68.950	111.176	133.273	84.857	289.936	365.812	467.086	523.910	1.371.026	1.256.561
VfL Osnabrück		54.041	77.940	121.723	149.583	205.134	371.875	439.714	463.263	691.065	659.868	3.528.235
SV Darmstadt 98		15.931	22.947	37.000	106.958	166.430	244.559	340.778	503.258	678.322	621.690	3.036.168
Rot-Weiß Erfurt		48.153	77.643	86.654	145.934	214.442	261.559	309.274	378.930	604.293	780.374	2.857.206
SV Wehen Wiesbaden			35.422	39.533	54.215	135.140	228.273	269.916	377.924	741.911	679.970	2.562.304
Stuttgarter Kickers			17.495	317.435	67.940	99.834	268.261	317.198	484.813	598.333	548.379	2.402.195
FC Carl Zeiss Jena		8.702	11.358	18.314	20.440	48.308	70.986	102.378	114.614	141.451	54.419	677.553
SSV Jahn Regensburg												
FC Ingolstadt 04				4.337	5.948	8.740	14.257	186.092	208.332	23.292	235.648	686.645
Eintracht Braunschweig		3.540	4.620	16.515	18.432	25.277	37.142	45.303	53.568	59.970	74.012	406.212
LFC Magdeburg												
Summe	1.886.297	5.572.909	13.499.974	20.283.995	34.068.085	58.538.890	85.566.566	114.342.253	142.453.590	192.650.375	191.508.579	860.371.513

4.1.3 Return on Investment Nachwuchsleistungszentrum

Der Berechnung des Return on Investment (ROI) für Nachwuchsleistungszentren im deutschen Profifußball erfolgt durch Anwendung der Formel 5 dargestellt in Kapitel 3.6.3 dieser Arbeit. Der Nettutzen in Euro dividiert durch die Kosten in Euro multipliziert mit 100 ergibt den ROI in % pro Nachwuchsleistungszentrum. Der ROI wird lediglich für einen Teil der Grundgesamtheit der Nachwuchsleistungszentren kalkuliert, da folgende Kriterien erfüllt sein müssen: Vereine müssen zur Grundgesamtheit der Nachwuchsleistungszentren gehören, Kosten ausweisen können (siehe Kapitel 4.1.1) und – wie in Kapitel 4.1.2.3 beschrieben – Spielerqualitäten hervorgebracht haben (n=39, siehe Tabelle 17).

Ein Ranking der Vereine anhand ihres Return on Investment zeigt Tabelle 18. Im Durchschnitt über alle 39 Vereine weisen die untersuchten Nachwuchsleistungszentren einen ROI von -44 % aus. Das negative Ergebnis bedeutet, dass die Kosten die Erträge überwiegen, für jeden in das Nachwuchsleistungszentrum investierten Euro sind 0,44 Euro zusätzlich aufzubringen.

Lediglich 4 Vereine weisen einen positiven ROI aus: TSV 1860 München (42 %), Eintracht Frankfurt (13%), VfB Stuttgart (7 %) und der Karlsruher SC (1 %). Diese 4 Vereine weisen in den ersten 5 bis 7 Jahren des Beobachtungszeitraums (11 Jahre) negative ROIs von -63 % bis -67% aus, während 4 bis 6 Jahre außerordentlich hohe ROIs erzielt werden konnten.

Der TSV 1860 München weist in den 6 Jahren mit positivem ROI im Mittel einen ROI von 162 % aus, stark geprägt durch den ROI von 366 % im Jahr 2018, aber auch von einem ROI von 182 % in 2 von diesen 6 Jahren. Im Jahr 2018 hat der TSV 1860 München bedingt durch die wirtschaftliche Lage lediglich 832 Tsd. Euro in das Nachwuchsleistungszentrum investiert, aber seinen höchsten Nutzen von 3,7 Mio. Euro erzielt. Auch der Karlsruher SC mit einem ROI von 1 % über die Gesamtzeit erzielt im Jahr 2018 einen sehr hohen positiven ROI (993 %) nach 7 Jahren in Folge mit negativen ROIs von im Mittel 63 % und 2 Jahren mit einem im Mittel positiven ROI von 110 %. Der Karlsruher SC hat in 2018, ebenso wie der TSV 1860 München, sehr niedrige Kosten für das Nachwuchsleistungszentrum ausgewiesen, aber auch den höchsten Nutzen im Beobachtungszeitraum hervorgebracht. Der Karlsruher SC, Hertha BSC Berlin und Bayer Leverkusen mit ihren ROIs um die +/- 1 % haben mit ihren Investitionen einen Ertrag in nahezu gleicher Höhe erzielen können. Die übrigen Vereine mit negativem Return on Investment sind mit der Ausbildung von Spielern für den Profifußball gemessen an ihrem ROI wirtschaftlich nicht erfolgreich.

Tabelle 17: Vereine mit Nachwuchsleistungszentrum, Kosten- und Nutzenkalkulation.

Verein	NLZ	Kosten	Nutzen	ROI Berechnung
1.FC Kaiserslautern	x	x	x	ja
1.FC Köln	x	x	x	ja
1.FC Magdeburg	x	x	x	ja
1.FC Nürnberg	x	x	x	ja
1.FSV Mainz 05	x	x	x	ja
1860 München	x	x	x	ja
Arm. Bielefeld	x	x	x	ja
Bay. Leverkusen	x	x	x	ja
Bayern München	x	x	x	ja
Bor. Dortmund	x	x	x	ja
Bor. M'gladbach	x	x	x	ja
Darmstadt 98	x	x	x	ja
Dynamo Dresden	x	x	x	ja
E. Frankfurt	x	x	x	ja
F. Düsseldorf	x	x	x	ja
FC Augsburg	x	x	x	ja
FC Ingolstadt	x	x	x	ja
FC Schalke 04	x	x	x	ja
FC St. Pauli	x	x	x	ja
Greuther Fürth	x	x	x	ja
Hamburger SV	x	x	x	ja
Hannover 96	x	x	x	ja
Hansa Rostock	x	x	x	ja
Hertha BSC	x	x	x	ja
Holstein Kiel	x	x	x	ja
Jahn Regensburg	x	x	x	ja
Karlsruher SC	x	x	x	ja
MSV Duisburg	x	x	x	ja
RB Leipzig	x	x	x	ja
SC Freiburg	x	x	x	ja
SC Paderborn	x	x	x	ja
TSG Hoffenheim	x	x	x	ja
Union Berlin	x	x	x	ja
VfB Stuttgart	x	x	x	ja
VfL Bochum	x	x	x	ja
VfL Osnabrück	x	x	x	ja
VfL Wolfsburg	x	x	x	ja
Wehen Wiesbaden	x	x	x	ja
Werder Bremen	x	x	x	ja
1.FC Heidenheim	x			
Carl Zeiss Jena	x		x	
Chemnitzer FC	x			
E. Braunschweig	x		x	
Energie Cottbus	x		x	
Erzgebirge Aue	x			
FSV Frankfurt	x		x	
K. Offenbach	x		x	
Rot-Weiß Erfurt	x		x	
RW Essen	x		x	
RW Oberhausen	x		x	
Stuttg. Kickers	x		x	
SV Elversberg	x			
SV Sandhausen	x			
Unterhaching	x		x	
Alemannia Aachen		x	x	
RW Ahlen		x		
TB Berlin		x		
TUS Koblenz		x	x	
VfR Aalen			x	
Wacker Burghausen			x	
Saarbrücken			x	
Trier			x	

Tabelle 18: ROI der Vereine mit Nachwuchsleistungszentrum.

		ROI (%) Nachwuchsleistungszentren													Summe
		2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2019	Summe	
Ranking	Liga in 2018/2019	Verein	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Summe	
1	3	TSV 1860 München	-94	-84	-74	-56	-24	54	69	126	187	348	187	42	
2	1	Eintracht Frankfurt	-99	-87	-71	-54	-42	-8	9	53	33	61	124	13	
3	1	VfB Stuttgart	-98	-93	-79	-69	-38	-25	27	86	64	86	211	7	
4	3	Karlsruher SC	-97	-96	-87	-77	-65	-16	-4	49	170	993	130	1	
5	1	Hertha BSC	-98	-83	-78	-37	-49	-17	-14	10	20	46	116	-1	
6	1	Bayer 04 Leverkusen	-98	-93	-72	-66	-56	-35	-2	36	30	79	70	-1	
7	1	FC Augsburg	-100	88	-67	-65	-38	-21	-29	-11	23	78	78	5	
8	1	1.FC Nürnberg	-99	-94	-88	-74	-63	-36	96	128	99	70	110	-9	
9	1	Hannover 96	-98	-94	-84	-84	-77	-63	-44	-16	23	1	59	148	-20
10	1	FC Schalke 04	-98	-94	-84	-77	-63	-44	-16	23	1	59	147	-21	
11	1	SC Freiburg	-99	-95	-84	-56	-53	-35	30	19	-11	18	6	-23	
12	1	TSG 1899 Hoffenheim	-95	-92	-86	-78	-53	-20	-28	0	-9	29	-14	-27	
13	2	Holstein Kiel	-96	-95	-92	-90	-55	-46	-6	3	14	18	18	-31	
14	1	1.FSV Mainz 05	-97	-91	-88	-85	-76	-54	-45	-29	6	19	28	-31	
15	3	VfL Osnabrück	-100	-100	-96	-96	-69	-43	-9	9	30	51	54	-31	
16	1	SV Werder Bremen	-99	-97	-92	-86	-69	-48	-28	-10	10	12	55	-37	
17	2	VfL Bochum	-97	-96	-89	-75	-66	-50	-35	-20	-8	41	33	-38	
18	3	1.FC Kaiserslautern	-98	-98	-84	-80	-68	-43	-31	-20	14	190	192	-39	
19	2	Arminia Bielefeld	-100	-75	-60	-100	-65	-52	-54	-54	-32	-36	-8	-41	
20	1	VfL Wolfsburg	-99	-95	-84	-86	-79	-63	-51	-6	-17	19	17	-44	
21	2	SC Paderborn 07	-100	-100	-99	-99	-85	-64	-22	167	96	13	12	-49	
22	3	FC Hansa Rostock	-93	-91	-87	-77	-65	-39	-39	-39	-3	-17	-32	-54	
23	1	MSV Duisburg	-100	-94	-85	-74	-62	-54	-48	-45	-46	-24	-22	-59	
24	2	SC Dynamo Dresden	-100	-100	-92	-92	-82	-54	-32	-58	-33	-33	-27	-59	
25	1	Borussia Dortmund	-93	-91	-87	-86	-79	-69	-69	-55	-54	-32	-33	-60	
26	2	1.FC Köln	-97	-94	-90	-83	-78	-72	-74	-69	-61	-16	10	-61	
27	2	Hamburger SV	-96	-96	-93	-92	-76	-72	-59	-47	-35	-18	4	-65	
28	1	Borussia Mönchengladbach	-97	-94	-93	-88	-87	-77	-74	-66	-50	-46	-31	-68	
29	2	SpVgg Greuther Fürth	-97	-94	-91	-91	-85	-83	-72	-68	-43	-34	-28	-70	
30	2	1.FC Union Berlin	-100	-93	-86	-87	-86	-79	-78	-79	-49	-67	-62	-73	
31	1	Fortuna Düsseldorf	-100	-96	-92	-95	-93	-85	-60	-60	-36	-52	-58	-73	
32	3	SV Wehen Wiesbaden	-100	-96	-94	-92	-85	-74	-72	-67	-61	-44	-19	-74	
33	1	Rasenballsport Leipzig	-100	-100	-100	-100	-100	-99	-89	-89	-88	-51	-51	-74	
34	2	FC St. Pauli	-97	-96	-94	-92	-95	-85	-84	-83	-79	-52	-54	-80	
35	1	FC Bayern München	-99	-98	-97	-96	-94	-87	-85	-81	-69	-62	-60	-81	
36	2	SV Darmstadt 98	-97	-97	-94	-92	-89	-86	-84	-84	-77	-61	-63	-84	
37	2	SSV Jahn Regensburg	-100	-100	-100	-100	-100	-100	-83	-82	-82	-81	-86	-90	
38	2	1.FC Magdeburg	-100	-100	-100	-100	-100	-100	-100	-100	-95	-89	-82	-95	
39	2	FC Ingolstadt 04	-100	-100	-100	-100	-100	-100	-100	-100	-95	-99	-78	-96	

Die Werte eines ROI für ein Programm für Personalentwicklungsmaßnahmen werden in der Praxis häufig recht hoch zwischen 25 % und 500 % ausgewiesen und widerspiegeln die möglichen Wirkungen eines erfolgreichen Programms (Phillips & Schirmer, 2008).

4.2 Ergebnis zur Forschungsfrage

Die Forschungsfrage *Wie kann der Erfolg von Nachwuchsleistungszentren im deutschen Profifußball mittels Investitionskennzahl und angemessenem Erhebungsaufwand valide evaluiert werden?* kann wie folgt beantwortet werden:

Zur Messung des Erfolgs von Nachwuchsleistungszentren im deutschen Profifußball ist es möglich, anhand von quantitativen Daten und dem Konstrukt geeigneter Indikatoren den Return on Investment als eine geeignete Investitionskennzahl zu ermitteln. Öffentlich verfügbare Daten in digitaler Form und grundsätzliche Schätzungen auf Basis von Informationen von Experten der Nachwuchsausbildung im deutschen Profifußball ermöglichen es, mit relativ geringem Erhebungsaufwand Daten für einen längeren Betrachtungszeitraum zu erheben.

Die verwendeten Daten weisen eine hohe Güte auf (siehe Kapitel 3). Für entwickelte Konstrukte – Vereinsqualität, Individuelle Spielerqualität, Hervorgebrachte Spielerqualität – kann ebenfalls eine hohe Validität nachgewiesen werden (siehe Kapitel 3).

4.3 Metaevaluation gemäß der Standards der DeGEval

Der Zweck der Metaevaluation besteht darin, Stärken und Schwächen in dem in dieser Studie entwickelten Evaluationsverfahren aufzudecken

Die retrospektive Metaevaluation der Evaluation des Erfolgs von Nachwuchsleistungszentren wird anhand der Checkliste zur Anwendung der Standards der DeGEval durchgeführt (siehe dazu DeGEval, 2017a). Einerseits scheint eine systematische Anwendung und Dokumentation aller Standards nicht sinnvoll oder möglich, andererseits bietet die Checkliste eine Reihe von Optionen in den Bereichen *Nützlichkeit*, *Durchführbarkeit*, *Fairness* und *Genauigkeit*, um zu bewerten, inwieweit der Standard eingehalten wurde (Sanders, 2000). In einem ersten Schritt werden für jeden Standard die bei dieser Evaluation gewählte Vorgehensweise dokumentiert und Stärken und Schwächen dokumentiert, im zweiten Schritt auf Basis des Expertenwissens und der Stärken-/Schwächen-Analyse bewertet, ob und in welchem Maße der Standard beachtet wurde (anwendbar und eingehalten, teilweise eingehalten oder nicht eingehalten sowie nicht anwendbar). Im nachfolgenden dritten Schritt werden zur Erhöhung von Qualität und Effektivität der Evaluation erforderliche Handlungsempfehlungen für die Verbesserung zukünftiger Evaluationen anhand des entwickelten Konzepts diskutiert. Im Folgenden wird das Ergebnis der Metaevaluation im Überblick dargestellt (siehe dazu Tabelle 19 am Ende dieses Kapitels).

4.3.1 Nützlichkeit

Die Standards der *Nützlichkeit* sollen sicherstellen, dass durch die Evaluation die Informationsbedürfnisse auf die vorgesehenen Evaluationsnutzer ausgerichtet werden:

Der Standard *Identifizierung der Beteiligten und Betroffenen* (N1) war anwendbar und wurde eingehalten. Die Ermittlung der beteiligten und betroffenen Personen oder Organisationen (Stakeholder) erfolgte für diese Studie zunächst über die Personen in führenden Positionen der Nachwuchsförderung bei DFB und DFL. Ihre Erfahrungen und Bedürfnisse flossen in die Stakeholder-Analyse ein, um herauszufinden, welche Bedeutung dieser Evaluation beigemessen werden kann, welche Informationen relevant sein und wie die Ergebnisse verwendet werden könnten.

Der Standard *Klärung der Evaluationszwecke* (N2) war anwendbar und wurde eingehalten. Der Zweck der Evaluation wurde deutlich bestimmt, so dass die Stakeholder und die Evaluierende eine klare Linie verfolgen konnten.

Der Standard *Kompetenz und Glaubwürdigkeit der Evaluator*innen* (N3) war anwendbar und wurde eingehalten. Die Evaluatorin ist als wissenschaftliche Mitarbeiterin mit langjähriger Berufserfahrung im Controlling (Value- und Portfolio-Management) sowohl vertrauenswürdig als auch kompetent, so dass den Evaluationsergebnissen ein hohes Maß an Glaubwürdigkeit und Akzeptanz bescheinigt werden kann.

Der Standard *Umfang und Auswahl der Informationen* (N4) war anwendbar und wurde eingehalten. Die durch die Evaluationen gewonnenen Informationen ermöglichen die Beantwortung der Fragen zum Erfolg der Nachwuchsleistungszentrum mittels Investitionskennzahlen und gehen gleichzeitig auf die Interessen und Bedürfnisse der Stakeholder ein.

Der Standard *Feststellung von Werten* (N5) war anwendbar und wurde eingehalten. Eine sorgfältige Beschreibung der Betrachtungsweisen und Verfahren, auf denen die Interpretationen der Ergebnisse beruhen, gewährleistet eine fundierte Grundlage der gefällten Werturteile.

Der Standard *Klarheit des Berichts* (N6) war anwendbar und wurde eingehalten. Die vorliegende Dissertationsschrift beschreibt ausführlich das evaluierte Programm Nachwuchsförderung einschließlich des wissenschaftlichen und wirtschaftlichen Kontextes sowie die Ziele, das Vorgehen und Ergebnisse der Evaluation, um die wesentlichen Informationen verständlich zur Verfügung zu stellen.

Der Standard *Rechtzeitigkeit und Verbreitung des Berichts* (N7) war anwendbar und wurde eingehalten. Als wichtiges Zwischenergebnis wurde ein Artikel zum Thema dieser Arbeit in einem Journal mit doppeltem Peer Review veröffentlicht und mit ausgewählten Stakeholdern der Evaluation diskutiert. Der Schlussbericht erfolgt mit Vorlage dieser Dissertationsschrift und ihrer Publikation und wird so den vorgesehenen Nutzern zur Kenntnis gebracht werden.

Der Standard *Wirkung der Evaluation* (N8) war anwendbar und wurde teilweise eingehalten. Die Planung, Durchführung und Darstellung der Evaluation führte nicht dazu, dass alle Stakeholder zeitnah zum Evaluationsprozess informiert werden konnten.

4.3.2 Durchführbarkeit

„Die *Durchführbarkeitsstandards* sollen sicherstellen, dass eine Evaluation realistisch, gut durchdacht, diplomatisch und kostenbewusst ausgeführt wird“ (Sanders, 2000, S. 64):

Der Standard *Angemessene Verfahren* (D1) war anwendbar und wurde eingehalten. Anhand der Auseinandersetzung mit dem Evaluationsmodell und dem Untersuchungsdesign sowie der Beschreibung der Datenerhebung konnte sichergestellt werden, dass die Evaluation störungsfrei durchgeführt und die benötigten Informationen beschafft werden konnten.

Der Standard *Diplomatisches Vorgehen* (D2) war anwendbar, wurde aber nicht eingehalten. Die Planung und Durchführung der Evaluation erfolgte auf Basis der Zielsetzung der Studie – Entwicklung eines Evaluationskonzepts – und berücksichtigte dazu insbesondere die Positionen der Stakeholder des DFB und der DFL. Eine Kooperation mit anderen Stakeholdern wurde nicht eingegangen.

Standard *Effizienz von Evaluationen* (D3) war anwendbar und wurde teilweise eingehalten. Ziel dieser Studie war es, ein Evaluationskonzept zu entwickeln, das unter anderem öffentlich zugängliche Daten nutzt und mit geringem Aufwand leicht anzuwenden ist. Die für die Beschaffung der Informationen eingesetzten Mittel sollten allerdings überdacht werden.

4.3.3 Fairness

Die Fairnessstandards sollen sicherstellen, dass in einer Evaluation respektvoll und fair mit allen Stakeholdern umgegangen wird (DeGEval, 2017b, S. 41):

Der Standard *Formale Vereinbarungen* (F1) war nicht anwendbar. Es wurden keine formalen Vereinbarungen bezüglich Finanzen, Zeit, Methodik und Veröffentlichungsrechte mit Auftraggebern oder anderen Stakeholdern getroffen. Die Evaluation erfolgt anhand öffentlich zugänglicher Daten. Die Methodik der Evaluation ist im Rahmen dieser Arbeit beschrieben worden und wird zu gegebener Zeit publiziert.

Der Standard *Schutz individueller Rechte* (F2) war anwendbar und wurde eingehalten. Evaluation kann elementare Personenrechte berühren, insbesondere das Aufzeigen von Schwächen oder ein Ranking kann zu massiven Beeinträchtigungen des Interessenschutzes führen; das gilt auch für Organisationen. Die Evaluatorin hat durch Anonymisierung der Daten darauf geachtet, dass Personen, mit denen sie im Verlauf der Evaluation in Kontakt getreten ist, nicht in ihrer Würde verletzt wurden. Darüber hinaus wurden die gesetzlichen und organisationsbezogenen Bestimmungen zum Datenschutz berücksichtigt.

Der Standard *Umfassende und faire Überprüfung* (F3) war anwendbar und wurde eingehalten. Die Methoden, die in der Evaluation benutzt wurden, sind von der Evaluatorin mit ihren Stärken und Schwächen beschrieben worden, um eine Bewertung der Auswirkun-

gen auf die Ergebnisse und Schlussfolgerungen vornehmen zu können. Bei der Beurteilung oder der Interpretation der Ergebnisse werden diese neutral dargestellt. Eine Diskussion der Ergebnisse erfolgt im Rahmen dieser Dissertationsschrift.

Der Standard *Unparteiische Durchführung und Berichterstattung* (F4) war anwendbar und wurde eingehalten. Der Evaluationsprozess wurde unparteiisch gestaltet und die Begründung der Ergebnisse nachvollziehbar dargelegt. Die Evaluation erfolgte auf Basis der Erhebung öffentlich zugänglicher Daten, die Konstrukte sind offen und transparent dargestellt worden. Somit war eine große Distanz zum Evaluationsgegenstand gewährleistet.

Der Standard *Offenlegung von Ergebnissen und Berichten* (F5) war anwendbar und wurde eingehalten. Die Evaluationsergebnisse und ihre relevanten Beschränkungen werden den Stakeholdern der Evaluation über die Publikation zugänglich gemacht werden. Der Evaluationsbericht enthält Bewertungen und Empfehlungen sowie die Dokumentation von Limitationen und eine Diskussion der Ergebnisse. Die Publikation der Evaluationsergebnisse erfolgt nach den Kriterien des wissenschaftlichen Arbeitens.

4.3.4 Genauigkeit

„Die Genauigkeitsstandards sollen sicherstellen, dass eine Evaluation gültige und nachvollziehbare Informationen und Ergebnisse zu dem jeweiligen Evaluationsgegenstand und den Evaluationsfragestellungen hervorbringt und vermittelt. In diesem Zusammenhang kommt wissenschaftlichen Gütekriterien eine besondere Bedeutung zu“ (DeGEval, 2017b, S. 44).

Der Standard *Beschreibung des Evaluationsgegenstands* (G1) war anwendbar und wurde eingehalten. Der Evaluationsgegenstand ist hinsichtlich der beteiligten Akteur*innen, seiner Ziele und Aktivitäten, Strukturen und Ressourcen (fokussiert auf die finanziellen) umfassend beschrieben worden. Darüber hinaus ist ein Wirkmodell des Evaluationsgegenstands dargestellt worden. Der Evaluationsgegenstand „Nachwuchsförderung“ ist klar und genau beschrieben und dokumentiert, sodass eine eindeutige Identifikation erfolgen kann.

Der Standard *Kontextanalyse* (G2) war anwendbar und wurde eingehalten. Der Kontext des Evaluationsgegenstandes – die Rahmenbedingungen, die die Nachwuchsleistungszentren umgeben – sind eingehend beschrieben worden. Fokussiert wurde auf die Faktoren, die die Wirkung des Evaluationsgegenstands maßgeblich beeinflussen können (ökonomische, soziale, gesetzliche Faktoren).

Der Standard *Beschreibung von Zwecken und Vorgehen* (G3) war anwendbar und wurde eingehalten. Bei der Beschreibung der Zwecke und der Fragestellungen der Evaluation sind mehrere Perspektiven berücksichtigt worden. Die geplante Vorgehensweise hat sich im Laufe des Evaluationsprozesses nicht verändert.

Der Standard *Angabe von Informationsquellen* (G4) war anwendbar und wurde eingehalten. Es sind mehrere unterschiedliche Informationsquellen herangezogen worden, um die gewonnenen Informationen vergleichen und abwägen zu können. Diese Informationsquellen sind beschrieben und insbesondere im Einzelnen auf ihre Güte hin bewertet worden.

Der Standard *Valide und reliable Informationen* (G5) war anwendbar und wurde eingehalten. Die eingesetzten, qualitativen (Expertenschätzung) wie quantitativen Instrumente der Datenerhebung sowie die Ermittlung der Indikatoren (Konstrukte) stellen konsistente, reproduzierbare, nachvollziehbare, zuverlässige und valide Informationen bereit. Zur Bewertung der Informationen wurden die wissenschaftlichen Gütekriterien herangezogen.

Der Standard *Systematische Fehlerprüfung* (G6) war anwendbar und wurde eingehalten. Der Evaluationsprozess wurde so gestaltet, dass potentielle Fehlerquellen frühzeitig erkannt und Fehler so weit wie möglich vermieden oder korrigiert werden konnten. Im Prozess der Evaluation fanden daher Peer-Reviews statt, um z. B. in Untersuchungsdesign, Stichprobe/Grundgesamtheit und Missachtung der Anwendungsvoraussetzungen statistischer Verfahren Fehlerquellen zu identifizieren und deren Konsequenzen im Rahmen der Berichterstattung zur Evaluation zu diskutieren.

Der Standard *Angemessene Analyse qualitativer und quantitativer Informationen* (G7) war anwendbar und wurde eingehalten. Die Auswahl der adäquaten Auswertungsverfahren wurde an der Fragestellung der Evaluation, nach dem Erkenntnisstand und kontextbezogen auf den Evaluationsgegenstand (Nachwuchsleistungszentrum) ausgerichtet. Die herangezogenen Regeln und deren methodische Grundlagen sind durch Literaturhinweise offengelegt worden. Auf die Aussagekraft der Methoden, wie auch auf ihre Begrenzungen, wird explizit hingewiesen.

Der Standard *Begründete Bewertungen und Schlussfolgerungen* (G8) war anwendbar und wurde eingehalten. Die Bewertung des Evaluationsgegenstands basiert auf einem Kriterium, in dem sich seine Güte – in dieser Studie der ROI des Nachwuchsleistungszentrums – darstellt. Die Höhe des ROI, als Indikator für die Messung des monetären Erfolgs, dient als Kriterium, um ein Ranking der Nachwuchsleistungszentren im deutschen Profifußball zu ermöglichen. Die Bewertung der Nachwuchsleistungszentren stützt sich damit auf die ermittelten Daten und nicht auf subjektive Einschätzungen der Evaluatorin. Die Schlussfolgerungen der Evaluation werden explizit begründet und zusammen mit den Annahmen und den eingesetzten Evaluationsverfahren transparent dargestellt.

Der Standard *Metaevaluation* (G9) war anwendbar und wurde eingehalten. Die systematische Metaevaluation wurde als Selbstevaluation durch die Evaluatorin retrospektiv am Evaluationsprozess durchgeführt. Die adäquate Dokumentation und Berichterstattung führten zu einem wissenschaftlichen Fortschritt und einem Zuwachs an Wissen in dem Themenbereich. Durch die wissenschaftliche Publikation dieser Arbeit werden Fortschritte in der Evaluationsforschung erreicht und die Qualität der Evaluation gesichert.

An diesen Ergebnisteil schließt sich die Diskussion der Ergebnisse an. Es erfolgt eine Bewertung der erzielten Ergebnisse im Hinblick auf die Forschungsfrage und die Einordnung dieser in den theoretischen Kontext der Arbeit. Des Weiteren erfolgen eine kritische Auseinandersetzung mit der eigenen Vorgehensweise sowie das Aufzeigen möglicher Verbesserungen und Vorschläge für weitere Forschungen.

Tabelle 19: Metaevaluation anhand der Checkliste zur Anwendung der Standards (DeGEval, 2017a).

Der Standard...		
...war	wurde eingehalten	
anwendbar	wurde teilweise eingehalten	
und	wurde nicht eingehalten	
	war nicht anwendbar	
Nützlichkeit		
N1	Identifizierung der Beteiligten und Betroffenen	
N2	Klärung der Evaluationszwecke	
N3	Kompetenz u. Glaubwürdigkeit des Evaluators/der Evaluatorin	
N4	Auswahl und Umfang der Informationen	
N5	Transparenz von Werthaltungen	
N6	Vollständigkeit und Klarheit der Berichterstattung	
N7	Rechtzeitigkeit der Evaluation	
N8	Nutzung und Nutzen der Evaluation	
Durchführbarkeit		
D1	Angemessene Verfahren	
D2	Diplomatisches Vorgehen	
D3	Effizienz von Evaluation	
Fairness		
F1	Formale Vereinbarungen	
F2	Schutz individueller Rechte	
F3	Umfassende und faire Überprüfung	
F4	Unparteiische Durchführung und Berichterstattung	
F5	Offenlegung von Ergebnisse und Berichten	
Genauigkeit		
G1	Beschreibung des Evaluationsgegenstandes	
G2	Kontextanalyse	
G3	Beschreibung von Zwecken und Vorgehen	
G4	Angabe von Informationsquellen	
G5	Valide und reliable Informationen	
G6	Systematische Fehlerprüfung	
G7	Angemessene Analyse qualitativer und quantitativer Informationen	
G8	Begründete Bewertungen und Schlussfolgerungen	
G9	Meta-Evaluation	

5 Abschlussdiskussion und Ausblick

Die vorliegende Arbeit liefert einen umfassenden Einblick in die Thematik der Nachwuchsleistungszentren im deutschen Profifußball, diskutiert verschiedene Ansätze zur Beurteilung des (betriebswirtschaftlichen) Erfolgs, entwickelt ein effektives und effizientes Evaluationskonzept, wendet dieses an und ermittelt anhand öffentlich verfügbarer Daten ein Ranking der Nachwuchsleistungszentren Deutschlands gemäß ihrem ROI. Im Folgenden wird die Arbeit zusammengefasst, Limitationen der Untersuchung, sowie Implikationen für die Evaluation von Personalentwicklungsmaßnahmen und für die Evaluation von Nachwuchsleistungszentren aufgezeigt. Im Anschluss daran werden Implikationen für und ein Ausblick auf zukünftige Forschungen beschrieben.

5.1 Zusammenfassung

Die vorliegende Arbeit hat ihr Ziel, die Entwicklung eines Evaluationskonzepts, das

- outcome- statt input-orientiert ist,
- Resultate im Zeitverlauf über 11 Jahre erhebt statt lediglich zu einem Zeitpunkt direkt nach der Maßnahme,
- öffentlich zugängliche objektive Daten nutzt statt Selbsteinschätzungen,
- Individual-Leistungs-Indikatoren berücksichtigt statt gesamtorganisatorische Erfolgskennzahlen,
- mit geringem Aufwand leicht anzuwenden ist,
- mittels geeigneter betriebswirtschaftlicher Kennzahlen statt nicht monetärer Indikatoren

den Erfolg von Nachwuchsleistungszentren im deutschen Profifußball bewerten kann, vollumfänglich erreicht.

Das verwendete mentale Modell zeigt die Wirkung der Aktivitäten von Nachwuchsleistungszentren anhand der *Hervorgebrachten Spielerqualität* sowie unter dieser Perspektive den Erfolg oder Misserfolg eines Nachwuchsleistungszentrums nach der monetären Bewertung in Relation zu den Kosten anhand der betriebswirtschaftlichen Kennzahl Return on Investment (ROI) an. Es werden Resultate im Zeitverlauf und über den gesamten Beobachtungszeitraum von 11 Jahren betrachtet. Die Resultate sind auf Basis öffentlich zugänglicher Daten und der Entwicklung eines geeigneten Konstrukts zur Berücksichtigung der individuellen Leistung eines Spielers (*Individuelle Spielerqualität*) ermittelt worden. Der Aufwand und die Komplexität der entwickelten Konstrukte ist transparent und mittels des einfachen Wirkmodells leicht nachvollziehbar.

Die vorliegende Arbeit liefert einen umfassenden Einblick in die Thematik der Nachwuchsleistungszentren im deutschen Profifußball, diskutiert verschiedene Ansätze zur Beurteilung des betriebswirtschaftlichen Erfolgs und entwickelt ein Evaluationskonzept für interessierte Stakeholder. Zur Prüfung der Validität der Resultate ordnet die Arbeit die Ergebnisse in einen theoretischen Kontext ein und prüft kausale Zusammenhänge.

Im internationalen Profifußball beobachtet man seit der Aufhebung der Ausländerbeschränkung im Jahr 1995 eine steigende Mobilität der Spieler. Die geänderten gesetzlichen Rahmenbedingungen haben zur Folge, dass sich die besten Spieler in den finanzstärksten Vereinen konzentrieren. Auch Vereine der Big-5-Ligen stehen durch diese Entwicklung vor der Herausforderung, ihre finanziellen Mittel möglichst effizient einzusetzen, um langfristig sowohl sportlich als auch finanziell erfolgreich zu sein. Wird die Nachwuchsförderung vernachlässigt, steigt die Gefahr, sich durch Investitionen beim Kauf von für einen Erfolg benötigten Spielern zu verschulden. Mit der Nachwuchsausbildung im deutschen Profifußball hat die Arbeit ein Thema aufgegriffen, das nicht nur relevant für die Praxis erscheint, sondern auch in der Erfolgsfaktorenforschung aus einer ökonomischen Perspektive Fragen über die Effizienz der eingesetzten Mittel in die Nachwuchsleistungszentren hervorruft.

Eine länderübergreifende Studie zur Untersuchung der Nachwuchsleistungszentren aller Vereine der Big-5-Ligen (England, Frankreich, Italien, Spanien und Deutschland) war aus zeitlichen und finanziellen Gründen im Rahmen der vorliegenden Dissertation nicht möglich. Die Reduzierung auf eine tiefgehende Analyse eines einzelnen Landes hat sich aufgrund der Komplexität des Themas und der Innovativität des Evaluationsansatzes angeboten. Der deutsche Profifußball kann jedoch nicht isoliert betrachtet werden, es hat daher eine tiefgehende Auseinandersetzung mit den erklärungsrelevanten nationalen sowie internationalen Rahmenbedingungen und Einflussfaktoren auf die Nachwuchsförderung stattgefunden.

Um den Erfolg der Nachwuchsleistungszentren in Deutschland zu überprüfen, setzte sich die Arbeit zunächst mit der Frage nach der Beurteilung des Erfolgs auseinander. Es wurde deutlich, dass sich bereits bestehende Indikatoren dazu eignen, den Erfolg der Nachwuchsleistungszentren zu beurteilen (FIFA-Weltrangliste, Anzahl an BIG-5-Spielern). Auf Ebene der Vereine zeigte sich u. a. aufgrund unterschiedlicher Ausbildungsorte und -zeiten der Spieler Anpassungsbedarf bei der Beurteilung des Erfolgs. Die im Rahmen der Untersuchung entwickelte Methode gibt keine detaillierte Beurteilung der Qualität des Ausbildungsprogramms oder Strukturen im NLZ wieder, jedoch war es über die Entwicklung des Indikators *Hervorgebrachte Spielerqualität* und der monetären Bewertung mittels des mittleren Marktwertes des Spielers im Betrachtungszeitraum möglich, Nachwuchsleistungszentren in Deutschland in eine klassifizierende Reihenfolge zu bringen. Eine Unterscheidung der erfolgreichen von weniger erfolgreichen Vereinen wurde ermöglicht.

Die Anwendung der Erfolgsfaktorenforschung im Bereich Investitionen erwies sich als geeignet, um die Ausbildung in Nachwuchsleistungszentren aus einer ökonomischen Perspektive zu betrachten. Über die Berechnung des ROI war es möglich, die Effizienz der Nachwuchsleistungszentren zu beurteilen und damit einen Beitrag zur betriebswirtschaftlichen Bewertung der NLZ zu leisten. Die Erkenntnisse können sowohl für Vereine in Deutschland als auch für Vereine in anderen nationalen Fußballligen praxisrelevant sein und damit einen wichtigen Beitrag zur Weiterentwicklung des Sports leisten.

5.2 Limitationen

Die vorliegende Untersuchung weist aufgrund der zur Verfügung stehenden Daten einige Einschränkungen sowie methodische Verbesserungspotentiale auf. Im Folgenden wird konkret auf einzelne Einschränkungen der Untersuchung eingegangen.

Erfolgsfaktorenforschung

Häufig wird die Erfolgsfaktorenforschung dahingehend kritisiert, dass eine Vielzahl miteinander in Verbindung stehender interner und externer Variablen, den Unternehmenserfolg bestimmen und die Erfolgswirksamkeit der einzelnen Variable nicht isoliert werden kann. Um diese Komplexität zu reduzieren, erfolgte die Berechnung des ROI der Nachwuchsleistungszentren auf Basis einiger weniger Variablen (Investition, Liga, Punktzahl Verein, Punktzahl Deutscher Meister der letzten 5 Jahre, Einsätze Spieler, max. Einsätze, Dauer / Verbleib im Nachwuchsleistungszentrum, Marktwert) und leicht anzuwendenden Konstrukten (Vereinsqualität, Spielerqualität individuell, Hervorgebrachte Spielerqualität). Ein Versuch, die Erfolgswirksamkeit der einzelnen Variablen zu isolieren, erfolgte nicht. Der ROI ist lediglich als ein Erfolgsfaktor für die Messung des Unternehmenserfolgs ermittelt worden und reiht sich damit in zahlreiche Studien ein, die allgemein dafür kritisiert wurden (Abfalter, 2020; Haenecke, 2002), dass es nicht möglich sei, eindeutige Erfolgsfaktoren zu identifizieren.

Ein *Key Informant Bias* konnte überwiegend durch die datenbasierte Konstruktion der Erfolgsfaktoren ausgeschlossen werden. Bei der Erhebung des Investments durch die Expertenbefragung nach den Budgets der Nachwuchsleistungszentren in 2019 konnte eine Verzerrung nicht ausgeschlossen werden, aber durch die Validierung mittels der Interraterreliabilität eingeordnet werden. In Zukunft vermeiden ließe sich dieser Key Informant Bias, indem die Datenerhebung auf der Basis einer Finanzdatenbank beim DFB oder der DFL erfolgt.

Ermittlung der Investitionen in ein Nachwuchsleistungszentrum

Zur Ermittlung der Höhe der Investitionen in den Nachwuchsbereich pro Saison und Nachwuchsleistungszentrum konnte nicht auf exakte Zahlen der Vereine zugegriffen werden. Unter der Annahme, dass die Investition einen prozentualen Anteil am Rohergebnis des Vereins darstellt, sind die eingehenden Daten zwar validiert worden, können aber im Ergebnis lediglich eine Annäherung beziehungsweise eine Schätzung der wahren Investitionen darstellen. Bei einer zukünftigen Anwendung des Evaluationskonzepts sollten für einen hinreichend genauen ROI Ist-Daten aus dem Controlling der Vereine verwendet werden.

Beurteilung der Vereinsqualität

Ein Faktor der Ermittlung des Indikators *Individuelle Spielerqualität* ist die konstruierte *Vereinsqualität*. Die Formel setzt lediglich die Vereinsqualitäten innerhalb der deutschen Fußballligen in Relation, dies schränkt die Vergleichbarkeit im internationalen Kontext ein und reduziert die Spannweite der Vereinsqualitätswerte. Die Ermittlung einer internationalen Vereinsqualität könnte hier Abhilfe schaffen.

Beurteilung der Individuellen Spielerqualität

Der ROI wurde als Indikator für die Beurteilung des Erfolgs eines Nachwuchsleistungszentrums im Vergleich mit anderen Nachwuchsleistungszentren in Deutschland herangezogen. In die Berechnung geht unter anderem die individuell erreichte Spielerqualität des in einem deutschen Nachwuchsleistungszentrum ausgebildeten Spielers ein, wobei die Nationalität des Spielers keine Rolle spielt. Diese Perspektive auf das Nachwuchsleistungszentrum lässt daher keine Aussage darüber zu, inwieweit explizit für die deutsche Nationalmannschaft Fußballspieler von hoher Qualität ausgebildet werden konnten.

Beurteilung der hervorgebrachten Spielerqualität

Die hervorgebrachte Spielerqualität ist retrospektiv für den Betrachtungsraum 2009-2019 ermittelt worden. Bei der Betrachtung der Daten entsteht ein Time-lag-Problem in Bezug auf die kurz zurückliegenden Jahre aufgrund der verzögerten Entwicklung der Qualität von Spielern, die noch am Beginn der Karriere stehen. Um diese Problematik einzugrenzen, sollten bei diesen Spielern die Daten rückwirkend für mindestens weitere 5 Jahre (bis 2024) erhoben oder aber ein Teil der Spieler unter einer definierten Altersgrenze von der Analyse ausgeschlossen werden.

Monetäre Bewertung der hervorgebrachten Spielerqualität

Die hervorgebrachte Spielerqualität ist mit einem angenommenen durchschnittlichen Marktwert pro Saison monetarisiert worden. Auf eine exakte Berechnung einer Ausbildungsentschädigung oder des Solidaritätsbeitrags ist aufgrund der anzulegenden Erfolgskriterien für das durch diese Arbeit entwickelte Evaluationskonzept – öffentlich zugängliche Daten, mit geringem Aufwand leicht anzuwenden – bewusst verzichtet worden.

Der damit ermittelte Nutzen stellt, wie die oben erwähnten Kosten, eine Annäherung beziehungsweise eine grobe Schätzung dar. Bei einer zukünftigen Anwendung des Evaluationskonzepts sollten für einen hinreichend genauen ROI, Ist-Daten aus dem Controlling der Vereine verwendet werden.

Beurteilung des betriebswirtschaftlichen Erfolgs der Nachwuchsleistungszentren

Die Kritik an zahlreichen Studien der Erfolgsfaktorenforschung hat gezeigt, dass es nicht möglich ist, eindeutige Erfolgsfaktoren zu identifizieren. Dennoch können bei komplexen Evaluationsgegenständen Einflussfaktoren bzw. potentielle Erfolgsfaktoren über kausale Zusammenhänge aufgedeckt und analysiert werden. Die sportökonomische Forschung, die aktuelle Entwicklung im Profifußball sowie die Rahmenbedingungen der Nachwuchsförderung stellen die externen Bedingungen des Handelns dar und beeinflussen die Ziele und Handlungsstrategien der einzelnen Stakeholder im Profifußball in Deutschland. Aus diesem Grund wurden die Rahmenbedingungen bereits zu Beginn der Untersuchung umfassend analysiert und mögliche Zusammenhänge aufgezeigt. Die präzise Beschreibung erhöht die Genauigkeit der Analyse und ermöglicht Aussagen zur wissenschaftlichen Korrektheit und Relevanz für die Praxis. In dieser Untersuchung wurde als Faktor für die Messung des Erfolgs eines Nachwuchsleistungszentrums lediglich die Investitionskennzahl *Return on Investment (ROI)* berücksichtigt. Eine Ermittlung weiterer Kennzahlen, die unternehmerische Entscheidungen vorbereiten oder im Nachhinein bewerten, könnten weitergehenden Analysen unterstützen – z. B. den *Return on Development-Investment (RODI)*.

Neben den wirtschaftlichen Wirkungen sind es aber vor allem gesellschaftliche Werte, die die hohen Investitionen in den Nachwuchssport rechtfertigen. Sport hat laut der Bundesregierung Deutschland (Bundestag, 2019) einen hohen Stellenwert und stellt „einen zentralen Bestandteil unseres gesellschaftlichen Zusammenlebens dar. Dies gilt sowohl für den Spitzensport als auch den Breitensport“ (Bundestag, 2019, S. 15). Erfolgreiche Sportler*innen haben häufig Vorbildfunktion und stehen für Leistungswillen, Ausdauer, Fairness und Teamgeist. „Dem Sport kommen damit diverse gesellschaftliche Funktionen zu, die in ihrer Bedeutung kaum hoch genug eingeschätzt werden können“ (Bundestag, 2019, S. 15). Eine Funktion von Nachwuchsleistungszentren im deutschen Profifußball wird daher insbesondere in der Erfüllung ihrer sozialen und erzieherischen Rolle gesehen. Einige Lizenzmannschaften legen einen hohen Wert auf das sozialerzie-

herische Konzept ihrer Nachwuchsspieler. So z. B. der 1. FC Kaiserslautern: „Die Erfüllung der sozialen und erzieherischen Rolle des NLZ kommt nicht nur dem Verein, sondern der Gesellschaft als Ganzes zu Gute“ (1. FC Kaiserslautern 2015). Damit wird die Wichtigkeit der Nachwuchsausbildung nicht nur im sportlichen Bereich, sondern auch hinsichtlich ihres Einflusses auf das soziale Verhalten unterstrichen.

5.3 Implikationen für die Evaluation von Personalentwicklungsmaßnahmen

In Zeiten knapper Budgets und Ressourcen sind Personalverantwortliche und Personalentwickler in der Verantwortung, den Nutzen von Personalentwicklungsmaßnahmen nachzuweisen und gegebenenfalls zu optimieren. Personalentwicklungsmaßnahmen sind zu evaluieren, also hinsichtlich ihres Erfolgs zu bewerten, um basierend auf den Ergebnissen, Entscheidungen für eine langfristige Steuerung der Investitionen in die Weiterbildung treffen zu können (Kauffeld, 2016). Aktuell und in Zukunft wird die Evaluation und das Controlling von Personalentwicklungsmaßnahmen eine Aufgabe sein, die durch Personalentwicklungs- oder Controllingabteilungen durchgeführt werden müssen, um die Wertschaffung der Personalentwicklungsmaßnahmen sichtbar zu machen und nachzuweisen. Die Evaluation von Personalentwicklungsmaßnahmen sollte als Teil der strategischen Entscheidungen in der Personalpolitik stärker ins Bewusstsein rücken. Mit dem in der vorliegenden Arbeit erarbeiteten und in der Praxis angewandten Evaluationskonzept können Vorurteile gegenüber den Methoden abgebaut, das Wissen über die Durchführung von Evaluationen im Unternehmen erweitert und Organisationen bei der Implementierung unterstützt werden.

In der Regel werden zur Erfolgskontrolle in Anlehnung an das Vier-Ebenen-Modell von Kirkpatrick (1967, 1994) im Bereich des Lernfelds, Zufriedenheit und Lernerfolg sowie im Bereich des Transferfelds, Transfererfolg und Unternehmenserfolg, gemessen (Kauffeld, 2016). Der Unternehmenserfolg wird in der Regel lediglich von 7% der Organisationen gemessen (van Buren & Erskine, 2002). Das Fehlen geeigneter Indikatoren bzw. Maße mit denen der Erfolg auf gesamtorganisatorischer Ebene mit einem ökonomisch vertretbaren Aufwand gemessen werden kann, gilt als einer der Hauptgründe diese geringe Anzahl an Evaluationen. Das in der vorliegenden Arbeit erarbeitete und in der Praxis angewandte Evaluationskonzept kann Unternehmen – nicht nur der Fußballbranche – aufzeigen, wie geeignete Indikatoren entwickelt und validiert werden können. Darüber hinaus wird aufgezeigt, wie ein standardisiertes Instrumentarium unabhängig von der Selbsteinschätzung der Teilnehmer, auf Basis von intern bereitstehenden Daten im Unternehmen oder extern vorliegenden über das Unternehmen, leicht entwickelt werden kann.

Die detaillierte Darstellung des mentalen Modells, der Konstruktion der Indikatoren und der Metaevaluation, kann helfen, eine spezifische in den organisationalen Kontext eingebettete Evaluationsmethode zu entwickeln und zu demonstrieren wie sie zur Erreichung einen Beitrag leisten kann.

Personalentwicklungsmaßnahmen werden sich in der Zukunft insbesondere auch in ihrer Durchführungsform und bezüglich ihrer organisationalen und unternehmensbezogenen Implementierung verändern. Die Digitalisierung wird mehr und mehr Einzug halten. Auch im Bereich der Evaluation geht der Trend von manuellen Happy-Sheets-Auswertungen zu einer digitalisierten, softwareunterstützten, datenbasierten Messung des Unternehmenserfolgs der Personalentwicklungsmaßnahmen. Das in dieser Studie erarbeitete Evaluationskonzept ist ein Anwendungsfall in der Disziplin Data Science in Human Resource: Öffentlich verfügbare Daten wurden in eine Datenbank importiert, aufbereitet, bereinigt, in einer konsistenten Form gespeichert, transformiert, modelliert und visualisiert sowie kommuniziert. Den Rahmen bildete die Programmierung mit der Programmiersprache „R“ (Wickham & Grolemund, 2016).

Darüber hinaus zeigt das erarbeitete Evaluationskonzept nicht nur neue technische Möglichkeiten, Daten zu erheben und auszuwerten, sondern auch die Möglichkeit der Verknüpfung von unternehmensinternen Daten und einer Vielfalt von frei verfügbaren, öffentlichen Daten. Viele Unternehmensbereiche nutzen bereits intensiv Big-Data-Ansätze um bessere Entscheidungen bei ihren strategischen Herausforderungen treffen zu können (zum Begriff *Big-Data* siehe Schermann et al., 2014). People-Analytics-Ansätze im Personalmanagement und in der Personalentwicklung sind eher selten (zu People-Analytics siehe Tursunbayeva et al., 2018). „People Analytics bezeichnen jede Form der Datenanalyse, die in Zusammenhang mit Menschen in ihrer Arbeitsumgebung steht“ (Reindl & Krügl, 2017, S. 17). Die vorgestellte Studie stellt eine weitere praktische Anwendung im Gebiet von People Analytics, der Nutzung von Big-Data, dar.

5.4 Implikationen für die Evaluation von Nachwuchsleistungszentren

Ergebnis dieser Arbeit ist ein in der Praxis einsetzbares Evaluationskonzept, das outcome-orientiert statt input-orientiert ist, die Resultate im Zeitverlauf erhebt statt zu einem Zeitpunkt direkt nach der Durchführung der Maßnahme, öffentlich zugängliche objektive Daten nutzt statt Selbsteinschätzungen, auf Individual-Leistungs-Indikatoren basierende Maßzahlen verwendet, um den Erfolg von Nachwuchsleistungszentrum im deutschen Profifußball zu bewerten – das EvalROINLZ. Es ist das aktuell einzige Konzept, das für Nachwuchsleistungszentren den betriebswirtschaftlichen Erfolg der Personalentwicklung – von Talenten zu Profifußballspielern – anhand einer Investitionskennzahl, dem Return on Investment (ROI), evaluiert.

Diese Methode fördert die Evaluation von Nachwuchsleistungszentren an einer Vielzahl von Stellen fundamental (inspiriert von V. Stein, 2006):

- Vor der EvalROINLZ war die Evaluation von Nachwuchsleistungszentren über den DFB/die DFL bzw. durch die Agentur Double PASS input-orientiert. Mit der Methode wurden die vorherrschenden Bedingungen in einem Nachwuchsleistungszentrum für acht Qualitätsdimensionen der Zertifizierung mit jeweils mehreren Subdimensionen anhand der darin enthaltenen Kriterien (Kroemer, 2015) bewertet. Mit der EvalROINLZ erfolgt die Bewertung outcome-orientiert, da die Wirksamkeit der Personalentwicklungsmaßnahmen an der Summe des durch die Arbeit des Nachwuchsleistungszentrums generierten Nutzens gemessen wird.
- Vor der EvalROINLZ wurde in vier Zyklen, mit einem Abstand von circa drei Jahren, eine Momentaufnahme der vorherrschenden Bedingungen evaluiert. Mit der EvalROINLZ gibt es eine Betrachtung der Wirksamkeit über einen längeren Zeitraum hinweg, die die Wertentwicklung der Spieler über die Jahre insgesamt, unabhängig von den vorherrschenden Bedingungen, berücksichtigt.
- Vor der EvalROINLZ gab es zur Bewertung der Nachwuchsleistungszentren einen Qualitäts-Score für die vorherrschenden Bedingungen des Nachwuchsleistungszentrums, der in der Verleihung von Sternen als Auszeichnung mündete. Mit der Eval-ROINLZ ist erstmals eine monetäre Bewertung der Leistung von Nachwuchsleistungszentren im deutschen Fußball möglich.
- Vor der EvalROINLZ wurde der Outcome eines Nachwuchsleistungszentrums lediglich an der Anzahl der Spieler gemessen, die als Fußballspieler in einer der lizenzierten Profiligen der Big-5-Länder mindestens ein Spiel absolvieren konnten. Eine Aussage über die Qualität der Arbeit im Nachwuchsleistungszentrum war so nicht möglich. Mit der EvalROINLZ kommen als Bewertungskomponente die Qualität der hervorgebrachten Spieler und der Anteil des Nachwuchsleistungszentrums an deren Qualität hinzu.
- Vor der EvalROINLZ ging das Nachwuchsleistungszentrum bei einer betriebswirtschaftlichen Bewertung der Gesamtorganisation *Verein* (z. B: in die GuV) lediglich als Kostenfaktor ein. Mit der EvalROINLZ kann für das Nachwuchsleistungszentrum eine eigene Performancemessung durchgeführt werden. Ein Nachwuchsleistungszentrum könnte im Controlling als Profitcenter geführt und es könnte von unternehmerischen Finanziers betriebswirtschaftlich eingeschätzt werden.
- Vor der EvalROINLZ basierte die Bewertung auf einer eher subjektiven Selbsteinschätzung durch Angestellte des Vereins sowie externe Auditoren einer Agentur und wurde somit auf der Basis interpretationsbedürftiger Kriterien vorgenommen. Mit der EvalROINLZ erfolgt die Bewertung nun formel- und faktenbasiert mit öffentlich verfügbaren validen Daten und damit weitgehend manipulationsresistent.

- Vor der EvalROINLZ wurde die Bewertung für das Nachwuchsleistungszentrum pauschal gesamtorganisatorisch vorgenommen, mit der EvalROINLZ können zunächst Indikatoren entwickelt werden, die die individuelle Qualität auf Basis der Leistungsdaten eines Spielers abbilden, um dann unter Berücksichtigung der individuellen Dauer im entsprechenden Nachwuchsleistungszentrum die Gesamtwirkung der Ausbildung darzustellen.
- Vor der EvalROINLZ wurden bis zu 1000 Seiten Dokumentation erstellt, abgefragte Daten standen erst nach umfangreichen Recherchen zur Verfügung, zwei bis drei Auditoren erschienen vor Ort zur Beobachtung und für Interviews und erstellten ebenfalls im Anschluss eine weitere umfangreiche Dokumentation und Bewertung. Mit der EvalROINLZ ist die Bewertung nun mit einem geringen Aufwand möglich: Die Daten stehen öffentlich verfügbar im Internet (www.transfermarkt.de); der Programm-Code für die Ermittlung der hervorgebrachten Spielerqualität ist in einer in der Praxis weit verbreiteten Sprache (R) erstellt; die verwendeten Finanzdaten liegen für die allgemeinen Berichtspflichten (Bilanz, GuV) in den Vereinen per se vor.
- Vor der EvalROINLZ war für Vereine die Erhebung von konkreten Zahlen bezüglich ihres Nachwuchsleistungszentrums schwierig, in der Regel wurde der Fokus auf Training und Infrastruktur gelegt (Rehm-Engel, 2019). Mit EvalROINLZ kann über die Berücksichtigung der zur Verfügung stehenden Budgets bzw. der Kosten und der Messung der Wirksamkeit der Maßnahmen (Hervorgebrachte Spielerqualität – Nutzen) eine einfache Betrachtung der Kosten-Nutzen-Relation, und damit der Personalentwicklungsmaßnahme, erfolgen.

Durch die Bewertung der Nachwuchsleistungszentren mit dem Return on Investment (ROI) ist es erstmals gelungen, den wirtschaftlichen Erfolg eines Nachwuchsleistungszentrums zu bewerten. Der ROI gilt in Managerkreisen als ein Key-Performance-Indikator (KPI), der für strategische Entscheidungen unersetzlich ist. Die Berechnung des ROI repräsentiert die Ergebnisorientierung des Managements und liefert Erkenntnisse darüber, wie die Wirksamkeit des Programms in monetären Geldeinheiten einzuschätzen ist. Der ROI des eigenen Nachwuchsleistungszentrums oder des Mitbewerbers kann als Benchmark dienen, um das Nachwuchsleistungszentrum weiterzuentwickeln. Der ROI kann zur periodischen Einschätzung erzielter Fortschritte genutzt werden und als nützliche Argumentationshilfe bei vereinsinternen Verhandlungen um Budgetfreigaben oder Verhandlungen mit externen Sponsoren bzw. potentiellen Investoren eingesetzt werden. Schlussendlich kommt der Messung des Erfolgs mithilfe der Hervorgebrachten Spielerqualität und dem ROI große Bedeutung bei der langfristigen strategischen Kaderplanung zu. Aus wirtschaftlichen Gesichtspunkten heraus können damit Make-or-Buy-Entscheidungen – d. h., ob Spieler aus der eigenen erfolgreichen Nachwuchsarbeit eine gute Qualität in die Mannschaften einbringen können oder zugekauft werden müssen – fakten- und ergebnisorientiert unterstützt werden.

Mit Anwendung dieser Evaluationsmethode sind interessierte Gruppen, wie der Deutsche Fußballbund (DFB), die Deutsche Fußball Liga e. V. (DFL), die United European Football Association (UEFA), potentielle Sponsoren, Spielerberater, Spieler, Vereine, Leiter*innen von Nachwuchsleistungszentren und Eltern von talentierten ambitionierten Talenten, in der Lage, die Programme in den Nachwuchsleistungszentren auf ihre Wirksamkeit hin zu bewerten.

Für die genannten interessierten Gruppen lassen sich diverse Handlungsempfehlungen ableiten:

Leistungsdiagnostik im Fußball ist nicht mehr wegzudenken. Hier geht es insbesondere darum die Übereinstimmung zwischen Wissen, Fertigkeiten und Kompetenzen einer Person in Erfahrung zu bringen. Ziel von Diagnostik ist es, Antworten auf zwei Fragestellungen zu finden: Passt der Spieler zu den vorherrschenden Bedingungen? Wie kann der Spieler entwickelt werden, um eine Passung zu den Bedingungen herzustellen. Im Allgemeinen umfasst die Diagnostik der fußballerischen Leistung die Messung vielfacher Faktoren: „Technik, Taktik, Physiologie, Anatomie, Konstitution, psychische Faktoren und auch soziometrische Faktoren“ (Mayer & Lanwehr, 2018, S. 5). Sie findet eher begleitend zum Entwicklungsprozess statt (Mayer & Lanwehr, 2018). Der Indikator *Individuelle Spielerqualität* berücksichtigt die Einsatzzeiten und die Qualität des Vereins, in dem der Spieler spielt, als Variablen für die Kalkulation des Wertes und bietet sich für eine Anwendung durch **Spieler, Spielerberater und Scouts (Recruiter)** an, um zu ausgewählten Zeitpunkten der Profifußballkarriere die Wirksamkeit der Entwicklung des Spielers empirisch nachzuweisen und eine evidenzbasierte gesamthafte Einschätzung der Leistung auf Basis objektiver Daten zu gewinnen. Die Einschätzung der Qualität des Spielers für den Zweck eines Recruitings oder Transfers basiert dann im Vergleich zu anderen Spielern, die in Frage kommen, nicht nur auf Bauchgefühl und Beobachtung der technischen Fähigkeiten, sondern kann mittels (öffentlich verfügbaren) Leistungsdaten erfolgen.

Für den **DFB** und die **DFL** bietet sich insbesondere die, mit geringem Aufwand verbundene, Kalkulation der *Hervorgebrachten Spielerqualität* zur Bewertung des Erfolgs der Nachwuchsarbeit als Bewertungskriterium aller Nachwuchsleistungszentren an. Es wird empfohlen, mithilfe (öffentlich) verfügbarer Daten (z.B. von transfermarkt.de oder anderen Anbietern von Leistungsdaten) und den in dieser Arbeit aufgestellten Formeln (siehe Seitenangaben Formelverzeichnis) eine Bewertung der Nachwuchsleistungszentren durch die Data Science-Abteilung vorzunehmen. Die datengetriebene Evaluation macht die Einbeziehung einer Agentur (wie bisher Double Pass) mit Besuchen der Nachwuchsleistungszentren und Vereine durch Auditoren obsolet. Auf Basis der Bewertung mittels der *Hervorgebrachten Spielerqualität* je Nachwuchsleistungszentrum kann sich eine tiefergehende Analyse anschließen, in der das *Warum?* geklärt werden kann, um darauf aufbauend wiederum Gemeinsamkeiten und Unterschiede zu identifizieren, die in Handlungsempfehlungen für die Kompetenzen von Nachwuchsleistungszentren ab-

zuleiten. Die Digitalisierung von Evaluationen im Personalentwicklungs- und Ausbildungsbereich sind mittlerweile Standard, sodass an dem Einsatz dieser Methode kein Weg vorbeiführt. Die Methode stellt eine Professionalisierung der Bewertung der Wirksamkeit der Nachwuchsarbeit dar.

Dem **DFB** und der **DFL** sowie den **Vereinen** wird empfohlen, den nächsten logischen Schritt im Modell zu gehen: die *Hervorgebrachte Spielerqualität* in Relation zu den Kosten der Nachwuchsarbeit zu setzen, den *ROI* zu berechnen. Mit der EvalROINLZ kommt eine Bewertungsmethode zur Anwendung, die einen Überblick darüber gibt, wie effektiv mit den zur Verfügung stehenden finanziellen Ressourcen umgegangen wurde. Insbesondere in den Vereinen sollte ein Umdenken in Richtung professioneller betriebswirtschaftlicher Betrachtung stattfinden. **Tatsächlichen und potentiellen Sponsoren** kann damit Rechenschaft über den Erfolg der Nachwuchsarbeit aufgezeigt werden. Die Sponsoren sollten sie in ihren Gesprächen vor Investitionsentscheidungen von den Vereinen einfordern. Die entwickelten Indikatoren sollten als Key Performance Indicator (KPI) in die Berichterstattung des Controllings, z.B. in einer Balanced Scorecard, aufgenommen werden. Diese KPIs können dazu dienen, in den **Vereinen** ein besseres Controllingverständnis zu fördern oder Controllingfunktionen in die Managementorganisation von Vereinen zu bringen. Sie ermöglichen den Vereinen langfristig rentabilitätsorientiert zu kalkulieren, statt aufgrund des Bauchgefühls der Manager Investitionsentscheidungen zu treffen (Götz, 2001).

Die **UEFA** hat 2020 Studienergebnisse zu Trainingseinrichtungen und Nachwuchsförderung in Europa veröffentlicht (UEFA, 2020a). Der Bericht „UEFA Training Facility and Youth Investment Landscape“ (UEFA, 2020b), bietet einen umfassenden Überblick zur Situation von ca. 950 Trainingsanlagen von 673 europäischen Fußballvereinen auf der Grundlage von Daten, die 2019 durch das UEFA Kompetenzzentrum erhoben wurden. Die in dieser Arbeit entwickelten Indikatoren könnten den Bericht der UEFA mit der betriebswirtschaftlichen Perspektive und der Wirksamkeitsmessung aufwerten. An ein Ranking anschließen, könnte sich eine tiefere Kausal-Analyse, die wiederum dazu beitragen könnte den „UEFA Best Practice Guide to Training Centre Construction and Management“ mitzugestalten.

Leiter*innen eines Nachwuchsleistungszentrums sollten diese Indikatoren kennen, um eine langfristige Erfolgskontrolle ihres Nachwuchsleistungszentrums zu dokumentieren. Sie sollten die Ergebnisse der *Hervorgebrachten Spielerqualität* und der Kosten-Nutzen-Relation (*ROI*) zum Anlass nehmen, sich mit ihren (direkten) Mitbewerbern zu vergleichen und weitere Analysen bezüglich der Stärken und Schwächen ihres Nachwuchsleistungszentrums durchzuführen.

Eltern von ambitionierten Talenten und deren Berater wird empfohlen diese KPIs vor den Gesprächen mit den Verantwortlichen eines Nachwuchsleistungszentrums zu erheben bzw. einzufordern. Sie stellen eine Entscheidungshilfe für die Auswahl eines Nachwuchsleistungszentrums für die Ausbildung ihres Sohnes bzw. ihres Schützlings

dar. Insbesondere die *Hervorgebrachte Spielerqualität*, aber auch der ROI, ermöglichen einen, auf dem Erfolg der Nachwuchsarbeit basierenden, Vergleich der in der engeren Auswahl stehenden verschiedenen Nachwuchsleistungszentren.

5.5 Implikationen für und Ausblick auf die zukünftige Forschung

Die Erkenntnisse aus dem vorliegenden Dissertationsprojekt bilden eine Grundlage für weitergehende Analysen über die Beurteilung des Erfolgs und der Erfolgsfaktoren der Nachwuchsleistungszentren beziehungsweise Ausbildungen im Fußball eines Landes. An diese Arbeit anknüpfende Forschung sollte sich auf der Ebene des Landes Deutschland intensiver mit den Strukturen zwischen den Vereinen, dem Deutschen Fußballbund (DFB) und der Deutschen Fußball Liga e. V. (DFL) beschäftigen.

Die Struktur der Nachwuchsleistungszentren wurde von den bis dato Evaluierenden über die Jahre als gut funktionierendes System bewertet. Mit der Analyse per ROI wird eine weitere Differenzierungsmöglichkeit geboten, die es ermöglicht, auch länderübergreifend Ausbildungsvereine zu vergleichen.

In einem internationalen Vergleich wäre es interessant, die Qualität der Ausbildung in Deutschland mit den Qualitäten der Ausbildung in den übrigen Big-5-Ländern und darüber hinaus eingehender zu betrachten, indem die nationalen Ausbildungssysteme verglichen und auch dort die zentralen Erfolgsfaktoren näher spezifiziert werden. Damit könnte die Wirksamkeit – absolut über die Hervorgebrachte Spielerqualität, in der Kosten-Nutzen-Relation und über die Zeit – der unterschiedlichen verbandseigenen Nachwuchsförderkonzepte – wie z.B. der in der englischen Premier League eingesetzte Elite Player Performance Plan (EPPP; <https://thepfsa.co.uk/de/eppp/>), das französische *Centre National du Football* (CNF; vgl. dazu Bächle, 2021), Jugendakademien in Spanien) evaluiert werden.

Um ein besseres Verständnis der Erfolgsdeterminanten von Nachwuchsarbeit im Rahmen der Erfolgsfaktorenforschung über den ROI hinaus auf der Grundlage eingehender Analysen einzelner Vereine oder Länder zu erlangen, könnte z. B. auch der Return on Development Investment für Nachwuchsleistungszentren ermittelt werden.

Den Untersuchungen dieser Arbeit könnten sich Analysen zu individuellen Karriereverläufen ausgebildeter Spieler und zur Wirksamkeit relevanter Variablen auf individuelle Karriereverläufe der Spieler anschließen – beispielhafte Fragen wären: Wie verlaufen typische Karrieren von Spielern, die im Profifußball ankommen? Gibt es Unterschiede in den Karriereverläufen in Bezug auf Positionen? Welche Auswirkungen haben Verletzungen auf den Karriereverlauf? Gibt es weitere Auffälligkeiten in den Karriereverläufen? Alle diese Fragen gelten auch für den internationalen Vergleich.

Diese Arbeit fokussierte auf die betriebswirtschaftliche Perspektive der Wirksamkeit von Nachwuchsausbildung. Untersuchungen zur Wirksamkeit der Nachwuchsausbil-

derung im deutschen Profifußball auf die Gesellschaft, könnten ein weiteres Forschungsfeld sein. Hierbei wäre ggf. das Thema Nachhaltigkeit der Nachwuchsarbeit mit Blick auf die beiden genannten Facetten – sportlich und menschlich – interessant.

Im Bereich der innovativen Forschungsmethoden in der Evaluation sind insbesondere zwei Trends auszumachen: *Weniger Pen- and Paper-Befragungen* und *mehr Partizipation und Immersion*. Weniger Pen- and Paper-Befragungen werden ermöglicht durch eine zunehmende Digitalisierung, z.B. in Form von Text-Mining, Data Analytics, Machine Learning / Algorithmen, Learning Analytics und softwarebasierte Evaluation. Mehr Partizipation und Immersion werden ermöglicht z.B. durch den Einsatz von Mixed Methods, Storytelling, concept Mapping, Remote Evaluation, Contribution Analysis. Durch die Anwendung dieser innovativen Methoden entstehen mehr implizite Daten als durch herkömmliche Evaluationsmethoden. Darüber hinaus wird eine Verlinkung von Daten in größerem Ausmaß ermöglicht, die in der Praxis solide Ergebnisse und umsetzbare Einsichten liefert (Tasliarmut, 2022). Das in dieser Arbeit entwickelte und angewendete Evaluationskonzept setzt auf diesem Trend auf und ist ein Anwendungsbeispiel, das das Potenzial automatisierter Leistungsbewertung und Wirksamkeitsmessung im Kontext der Evaluationsforschung aufzeigt. Zukünftige Forschungen könnten diesen Ansatz – z.B. in Analogie zur aktuellen Evaluationspraxis in der Marktforschung – ausbauen, indem weitere öffentlich verfügbare für die Messung des Erfolgs relevanten Daten in den Sozialen Medien identifiziert und verarbeitet werden. Darüber hinaus könnten vereinspezifische Daten in die Konstrukte einfließen.

Um die vielseitigen Chancen im Umgang mit systematisch erhobenen Daten zu heben, könnte die gezeigte Evaluationsmethode um non-reaktive digitale Erhebungsverfahren wie z.B. Tracking-Verfahren bei der Leistungsanalyse von Spielern und Data Analytics erweitert werden, um die Tiefe der Analyse zu erhöhen (Stützer et al., 2018).

In enger Kooperation mit den nationalen Fußballorganisationen (DFB und DFL) könnte eine brancheninterne Diskussion, zu Gütekriterien und Qualitätsstandards in Bezug auf den Umgang mit digitalen Methoden angestrebt werden. Perspektivisch könnte dies zu einer weiteren Standardisierung der Evaluationspraxis, zu methodischen Arbeitsabläufen und dem Umgang mit Daten sowie der Kommunikation zu neuen digitalen Methoden in der Fußballbranche führen (Stützer, 2020).

Damit Nachwuchsleistungszentren ihre Wirksamkeit realistischer einschätzen können, bedarf es seitens der Forschung praxisnaher Handlungsempfehlungen, wie eine Evaluation effektiv und effizient durchgeführt werden kann. Die bisherige Praxis der Evaluationen, durchgeführt durch eine externe Agentur beauftragt vom DFB und der DFL, führte dazu, dass Nachwuchsleistungszentren bzw. Vereine, keine umfangreichen Kompetenzen für eine aktive Qualitätssicherung und ein proaktives Controlling aufbauten. Dies führt in der Folge dazu, dass Nachwuchsleistungszentren die wirtschaftlichen Zusammenhänge bei der Ausbildung der Nachwuchsfußballspieler nicht in Betracht ziehen. Da von den Nachwuchsleistungszentren nicht erwartet werden kann, eigene Forschung in dieser Richtung zu betreiben, gilt es Schnittstellen mit der Wissenschaft zu etablieren. Hierzu müsste die Forschung auch über die bestmögliche Distribution ihrer

Wissenschaftskommunikation für die Fußballbranche reflektieren, sodass ihre Erkenntnisse flächendeckend Verbreitung finden.

Literaturverzeichnis

- AAPOR. (2021, 11. Oktober). *Response Rates - An Overview - AAPOR*. <https://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx>
- Abbott, A. & Collins, D. (2004). Eliminating the dichotomy between theory and practice in talent identification and development: considering the role of psychology. *Journal of Sports Sciences*, 22(5), 395–408.
<https://doi.org/10.1080/02640410410001675324>
- Abfalter, D. (2010). *Das Unmessbare messen?* Zugl.: Innsbruck, Univ., Diss., 2008 (1. Aufl.). *VS research* [391 S]. VS Verlag für Sozialwissenschaften.
<https://doi.org/10.1007/978-3-531-92163-1>
- Ackermann, P. & Follert, F. (2018). Einige bewertungstheoretische Anmerkungen zur Marktwertanalyse der Plattform transfer-markt.de. *SCLAMUS – Sport und Management*(3), 21–41.
- agof. (2021). *Ddf Februar 2021 Angebote Ranking Internet*. Studienarchiv daily digital facts 2021 - agof. <https://www.agof.de/studien/daily-digital-facts/studienarchiv-daily-digital/studienarchiv-daily-digital-facts-2021/#09.03.2021>
- Alagaraja, M. (2013). HRD and HRM Perspectives on Organizational Performance. *Human Resource Development Review*, 12(2), 117–143.
<https://doi.org/10.1177/1534484312450868>
- Ali, A. (2011). Measuring soccer skill performance: a review. *Scandinavian Journal of Medicine & Science in Sports*(21), 170–183.
- Alliger, G. M. & Janak, E. A. (1989). Kirkpatrick's levels of training criteria: thirty years later. *Personnel Psychology*(42), 331–342.
- Alliger, G. M., Tannenbaum, S. I., Bennett, W., Traver, H. & Shotland, A. (1997). A Meta-analysis of the relations among training criteria. *Personnel Psychology*, 50(2), 341–358.
- Apel, D., Behme, W., Eberlein, R. & Merighi, C. (2015). *Datenqualität erfolgreich steuern: Praxislösungen für Business-Intelligence-Projekte* (3. Aufl.). Edition TDWI. Dpunkt-Verl.
- Araújo, D., Fonseca, C., Davids, K., Garganta, J., Volossovitch, A., Brandão, R. & Krebs, R. (2010). The Role of Ecological Constraints on Expertise Development. *Talent Development & Excellence*, 2(2), 165–179.
- Arthur, W., Tubré, T. P., Paul, D. S. & Edens, P. S. (2003). Teaching effectiveness: The relationship between reactions and learning evaluation criteria. *Educational Technology*, 23, 275–285.
- ASTD. (2002). *The 2002 ASTD state of the industry report*. <https://www.voiced.edu.au/content/ngv:39895>
- ATD Research. (2016). *Evaluating Learning: Getting to Measurements That Matter*. ATD Press.

- Avolio, B. J., Avey, J. B. & Quisenberry, D. (2010). Estimating return on leadership development investment. *The Leadership Quarterly*, 21(4), 633–644.
<https://doi.org/10.1016/j.leaqua.2010.06.006>
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R [R.]. (2000). *Multivariate Analysemethoden-Eine anwendungsorientierte Einführung*. (9. überarbeitete und erweiterte Auflage).
- Baker, J., Schorer, J. & Cobley, S. (2010). Relative age effects. *Sportwissenschaft*, 40(1), 26–30. <https://doi.org/10.1007/s12662-009-0095-2>
- Bächle, M., (2021). Was die französische Nachwuchsförderung erfolgreich macht. Spurensuche im Wald: Wie Frankreich so viele Superstars ausbildet
<https://www.kicker.de/spurensuche-im-wald-wie-frankreich-so-viele-superstars-ausbildet-806562/artikel>
- Balzer, L. (2005). *Wie werden Evaluationsprojekte erfolgreich? – Ein integrierender theoretischer Ansatz und eine empirische Studie zum Evaluationsprozess*. Empirische Pädagogik.
- Balzer, L. & Beywl, W. (2018). *Evaluiert: Erweitertes Planungsbuch für Evaluationen im Bildungsbereich* (2., überarbeitete Auflage). hep, der bildungsverlag.
<https://irf.fhnw.ch/handle/11654/26832>
- Bartscher, T. & Nissen, R. (2018). *Wirtschaftslexikon Gabler - Definition Personalentwicklung*. <https://wirtschaftslexikon.gabler.de/definition/personalentwicklung-52604/version-330100>
- Bates, R. (2004). A critical analysis of evaluation practice: the Kirkpatrick model and the principle of beneficence. *Evaluation and Program Planning*(27), 341–347.
- Batini, C. & Scannapieca, M. (2006). *Data Quality: Concepts, Methodologies and Techniques. Data-Centric Systems and Applications*. Springer Berlin Heidelberg.
<http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10145282>
<https://doi.org/10.1007/3-540-33173-5>
- Battis, U., Ingold, A. & Kuhnert, K. (2010). Zur Vereinbarkeit der „6+5“-Spielregel der FIFA mit dem Unionsrecht. *Europarecht*, 45(1), 3–29.
<https://doi.org/10.5771/0531-2485-2010-1-3>
- Battis, U., Ingold, A. & Kuhnert, K. *Zur Vereinbarkeit der „6+5“-Spielregel der FIFA mit dem Unionsrecht*. <https://www.europa-recht.nomos.de/?id=1268>
- Becker, F. G. (2005). Den Return on Development messen: Möglichkeiten und Grenzen der Evaluation. *Personalführung*(4), 48–53.
- Becker, G. S. (1993). *Ökonomische Erklärung menschlichen Verhaltens: Die Einheit der Gesellschaftswissenschaften, Bd. 32. (2. Aufl.)*. Mohr.
- Becker, M. (2013). *Personalentwicklung: Bildung Förderung und Organisationsentwicklung in Theorie und Praxis* (6., überarb. und aktualisierte Aufl.). Schäffer-Poeschel.
- Becker, R. (2017). *Lehrbuch der Bildungssoziologie*. Springer Fachmedien Wiesbaden.
<https://doi.org/10.1007/978-3-658-15272-7>
- Behnke, N [N.]. (2009). Datenarchive. In K.-U. Schnapp, N. Behnke & J. Behnke (Hrsg.), *Forschungsstand Politikwissenschaft. Datenwelten: Datenerhebungen und Datenbestände in der Politikwissenschaft* (1. Aufl., S. 392–408). Nomos.

- Bergmann, G. & Meurer, G. (2003). Zum Einstieg: Best Patterns. In G. Bergmann & G. Meurer (Hrsg.), *Best Patterns Marketing, Erfolgsmuster für Innovations-, Kommunikations- und Markenmanagement*. (S. 15–39). Luchterhand.
- Bialas, J. (2021). *Herz über Kopf bei Transfermarkt.de? – Eine Prüfung der Datenqualität von Transfermarkt.de der Major League Soccer mit anderen Daten amerikanischer Fußballstatistiken* [unveröffentlichte Hausarbeit]. Fachhochschule Südwestfalen, Meschede.
- Birgmayr, R. (2011). Eine praxisnahe Einführung in Bildungscontrolling. Das Modell von Kirkpatrick und seine Erweiterungen durch Phillips und Kellner. *Magazin erwachsenenbildung.at. Das Fachmedium für Forschung, Praxis und Diskurs*.(12). <http://www.erwachsenenbildung.at/magazin/11-12/meb11-12.pdf>
- Bledsoe, M. D. (2000). *Correlations in Kirkpatrick's training evaluation model* [Doctoral dissertation]. University of Cincinnati.
- Bloom, B. S. (1985). *Developing talent in young people* (1st ed.). Ballantine Books.
- Bomberger, D. W. (2003). *Evaluation of training in human service organizations: A qualitative case study*. (3114815) [Doctoral Dissertation]. The Pennsylvania State University.
- Bortz, J. & Lienert, G. A. (2008). *Kurzgefasste Statistik für die klinische Forschung: Leitfaden für die verteilungsfreie Analyse kleiner Stichproben ; mit 13 Abbildungen und 97 Tabellen sowie zahlreichen Formeln* (3. Aufl.). *Springer-Lehrbuch Bachelor, Master*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-540-75738-2>
- Bortz, J. & Schuster, C. (2010). Tests zur Überprüfung von Unterschiedshypothesen. In J. Bortz & C. Schuster (Hrsg.), *Springer-Lehrbuch. Statistik für Human- und Sozialwissenschaftler: Mit 163 Tabellen* (7. Aufl., S. 117–136). Springer. https://doi.org/10.1007/978-3-642-12770-0_8
- Bortz, J. & Schuster, C. (2016). *Statistik für Human- und Sozialwissenschaftler: Mit 163 Tabellen* (7. Aufl.). *Springer-Lehrbuch*. Springer.
- Bosscher, V. de [Veerle], Knop, P. de [Paul], van Bottenburg, M. & Shibli, S. (2006). A Conceptual Framework for Analysing Sports Policy Factors Leading to International Sporting Success. *European Sport Management Quarterly*, 6(2), 185–215. <https://doi.org/10.1080/16184740600955087>
- Bott, D., Hoppe, G. & Breitner, M. H. (2003). *Nutzenanalyse im Rahmen der Evaluation von E-Learning Szenarien.: IWI Discussion Paper*. (Rep. No. 6). Universität Hannover. Institut für Wirtschaftsinformatik.
- Boudreau, J. W. & Ramstad, P. M. (2002). *Strategic HRM Measurement in the 21st Century: From Justifying HR to Strategic Talent Leadership*. <https://ecommons.cornell.edu/handle/1813/77386>
- Böwing-Schmalenbrock, M. & Jurczok, A. (2012). *Multiple Imputation in der Praxis : ein sozialwissenschaftliches Anwendungsbeispiel* (urn:nbn:de:kobv:517-opus-58111) [Wissenschaftlicher Artikel, Universität Potsdam, Potsdam]. BibTeX.
- Brand, S., Emrich, E., Güllich, A., Prohl, R., Turbanski, S. (2006). Bericht über den Arbeitskreis Qualitätsentwicklung als pädagogische Dimension des (Nachwuchs-)Leistungssports. In V. Scheid (Hrsg.), *Sport und Bewegung vermitteln* (Vol.

- 165). Czwalina. https://www.researchgate.net/profile/Arne-Guellich/publication/259737674_Bericht_uber_den_Arbeitskreis_Qualitaetsentwicklung_als_paedagogische_Dimension_des_Nachwuchs-Leistungssports/links/00b4953140f3ee082b000000/Bericht-ueber-den-Arbeitskreis-Qualitaetsentwicklung-als-paedagogische-Dimension-des-Nachwuchs-Leistungssports.pdf
- Briel, F. & Lehmann, J. (2017). *TSG 1899 Hoffenheim Fußball-Spielbetriebs GmbH, Sinsheim, Jahresabschluss zum Geschäftsjahr vom 01.07.2015 bis zum 30.06.2016.: Lagebericht für das Wirtschaftsjahr 01. Juli 2015 bis 30. Juni 2016*. Rechnungslegung / Finanzberichte. <https://www.bundesanzeiger.de/pub/de/suchergebnis?20>
- Brinkerhoff, R. O. (1987). *Achieving results from training*. Jossey-Bass.
- Bryson, A., Frick, B. & Simmons, R. (2013). The Returns to Scarce Talent. *Journal of Sports Economics*, 14(6), 606–628. <https://doi.org/10.1177/1527002511435118>
- Bühl, A. (2019). *SPSS: Einführung in die moderne Datenanalyse ab SPSS 25* (16., aktualisierte Auflage). Pearson.
- Bundesliga. (2021). *Ewige Tabelle*. <https://www.bundesliga.com/de/bundesliga/ewige-tabelle>
- Bundestag, D. (2019). 14. Sportbericht der Bundesregierung: Drucksache 19/9150. <https://dserver.bundestag.de/btd/19/091/1909150.pdf>
- Burzan, N. (2014). Indikatoren. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 1029–1036). Springer VS.
- Bushnell, D. S. (1990). Input, process, output: a model for evaluating training. *Training & Development Journal*, 44(3), 41–43. <https://link.gale.com/apps/doc/A8254390/AONE?u=anon~6ed5e605&sid=googleScholar&xid=3f3093d1>
- Cachay & Riedl. (2002). *Bosman-Urteil und Nachwuchsförderung. Folgen der Abschaffung von Ausländerklauseln für die Nachwuchsrekrutierung in den Sportspielen*. http://www.bisp.de/nn_18766/shareddocs/downloads/publikationen/jahrbuch/jb__2000__artikel/cachay,templateid=raw,property=publication-file.pdf/cachay.pdf
- Cannon-Bowers, J. A., Salas, E., Tannenbaum, S. I. & Mathieu, J. E. (1995). Toward Theoretically Based Principles of Training Effectiveness: A Model and Initial Empirical Investigation. *Military Psychology*, 7(3), 141–164. https://doi.org/10.1207/s15327876mp0703_1
- Cascio, W. F., Boudreau, J. & Fink, A. A. (2019). *Investing in people: Financial impact of human resource initiatives* (3rd edition). Society for Human Resource Management.
- Chadwick, S. & Parker, G. (1996). Strategy and Success in the English Football Premier League. In C. Pigeassou & R. Ferguson (Hrsg.), *Fourth Congress of the European Association for Sport Management: Proceedings* (S. 322–342).
- Charness, G. & Sutter, M. (2012). Groups Make Better Self-Interested Decisions. *Journal of Economic Perspectives*, 26(3), 157–176. <https://doi.org/10.1257/jep.26.3.157>
- CIES. (2018a). *A comparative analysis of club-trained players in Europe*. CIES Football Observatory. <https://football-observatory.com/IMG/sites/mr/mr33/en/>.

- CIES. (2017a). *CIES Football Observatory Monthly Report n°29 - November 2017: Demographic study of European football (2009-2017)*. CIES Football Observatory. <https://football-observatory.com/IMG/sites/mr/mr29/en/>
- CIES. (2017b). *Swiss Football Study: Swiss Football League*. https://www.sfl.ch/fileadmin/user_upload/www.sfl.ch/downloads/publikationen/SL2017_v01_de.pdf
- CIES. (2020). *Scientific evaluation of the transfer value of football players*. Monthly Report 53. CIES Football Observatory.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. <http://www.utstat.toronto.edu/~brunner/oldclass/378f16/readings/Cohen-Power.pdf>
- Colquitt, J. A., LePine, J. A. & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology*(85), 678–707.
- Côté, J. (1999). The Influence of the Family in the Development of Talent in Sport. *The Sport Psychologist*, 13(4), 395–417. <https://doi.org/10.1123/tsp.13.4.395>
- Côté, J., Baker, J. & Abernethy, B. (2007). Practice and Play in the Development of Sport Exercise. In R. Eklund & G. Tenenbaum (Hrsg.), *Handbook of Sport Psychology* (S. 184–202). Wiley.
- Dalziel, M., Downward, P., Parrish, R., Pearson, G. & Semens, A. (2013, 30. April). *Study on the Assessment of UEFA's 'Home Grown Player Rule': Negotiated procedure EAC/07/2012*. <https://ec.europa.eu/assets/eac/sport/library/studies/final-rpt-april2013-homegrownplayer.pdf>
- Day, E. A., Arthur, W. & Gettman, D. (2001). Knowledge structures and the acquisition of a complex skill. *The Journal of applied psychology*, 86(5), 1022–1033. <https://doi.org/10.1037/0021-9010.86.5.1022>
- DeGEval. (2017a). *Checkliste zur Anwendung der Standards: Begleitmaterial zu den Standards für Evaluation, Gesellschaft für Evaluation e.V., erste Revision 2016*. https://www.degeval.org/fileadmin/DeGEval-Standards/2018_Checkliste_zur_Anwendung_der_Standards.pdf
- DeGEval. (2017b). *DeGEval-Standards_fuer_Evaluation*. Langfassung. https://www.degeval.org/fileadmin/Publikationen/DeGEval-Standards_fuer_Evaluation.pdf
- DeGEval. (2017c). *Standards für Evaluation: Erste Revision 2016*.
- DeGEval. (2021). *Glossar der Standards für Evaluation*. <https://www.degeval.org/degeval-standards/glossar-der-standards-fuer-evaluation/>
- Delahaye, B. L. (2000). *Human resource development: Principles and practice*. J. Wiley.
- Deutscher Sportbund. (2006). *Nachwuchsleistungssportkonzept 2012: Leitlinien zur Weiterentwicklung des Nachwuchsleistungssports*.
- DFB. (2013). *DFL/Ligaverband*. <https://www.dfb.de/bundesliga/liga-information/dfligaverband/>
- DFB. (2014a). *DFB Ausbildungskonzeption: Onlineblätterfunktion*. <https://www.dfb.de/trainer/artikel/online-blaetterfunktion-dfb-ausbildungskonzeption-651/>

- DFB. (2014b). *Leistungszentren*. <https://www.dfb.de/sportl-strukturen/talentfoerderung/leistungszentren/>
- DFB. (2018). *Verbandsstruktur*. <https://www.dfb.de/verbandsstruktur/>
- DFB. (2020). *Talentförderung - Förderung der Spitztalente*. <https://www.dfb.de/sportl-strukturen/talentfoerderung/leistungszentren/>
- DFB. (2021a). *Aktuelle Statistik: DFB-Mitgliederstatistik*. <https://www.dfb.de/verbandsstruktur/mitglieder/aktuelle-statistik/>
- DFB. (2021b). *Ewige Tabelle*. Deutscher Fußball-Bund. <https://www.dfb.de/bundesliga/statistik/ewige-tabelle/>
- DFB. (2021c). *Talentförderung*. <https://www.dfb.de/sportl-strukturen/talentfoerderung/einfuehrung/>
- DFL. (2011, 13. September). *61_Spieldaten: Medien erhalten weiterhin alle offiziellen Spieldaten der Bundesliga und 2. Bundesliga* [Press release]. Frankfurt/Main. https://media.dfl.de/sites/2/2018/11/2011_Pressemitteilung_61_Spieldaten.pdf
- DFL (2016). *DFL Wirtschaftsreport 2016*. https://media.dfl.de/sites/2/2018/11/BL_Report_2016.pdf
- DFL (2017). *DFL Wirtschaftsreport_2017*. https://media.dfl.de/sites/2/2018/11/DFL_Report_2017.pdf
- DFL (2018). *DFL Wirtschaftsreport 2018*. https://media.dfl.de/sites/2/2018/11/DFL_Report_2018_S.pdf
- DFL. (2019a). *DFL Hintergrund Leistungszentren*. <https://www.dfl.de/de/hintergrund/nachwuchsleistungszentren/wende-am-tiefpunkt-die-nachwuchsleistungszentren>
- DFL. (2019b). *DFL Wirtschaftsreport 2019*. https://www.dfl.de/de/dfl_wirtschaftsreport_2019_de_m-2/
- DFL. (2020a). *DFL Wirtschaftsreport 2020*. <https://www.dfl.de/de/aktuelles/dfl-wirtschaftsreport-2020/>
- DFL. (2020b). *Mindestens 12 deutsche Profis und Local-Player-Regelung*. <https://www.dfl.de/de/hintergrund/transferwesen/local-player-regelung/>
- DFL. (2020c). *Die Vorgaben der Leistungszentren*. <https://www.dfl.de/de/hintergrund/nachwuchsleistungszentren/vielfaeltige-aspekte-die-vorgaben-nachwuchsleistungszentren/>
- DFL (2021a). *DFL Wirtschaftsreport 2021*. https://media.dfl.de/sites/2/2021/03/D_DFL_Wirtschaftsreport_2021_S.pdf
- DFL. (2021b). *Die Entwicklung der Leistungszentren*. <https://www.dfl.de/de/hintergrund/nachwuchsleistungszentren/wende-am-tiefpunkt-die-nachwuchsleistungszentren/>
- DFL. (2021c). *Richtlinien für die Errichtung und Unterhaltung von Leistungszentren der Teilnehmer der Lizenzligen: (Anhang V zur Lizenzierungsordnung des DFL e.V.)*. <https://media.dfl.de/sites/2/2021/01/Anhang-V-zur-LO-2021-01-01-Stand-Geltung-ab-1.-Januar-2021.pdf>
- Lizenzierungsordnung des DFL Deutsche Fußball Liga e.V. (2021). <https://media.dfl.de/sites/2/2021/03/Lizenzierungsordnung-LO-2021-03-05-Stand.pdf>

- Diebolt, C., Hippe, R. & Jaoul-Grammare, M. (2017). Die traditionelle Humankapitaltheorie. In C. Diebolt, R. Hippe & M. Jaoul-Grammare (Hrsg.), *Bildungsökonomie. Eine Einführung aus historischer Perspektive.: Lehrbuch, Bd. 95* (S. 5–13). Springer Gabler. https://doi.org/10.1007/978-3-658-16147-7_2
- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Springer. <https://doi.org/10.1007/978-3-642-41089-5>
- Double Pass. (2021). *About*. <https://www.doublepass.com/about/>
- Duden. (2016). *Wirtschaft von A bis Z: Grundlagenwissen für Schule und Studium, Beruf und Alltag*.
- Durand-Bush, N. & Salmela, J. H. (2001). The development of talent in sport. In R. N. Singer, H. A. Hausenblas & C. Janelle (Hrsg.), *Handbook of Sport Psychology*. (S. 269–289). Wiley.
- Durand-Bush, N. & Salmela, J. H. (2002). The Development and Maintenance of Expert Athletic Performance: Perceptions of World and Olympic Champions. *Journal of Applied Sport Psychology*, 14(3), 154–171. <https://doi.org/10.1080/10413200290103473>
- ECA. (2012). *Report on Youth Academies in Europe*. Nyon. European Club Association. <https://www.ecaeurope.com/media/2730/eca-report-on-youth-academies.pdf>
- ECA. (2019). *Youth Academy Study: Condensed Version*. European Club Association. <https://www.ecaeurope.com/media/4546/eca-youth-academy-study-condensed-version-2019.pdf>
- Eid, M., Gollwitzer, M. & Schmitt, M. (2010). *Statistik und Forschungsmethoden* (3. Aufl.). Beltz.
- Elter, V. V. C. (2011). Bewertung von Fußballunternehmen. In J. Drukarczyk & D. Ernst (Hrsg.), *Branchenorientierte Unternehmensbewertung* (3. Aufl., 381–40). Vahlen.
- Emrich, E., Güllich, A. & Büch, M.-P. (Hrsg.). (2005). *Beiträge zum Nachwuchsleistungssport*. Hofmann.
- Ericsson, K. A. & Lehmann, A. C. (1996). Expert and exceptional performance: evidence of maximal adaptation to task constraints. *Annual review of psychology*, 47, 273–305. <https://doi.org/10.1146/annurev.psych.47.1.273>
- Eschweiler, M. & Vieth, M. (2004). Preisdeterminanten bei Spielertransfers in der Fußball-Bundesliga. *Die Betriebswirtschaft*, 64(6), 671–692. <https://www.proquest.com/scholarly-journals/preisdeterminanten-bei-spielertransfers-der/docview/208919706/se-2?accountid=203855>
- EUGH. (2010). *Ausbildungsentschädigungen: Art. 39 EG - Freizügigkeit der Arbeitnehmer - Beschränkung - Berufsfußballspieler - Verpflichtung zum Abschluss des ersten Vertrags als Berufsspieler mit dem ausbildenden Verein - Verurteilung des Spielers zu Schadensersatz wegen Verstoßes gegen diese Verpflichtung - Rechtfertigung - Zweck, die Anwerbung und Ausbildung von Nachwuchsspielern zu fördern*.
- Fahrmeir, L., Künstler, R., Pigeot, I. & Tutz, G. (2004). *Statistik: Der Weg zur Datenanalyse ; mit 25 Tabellen* (5. Aufl.). Springer-Lehrbuch. Springer.
- Fandel, G. (2005). *Produktion I. Produktions- und Kostentheorie*. Springer.

- Farrokhzad, S. & Mäder, S. (2014). *Nutzenorientierte Evaluation: Ein Leitfaden für die Arbeitsfelder Integration, Vielfalt und Toleranz* (1. Aufl.). Waxmann Verlag GmbH.
- Farrow, D. (2010). A Multi-Factorial Examination of the Development of Skill Expertise in High Performance Netball. *Talent Development & Excellence*, 2(2), 123–135.
- Faude, O., Schlumberger, A., Fritsche, T., Treff, G. & Meyer, T. (2010). Leistungsdiagnostische Testverfahren im Fußball – methodische Standards. *Deutsche Zeitschrift für Sportmedizin*(61), 129–133.
- FIFA. (2019a). *Global Transfer Market Report 2018*. <https://digitalhub.fifa.com/m/2f488488c4367b36/original/wtrsayyoponlu1tknvc-pdf.pdf>
- FIFA (2019b). Reglement: bezüglich Status und Transfer von Spielern. <https://digitalhub.fifa.com/m/2b3d72a0fa811051/original/uhcoqzltalbars1vjomk-pdf.pdf>
- FIFA. (2017). *Global Transfer Market Report 2016*. <https://digitalhub.fifa.com/m/480f312fac79d3b8/original/a1jezc9xn78nzivqpgae-pdf.pdf>
- FIFA. (2018). *Global Transfer Market Report 2017*. <https://digitalhub.fifa.com/m/480f312fac79d3b8/original/a1jezc9xn78nzivqpgae-pdf.pdf>
- FIFA. (2020). *Global Transfer Market Report 2019*. <https://digitalhub.fifa.com/m/248987d86f2b9955/original/x2wrqjstwjoiainncnod-pdf.pdf>
- FIFA. (2021). *Global Transfer Market Report 2020*. <https://digitalhub.fifa.com/m/482e6b2d76404434/original/ijiz9rtpkfnbhwqqr70-pdf.pdf>
- Fitz-enz, J. (2009). *The ROI of human capital: Measuring the economic value of employee performance* / Jac Fitz-enz (2nd ed.). AMACOM; London : McGraw-Hill [distributor].
- Fitzpatrick, J., Sanders, J. & Worthen, B. (2011). *Program evaluation. Alternative approaches and practical guidelines*. (4th Edition). Pearson.
- Fleiss, J. L. (2011). *Design and Analysis of Clinical Experiments* (1. Aufl.). *Wiley Classics Library*. John Wiley & Sons.
- Follert, F. (2018). Ökonomisierung des Fußballs. *Das Wirtschaftsstudium*, 47, 668–670.
- Franck, E. & Nüesch, S. (2012). Talent and/or popularity: what does it take to be a superstar? *Economic Inquiry*, 50(1), 202–216. <https://doi.org/10.1111/j.1465-7295.2010.00360.x>
- Frick, B. (2001). Die Einkommen von „Superstars“ und „Wasserträgern“ im professionellen Teamsport: Ökonomische Analyse und empirische Befunde. *Zeitschrift für Betriebswirtschaft*, 71, 701–720.
- Fritz, T. (2012). *Die monetäre Bewertung von Bildungsmaßnahmen als Teilaspekt des betrieblichen Bildungscontrollings* [Dissertation, Universität, Frankfurt/Main]. Deutsche Nationalbibliothek.
- Gabler Wirtschaftslexikon. (2018). *Erfolg: Definition*. "<https://wirtschaftslexikon.gabler.de/definition/erfolg-33857/version-257374>"
- Gabler Wirtschaftslexikon. (2022, 31. März). *Revision von Investitionsrechnung vom Do., 15.02.2018 - 15:01 • Definition* | *Gabler Wirtschaftslexikon*. <https://wirtschaftslexikon.gabler.de/definition/investitionsrechnung-41465/version-264829>
- Gebel, T. & Rosenbohm, S. (2017). Forschungsdatenmanagement in der Organisationsforschung. In S. Liebig, W. Matiaske & S. Rosenbohm (Hrsg.), *Handbuch*

- Empirische Organisationsforschung* (S. 157–183). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-08493-6_9
- Gellert-Beckmann, S. (2020). *Neue Begriffe in der Eingliederungshilfe. Wirkung, Wirkungskontrolle und Wirksamkeit*. <https://www.socialnet.de/materialien/29122.php>
- Gerhards, J., Mutz, M. & Wagner, G. G. (2014). Die Berechnung des Siegers: Marktwert, Ungleichheit, Diversität und Routine als Einflussfaktoren auf die Leistung professioneller Fußballteams / Predictable Winners. Market Value, Inequality, Diversity, and Routine as Predictors of Success in European Soccer Leagues. *Zeitschrift für Soziologie*, 43(3), 231–250. <https://doi.org/10.1515/zfsoz-2014-0305>
- Gessler, M. & Sebe-Opfermann, A. (2011). Der Mythos "Wirkungskette" in der Weiterbildung. Empirische Prüfung der Wirkungsannahmen im "Four Levels Evaluation Model" von Donald Kirkpatrick. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 107(2), 270–279.
- Geyer, H. & Dilger, A. (2009). *Folgen des Bosman-Urteils für die 1. Fußball-Bundesliga* (IÖB-Diskussionspapier 2/09). Münster: Westfälische Wilhelms-Universität Münster, Institut für Ökonomische Bildung (IÖB). <https://www.econstor.eu/handle/10419/39074>
- Gladen, W. (2014). *Performance Measurement*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-05138-9>
- Goldstein, I. L. & Ford, J. K. (2002). *Training in organizations: Needs assessment, development, and evaluation* (4. Aufl.). Wadsworth.
- Gollwitzer, M. & Jäger, R. S. (2009). *Evaluation kompakt*. Beltz.
- Götz, J. (2001). Realloptionsbasierte Bewertung von Investitionen im Profifußball. In J. Sieglöcher & C. Klimmer (Hrsg.), *Unternehmen Profifußball* (S. 103–135).
- Greguras, G. J. & Robie, C. (1998). A new look at within-source interrater reliability of 360-degree feedback ratings. *The Journal of applied psychology*, 83(6), 960–968. <https://doi.org/10.1037/0021-9010.83.6.960>
- Grossman, R. & Salas, E. (2011). The transfer of training: what really matters. *International Journal of Training and Development*, 15(2), 103–120. <https://doi.org/10.1111/j.1468-2419.2011.00373.x>
- Grünig, R., Heckner, F. & Zeus, A. (1996). Methoden zur Identifikation strategischer Erfolgsfaktoren. *Die Unternehmung*, 50(1), 3–12.
- Gulbin, J. P., Oldenziel, K. E., Weissensteiner, J. R. & Gagné, F. (2010). A Look Through the Rear View Mirror: Developmental Experiences and Insights of High Performance Athletes. *Talent Development & Excellence*, 2(2), 149–164.
- Güllich, A. (2014). Selection, de-selection and progression in German football talent promotion. *European journal of sport science*, 14(6), 530–537. <https://doi.org/10.1080/17461391.2013.858371>
- Güllich, A., Anthes, E. & Emrich, E. (2005). Talentförderung im Sportverein. Teil 2: Interventionen zur Talentsuche und Talentförderung. *Leistungssport*(6), 48–55.
- Gülpen, B. (1996). *Evaluation betrieblicher Verhaltenstrainings unter besonderer Berücksichtigung des Nutzens*. Hampp.

- Haenecke, H. (2002). Methodenorientierte Systematisierung der Kritik an der Erfolgsfaktorenforschung. *Journal of Business Economics (JBE)*, 72(2), 165-183.
- Hagen, K. & Wenning, H. (2007). Evaluation des Management-Entwicklungs-Programms der Hamburger Sparkasse. In A. Süßmair & J. Rowold (Hrsg.), *Kosten-Nutzen-Analysen und Human Resources*. (S. 122–137). Beltz.
- Hager, W., Patry, J.-L. & Brezing, H. (Hrsg.). (2000). *Handbuch Evaluation psychologischer Interventionsmaßnahmen*. Huber.
- Haidl, V. (2021). *Qualität der Leistungsdaten der spanischen La Liga des Internetportals ‚Transfermarkt.de‘* [Bachelor Thesis]. Fachhochschule Südwestfalen, Meschede.
- Hamblin, A. C. (1974). Evaluation and Control of Training. *Industrial Training International*, 9(5), 154–156.
- Hanna, D. & Dempster, M. (2017). *Statistik für Psychologen für Dummies* (1. Aufl.). *Lernen leichter gemacht*. Wiley-VCH Verlag GmbH & Co. KGaA.
- Hansmann, K. W. & Kehl, M. (2000). *Studie zum Shareholder-Value in deutschen Unternehmen*.
- Heaton, J. (2008). *Secondary Analysis of Qualitative Data. An Overview*.
<https://doi.org/10.12759/hsr.33.2008.3.33-45>
- Hedderich, J. & Sachs, L. (2018). *Angewandte Statistik: Methodensammlung mit R* (16., überarb. u. erw. Auflage 2018). Springer Berlin Heidelberg.
<https://doi.org/10.1007/978-3-662-56657-2>
- Heermann, P. W. (2022). *Verbandsautonomie im Sport*. Nomos Verlagsgesellschaft mbH & Co. KG. <https://doi.org/10.5771/9783748933298>
- Heibel (30. Januar 2012). Wertsachen oder Muster ohne Wert? – Über die Macht von transfermarkt.de. *Aktives Abseits - Fussball Blog*.
- Heinrich, B. & Klier, M. (2015). Datenqualitätsmetriken für ein ökonomisch orientiertes Qualitätsmanagement. In K. Hildebrand, M. Gebauer, H. Hinrichs & M. Mielke (Hrsg.), *Daten- und Informationsqualität: Auf dem Weg zur Information Excellence*. (3. Aufl., S. 49–67). Springer Vieweg.
- Helfferich, C. (2011). *Die Qualität qualitativer Daten: Manual für die Durchführung qualitativer Interviews* (4. Aufl.). *Lehrbuch*. VS-Verl. <https://doi.org/10.1007/978-3-531-92076-4>
- Helfrich, H. (2016). *Wissenschaftstheorie für Betriebswirtschaftler* (1. Aufl. 2016). Springer Gabler.
- Herm, S., Callsen-Bracker, H.-M. & Kreis, H. (2014). When the crowd evaluates soccer players' market values: Accuracy and evaluation attributes of an online community. *Sport Management Review*, 17(4), 484–492.
<https://doi.org/10.1016/j.smr.2013.12.006>
- Herr, C. (2007). *Nicht-lineare Wirkungsbeziehungen von Erfolgsfaktoren der Unternehmensgründung*. Zugl.: Duisburg, Essen, Univ., Diss., 2006 (1. Aufl.). *Entrepreneurship* [XX, 349 S]. Springer-Verlag.
- Herterich, H. (2016). *"Geld schießt Tore" als Basis für sportlichen Erfolg in der Fußball-Bundesliga? eine Langzeitanalyse* [Bachelorarbeit, Hochschule Mittweida, University of

- Applied Sciences, Mittweida]. RIS. https://monami.hs-mittweida.de/files/7423/ba_hannes+herterich.pdf
- Heuer, A. (2012). *Der perfekte Tipp: Statistik des Fussballspiels: Erlebnis Wissenschaft*. (1. Aufl.). Wiley-VCH.
- Hildebrand, K., Gebauer, M., Hinrichs, H [Holger] & Mielke, M. (Hrsg.). (2015). *Daten- und Informationsqualität: Auf dem Weg zur Information Excellence*. (3., erweiterte Auflage). Springer Vieweg.
- Hinrichs, H [H.]. (2002). *Datenqualitätsmanagement in Data Warehouse-systemen*. <http://oops.uni-oldenburg.de/id/eprint/279>
- Hinz, A. (2000). *Psychologie der Zeit*. Waxmann Verlag.
- Hochholdinger, S., Rowold, J. & Schaper, N. (2008). Ansätze zur Trainings- und Transferevaluation. In J. Rowold, S. Hochholdinger, S. & N. Schaper (Hrsg.), *Evaluation und Transfersicherung betrieblicher Trainings. Modelle, Methoden und Befunde* (S. 30–49). Hogrefe.
- Hoecke van, J., Schoukens, H. & Knop de, P. (2006). Foot PASS: a constructive and distinctive Quality System for Youth Academies of Professional Fottball Clubs. In Papadimitriou (Hrsg.), *Proceedings of the 14th Congress of the European Association for Sport Management. Nicosia: EASM* (S. 278–279). EASM.
- Höft, S. (2001). Erfolgsüberprüfung personalpsychologischer Arbeit: (. In H. Schuler (Hrsg.), *Lehrbuch der Personalpsychologie* (S. 617–651). Hogrefe/Verlag für Psychologie.
- Hohmann, A. & Seidel, I. (2003). Scientific aspects of talent development. *International Journal of Physical Education*, 40(1), 9–20. <https://eref.uni-bayreuth.de/3360/>
- Hohmann, A., Wick, D. & Carl, K. (2002). Talent im Sport: Versuch einer Bilanz. In A. Hohmann, D. Wick & K. Carl (Hrsg.), *Schriftenreihe: Bd. 110. Talent im Sport*. Hofmann. <https://d-nb.info/96516182X/04>
- Holton, E. F. (1996). The flawed four-level evaluation model. *Human Resource Development Quarterly*, 7, 5–21.
- Höner, O. (2012). Herausforderungen an die Talentforschung im Fußball. *Deutsche Zeitschrift für Sportmedizin*(09), 268–269. <https://doi.org/10.5960/dzsm.2012.033>
- Hornburg, C. & Giering, A. (1996). Konzeptualisierung und Operationalisierung komplexer Konstrukte. Ein Leitfaden für die Marketingforschung. *Marketing ZFP*, 18(1), 5–24. <https://doi.org/10.15358/0344-1369-1996-1-5>
- INEA. (2008). *Rechtsgutachten zur Vereinbarkeit der „6+5-Regel“ mit europäischem Gemeinschaftsrecht*. https://www.rechthaber.com/wp-content/uploads/2010/08/INEA_Gutachten_zu_6_plus_5_Regel_2008.pdf
- IVW. (2021). *Transfermarkt.de - Das Fußballportal*. Goldfish IT- und Medien-Service - Peter Bratfisch. <http://ausweisung.ivw-online.de/index.php?i=1161&a=o90838&it=1>
- Jaeger, B. (2004). *Humankapital und Unternehmenskultur*. Deutscher Universitätsverlag. <https://doi.org/10.1007/978-3-322-81856-0>
- Joch, W. (2001). *Das sportliche Talent: Talenterkennung - Talentförderung - Talentperspektiven* (4., unveränd. Aufl.). Meyer und Meyer.

- Kabst, R. & Wehner, M. C. (2010). Institutionalisation der Personalentwicklung - Ist der Patient auf dem Weg der Besserung? In M. T. Meifert (Hrsg.), *Strategische Personalentwicklung: Ein Programm in acht Etappen ; [Bonusmaterial im Web]* (2. Aufl., S. 45–60). Springer. https://doi.org/10.1007/978-3-642-04401-4_3
- Kasperek, A. (2020). *Kann man transfermarkt.de trauen? Eine Datenqualitätsprüfung der Fußballstatistiken von transfermarkt.de auf Basis eines Vergleichs mit den Daten polnischer Fußballstatistiken* [Bachelor Thesis]. Fachhochschule Südwestfalen, Meschede.
- Kauffeld, S [S.]. (2010). *Nachhaltige Weiterbildung: Betriebliche Seminare und Trainings entwickeln, Erfolge messen, Transfer sichern*. Springer.
- Kauffeld, S [S.]. (2014). *Arbeits-, Organisations- und Personalpsychologie für Bachelor*. Springer.
- Kauffeld, S [S.]. (2016). *Nachhaltige Personalentwicklung und Weiterbildung*. Springer. <https://doi.org/10.1007/978-3-662-48130-1>
- Kauffeld, S [S.] & Paulsen, H. (2018). *Kompetenzmanagement in Unternehmen: Kompetenzen beschreiben messen entwickeln und nutzen* (1. Aufl.). *Arbeits-, Organisations- und Wirtschaftspsychologie*.
- Kaufman, R. & Keller, J. M. (1994). Levels of evaluation: Beyond Kirkpatrick. *Human Resource Development Quarterly*, 5(4), 371–380. <https://doi.org/10.1002/hrdq.3920050408>
- Kaufman, R., Keller, J., Watkins, R. (1996). What works and what doesn't work: Evaluation beyond Kirkpatrick. *Performance & Instruction*, 35(2), 8–12.
- Kellner, H. J. (2006). *Value of Investment: Neue Evaluierungsmethoden für Personalentwicklung und Bildungscontrolling*. Gabal.
- Késenne, S. (2011). Youth Development and Training after the Bosman Verdict (1995) and the Bernard Case (2010) of the European Court of Justice. *European Sport Management Quarterly*, 11(5), 547–553. <https://doi.org/10.1080/16184742.2011.624111>
- Kirkpatrick, D. L. (1959_2). Techniques for evaluating training programs: Part 2 - learning. *Journal of the American Society of Training Directors*, 13(12), 21–26.
- Kirkpatrick, D. L. (1960_1). Techniques for evaluating training programs: Part 3 - behavior. *Journal of the American Society of Training Directors*, 14(1), 13–18.
- Kirkpatrick, D. L. (1960_2). Techniques for evaluating training programs: Part 4 - results. *Journal of the American Society of Training Directors*, 14(2), 28–32.
- Kirkpatrick, D. L. (1959_1). Techniques for evaluation training programs. *Journal of the American Society of Training Directors*, 13, 3–9.
- Kirkpatrick, D. L. (1967). Evaluation of Training. In R. L. Craig (Hrsg.), *Training and development handbook: A guide to human resources development*. (18.1 - 18.27). McGraw-Hill.
- Kirkpatrick, D. L. (1998). *Evaluating training programs: The four levels*. (2. Aufl.). Berrett-Koehler.
- Kirkpatrick, D. L. & Kirkpatrick, J. D. (2005). *Transferring learning to behavior: Using the four levels to improve performance*. Berrett-Koehler Publishers.

- Kirkpatrick, D. L. & Kirkpatrick, J. D. (2006). *Evaluating Training Programs: The Four Levels*. (3. Aufl.). Berrett-Koehler Publishers.
- Kirkpatrick, J. D. & Kirkpatrick W. (2016). *Four Level's of Training Evaluation*. ATD Press.
- Klier, M. (2008). Metriken zur Bewertung der Datenqualität – Konzeption und praktischer Nutzen. *Informatik-Spektrum*, 31(3), 223–236.
<https://doi.org/10.1007/s00287-007-0206-0>
- Kliesch, J. (2017). *Der Status des Profifußballers im europäischen Recht*. Dissertation (1. Aufl.). *Schriften zum Sportrecht: Band 42* [360 Seiten].
- Klimmer, I. (2003). *Profifußballunternehmen an der Börse: Analyse des Wirkungszusammenhangs zwischen sportlichem und wirtschaftlichem Erfolg im Berufsfußball*. Zugl.: Bayreuth, Univ., Dipl.-Arb., 2003 (1. Aufl.). *Sportökonomie Uni Bayreuth - Edition Wissenschaft und Praxis*. Sportökonomie Uni Bayreuth.
- Knop, P. de [P.], van Hoecke, J. & Bosscher, V. de [V.] (2004). Quality Management in Sports Clubs. *Sport Management Review*, 7(1), 57–77.
[https://doi.org/10.1016/S1441-3523\(04\)70045-5](https://doi.org/10.1016/S1441-3523(04)70045-5)
- Koo, T. K. & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Koontz, H., O'Donnell, C. & Wehrich, H. (1984). *Management* (8. Aufl.). *McGraw-Hill series in management*. McGraw-Hill.
- Kraiger, K. (2002). Decision-based evaluation. In K. Kraiger (Hrsg.), *Creating, implementing, and managing effective training and development. Stat-of-the-art lessons for practice* (S. 331–375). CA: Jossey-Bass.
- Kraiger, K. (2003). Perspectives on training and development. In W. C. Borman, D. R. Ilgen & R. J. Klimoski (Hrsg.), *Handbook of psychology: Industrial and organizational psychology* (12. Aufl., S. 171–192). John Wiley & Sons Inc.
- Kraiger, K., Ford, J. K. & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *The Journal of applied psychology*(78), 311–328.
- Kraiger, K. & Jung, K. M. (1997). Linking training objectives to evaluation criteria. In M. A. Quinones & A. Dutta (Hrsg.), *Training for the 21st century technology: Applications of psychological research*. (S. 151–175). American Psychological Association.
<https://doi.org/10.1037/10260-006>
- Krebs, D. & Menold, N. (2014). Gütekriterien quantitativer Sozialforschung. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 425–438). Springer VS. https://doi.org/10.1007/978-3-531-18939-0_30
- Kroemer, U. (2015). „Wir arbeiten zusammen, nicht gegeneinander!“ . *fussballtraining*(08), 2–5. https://www.dfb.de/fileadmin/_dfbdam/78258-fussballtraining_double_pass.pdf
- Kromrey, H. (2001). Evaluation - ein vielschichtiges Konzept: Begriff und Methodik von Evaluierung und Evaluationsforschung ; Empfehlungen für die Praxis.: Sozialwissenschaften und Berufspraxis, 24(2), 105–131.

- Kromrey, H. (2002). *Empirische Sozialforschung* (10. Aufl.). Leske + Budrich.
- Kühl, S. (2006). *Welche Daseinsberechtigung hat Personalarbeit? Ein Thesenpapier mit Begründungen*. Bielefeld. <https://www.uni-bielefeld.de/soz/personen/kuehl/pdf/Paper6.2006-Welche-Daseinsberechtigung-hat-Personalarbeit.pdf>
- Lang, A. & Otto, B. (2010). Erreichen vorbereitete Verhandlungsführer mehr? Überprüfung eines Trainings zur Planung von Verhandlungen. *Zeitschrift für Arbeits- u. Organisationspsychologie*, 54(2), 92–102.
- Lange, B. (1982). Bestimmung strategischer Erfolgsfaktoren und Grenzen ihrer empirischen Fundierung: Dargestellt am Beispiel der PIMS-Studie. *Die Unternehmung*, 36(1), 27–41. <http://www.jstor.org/stable/24178509>
- Lanwehr, R., Honsel, M. & Wilms, R. (2021). Die Evaluation von Qualität – Ein Methodenvergleich anhand von Nachwuchsleistungszentren im deutschen Profifußball. *Gruppe. Interaktion. Organisation. Zeitschrift für Angewandte Organisationspsychologie (GIO)*, 52(1), 25–35. <https://doi.org/10.1007/s11612-021-00556-y>
- Laupus, R. (2019). *Einfluss der Kommerzialisierung auf den Nachwuchsbereich im Fußball: Herausforderungen und Perspektiven (190227)* [Bachelorthesis, Hochschule Mittweida, University of Applied Sciences, Mittweida]. RIS. https://monami.hs-mittweida.de/files/11755/190227_bachelorthesis_robin_laupus.pdf
- Leicht, U. (2022). Erfolgsfaktoren des Supply-Chain Risikomanagements in der Krise. Eine empirische Analyse kleiner und mittlerer B2C-Unternehmen in Deutschland am Beispiel der Coronapandemie. Springer, Neu-Isenburg. <https://doi.org/10.1007/978-3-658-39505-6>.
- Lidor, R., Côté, J., Arnon, M., Zeev, A. & Cohen-Maoz, S. (2010). Relative age and birthplace effects in division 1 players—Do they exist in a small country: Vol. 2, No. 2, 2010, 181–192. *Talent Development & Excellence*, 2(2), 181–192. <https://d-nb.info/1008285544/34#page=67>
- Littkemann, J [J.], Fietz, A. & Krechel, S. (2012). Instrumente zum Controlling von Spielerinvestitionen im Profifußball. In G. Schewe & J. Littkemann (Hrsg.), *Sportmanagement: Der Profifußball aus sportökonomischer Perspektive* (3. Aufl., S. 165–186). Hofmann.
- Littkemann, J [Jörn] & Kleist, S. (2002). Sportlicher Erfolg in der Fußball-Bundesliga: eine Frage der Auf- oder der Einstellung? In H. Albach & B. Frick (Hrsg.), *Sportökonomie* (S. 181–202). Gabler Verlag, Wiesbaden. https://doi.org/10.1007/978-3-663-07711-4_10
- Lorenz, J., Rauhut, H., Schweitzer, F. & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, 108(22), 9020–9025. <https://doi.org/10.1073/pnas.1008636108>
- Martinez-Arastey, G. (2018). *Opta Sports. The leading sports data provider*. <https://www.sportperformanceanalysis.com/article/opta-leading-sport-data-provider>

- Mayntz, R. (1980). Einleitung: Die Entwicklung des analytischen Paradigmas der Implementationsforschung. In R. Mayntz (Hrsg.), *Implementation politischer Programme: Empirische Forschungsberichte* (S. 1–19). Verlagsgruppe Athenäum, Hain, Scriptor, Hanstein.
- McGovern, J., Lindemann, M., Vergara, M., Murphy, S., Barker, L. & Warrenfeltz, R. (2001). Maximising the impact of executive coaching: Behavioral change, organizational outcomes, and Return on Investment. *Manchester Review*, 6(1), 1–9.
- McKinsey & Company. (2020, 24. September). *Profifußball wächst in Deutschland – aber Kadermanagement muss professioneller werden* [Press release]. <https://www.mckinsey.de/news/presse/2020-09-21-bundesliga>
- McKnight, P. & McKnight, K. (2007). *Missing Data: A Gentle Introduction*. Guilford Publications.
- Media Impact. (2019). *Transfermarkt.de - Transferphase*. https://www.mediaimpact.de/data/uploads/2019/11/191105_transferphasetransfermarkt.de-1.pdf
- Media Impact. (2021). *Sportvermarktung: Transfermarkt.de*. <https://www.mediaimpact.de/de/portfolio/transfermarkt-de>
- Meffert, H. (1994). Das PIMS-Programm. In H. Meffert (Hrsg.), *Marketing-Management* (S. 57–77). Gabler Verlag. https://doi.org/10.1007/978-3-322-94537-2_5
- Meuse, K. P. de, Dai, G. & Lee, R. J. (2009). Evaluating the effectiveness of executive coaching: beyond ROI? Meuse, Kenneth P. de; Dai, Guangrong; Lee, Robert J. (2009): Evaluating the effectiveness of executive coaching: beyond ROI? In: *Coaching: An International Journal of Theory, Research and Practice* 2 (2), S. 117–134. DOI: *Coaching: An International Journal of Theory, Research and Practice*, 2(2), 117–134. <https://doi.org/10.1080/17521880902882413>
- Meylan, C., Cronin, J., Oliver, J. & Hughes, M. (2010). Talent Identification in Soccer: The Role of Maturity Status on Physical, Physiological and Technical Characteristics. *International Journal of Sports Science & Coaching*, 5(4), 571–592. <https://doi.org/10.1260/1747-9541.5.4.571>
- Mittag, W. & Hager, W. (2000). Ein Rahmenkonzept zur Evaluation psychologischer Interventionsmaßnahmen. In W. Hager, J.-L. Patry & H. Brezing (Hrsg.), *Handbuch Evaluation psychologischer Interventionsmaßnahmen*. (S. 102–128). Huber.
- Mittag, W. & Hasselhorn, M. (2000). Psychologische Interventionsmaßnahmen: Was sollen sie bewirken können? In W. Hager, J.-L. Patry & H. Brezing (Hrsg.), *Handbuch Evaluation psychologischer Interventionsmaßnahmen*. (S. 41–85). Huber.
- Mittag, W. & Jerusalem, M. (1997). Evaluation von Präventionsprogrammen. In R. Schwarzer (Hrsg.), *Gesundheitspsychologie. Ein Lehrbuch*. (S. 595–611). Hogrefe.
- Mochmann, E. (2014). Quantitative Daten in der Sekundäranalyse. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 233–244). Springer VS.
- Morgan, R. B. & Casper, W. J. (2000). Examining the factor structure of participant reactions to training: A multidimensional approach. *Human Resource Development Quarterly*, 11(3), 301–317. [https://doi.org/10.1002/1532-1096\(200023\)11:3<301::AID-HRDQ7>3.0.CO;2-P](https://doi.org/10.1002/1532-1096(200023)11:3<301::AID-HRDQ7>3.0.CO;2-P)

- Müller, J. (2017). Kommunikative Begleitung als Erfolgsfaktor in der Nutzendarstellung von Investitionen in die Personalentwicklung. *Corporate Communications Journal*, 2(3), 12–21.
- Müller, O., Simons, A. & Weinmann, M. (2017). Beyond crowd judgments: Data-driven estimation of market value in association football. *European Journal of Operational Research*, 263(2), 611–624. <https://doi.org/10.1016/j.ejor.2017.05.005>
- Müller-Hagedorn, L., Toporowski, W. & Zielke, S. (2012). *Der Handel: Grundlagen - Management - Strategien* (2. Auflage). Kohlhammer Verlag.
- Müller-Römer, S. (2006). *Ausbildungsentschädigung für junge Fussballer - § 23 a der DFB Spielordnung im Aus?* https://web.archive.org/web/20111007101301/http://www.anwaltzentrale.de/rechtswalt_fachartikel/fachartikel_detail.php?id=97&Fachgebiet_id=140
- Murr, D. (2019). *Prognoserelevanz personbezogener Talentprädiktoren im Fußball - Eine systematische Analyse zur empirischen Evidenz unter Berücksichtigung von Merkmalen des Studiendesigns*. <https://doi.org/10.15496/PUBLIKATION-26360>
- Neuberger, O. (1994). *Personalentwicklung: Basistexte Personalwesen*. 2. durchges. Aufl. Enke.
- Nicolai, A. & Kieser, A. (2002). Trotz eklatanter Erfolglosigkeit: Die Erfolgsfaktorenforschung weiter auf Erfolgskurs. *DBW - Die Betriebswirtschaft*, 62(6), 579–596. <http://dialog-erfolgsfaktorenforschung.de/ErfolgsfaktorenforschungNicolai-KieserDBW2002.pdf>
- Nork, M. E. (1991). *Management-Training: Evaluation - Probleme - Lösungsansätze* (2. Aufl.). *Hochschulschriften zum Personalwesen: Bd. 9*. Hampp.
- Optasports. (2020). *Opta Daten Feed Übersicht - Optasports*. <https://optasports.de/produkte-medien-daten-feeds-opta-daten-feed-ubersicht.aspx/>
- Otto, B. & Österle, H. (2016). *Corporate Data Quality: Voraussetzung erfolgreicher Geschäftsmodelle*. Springer Gabler.
- Pabst. (o.J.). *Wie die Leistungszentren kontinuierlich verbessert werden*. https://assets.dfb.de/uploads/000/221/315/original_Artikel_IntraNet_-_Wie_die_Leistungszentren_kontinuierlich_verbessert_werden.MPA.pdf?1586357322
- Parrish, R. & Miettinen, S. (2008). *The Sporting Exception in European Union Law*. T.M.C. Asser Press.
- Patry, J.-L. & Perrez, M. (2000). Theorie-Praxis-Probleme und die Evaluation von Interventionsprogrammen. In W. Hager, J.-L. Patry & H. Brezing (Hrsg.), *Handbuch Evaluation psychologischer Interventionsmaßnahmen*. (S. 19–40). Huber.
- Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text*. Sage.
- Peeters, T. (2018). Testing the Wisdom of Crowds in the field: Transfermarkt valuations and international soccer results. *International Journal of Forecasting*, 34(1), 17–29. <https://doi.org/10.1016/j.ijforecast.2017.08.002>
- Perkhofer, S. (2016). Grundlagen und anwendungsbezogene Wissenschaft in den Gesundheitsberufen. In V. Ritschl, R. Weigl & T. Stamm (Hrsg.), *Wissenschaftliches Arbeiten und Schreiben: Verstehen, Anwenden, Nutzen für die Praxis* (S. 3–7). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-49908-5_1

- Peters, T. J. & Waterman, R. H. (1982). *In search of excellence: Lessons from America's best-run companies / by Thomas J. Peters and Robert H. Waterman, Jr.* Harper and Row.
- Pfeil, P. & Müller, M. (2020). Die einzelnen Schritte eines Evaluationsprozesses. In P. Pfeil & M. Müller (Hrsg.), *Lehrbuch. Evaluation in der Sozialwirtschaft: Eine Einführung* (S. 27–92). Springer VS. https://doi.org/10.1007/978-3-658-26322-5_2
- Pflüger, J., Pongratz, H. J. & Trinczek, R. (2017). Fallstudien in der Organisationsforschung. In S. Liebig, W. Matiaske & S. Rosenbohm (Hrsg.), *Handbuch Empirische Organisationsforschung* (S. 389–413). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-08493-6_19
- Phillips, J. J. (1996). ROI: The search for best practices. *Training & Development*, 50(2), 42–47.
- Phillips, J. J. (1998). The return-on-investment (ROI) process: Issues and trends. *Educational Technology*, 38(4), 7–14.
- Phillips, J. J. (1999). *HRD trends worldwide: Shared solutions to compete in a global economy*. Butterworth-Heinemann.
- Phillips, J. J. (2003). *Return on investment in training and performance improvement programs* (2. Aufl.). Butterworth-Heinemann.
- Phillips, J. J. & Phillips, P. P. (1994-2001). *Measuring return on investment. In action*. American Society for Training and Development.
- Phillips, J. J. & Schirmer, F. C. (2008). *Return on Investment in der Personalentwicklung: Der 5-Stufen-Evaluationsprozess* (2. Aufl.). Springer.
- Prockl, F. & Frick, B. (2018). Information precision in online communities: player valuations on www.transfermarkt.de. *International Journal of Sport Finance*, 13(4), 319–335.
- Rae, L. (2002). *Assessing the value of your training: The evaluation process from training needs to the report to the board / Leslie Rae*. Gower.
- Rässler, S. (2000). Ergänzung fehlender Daten in Umfragen / Imputation of Missing Data in Surveys. *Jahrbücher für Nationalökonomie und Statistik*, 220(1), 64–94. <https://doi.org/10.1515/jbnst-2000-0106>
- RatSWD. (2014). *Anwendung und Bewertung von Messinstrumenten in der sozialwissenschaftlichen Umfrageforschung. Rat für Sozial und Wirtschaftsdaten - Working Paper Series: Bd. 230*. Scivero. https://www.konsortswd.de/wp-content/uploads/RatSWD_WP_230.pdf
- Regnier, G., Salmela, J. & Russell, S. J. (1993). Talent Detection and Development in Sport. In R. N. Singer, M. Murphey & L. K. Tennant (Hrsg.), *Handbook of Research on Sport Psychology*. (S. 290–313). MacMillan.
- Rehm-Engel, H. (28. Januar 2019). Nachwuchsarbeit in der Bundesliga: Ein Geschäftsmodell? SPONSORS. <https://www.sponsors.de/news/themen/nachwuchsarbeit-der-bundesliga-ein-geschaeftsmodell>
- Reilly, T., Williams, A. M., Nevill, A. & Franks, A. (2000). A multidisciplinary approach to talent identification in soccer. *Journal of Sports Sciences*, 18(9), 695–702. <https://doi.org/10.1080/02640410050120078>

- Reindl, C. & Krügl, S. (2017). *People Analytics in der Praxis - inkl. Arbeitshilfen online: Mit Datenanalyse zu besseren Entscheidungen im Personalmanagement* (1. Aufl.). Haufe Fachbuch. Haufe Lexware.
- Reio, T. G., Rocco, T. S., Smith, D. H. & Chang, E. (2017). A Critique of Kirkpatrick's Evaluation Model. *New Horizons in Adult Education and Human Resource Development*, 29(2), 35–53. <https://doi.org/10.1002/nha3.20178>
- Röderstein, R [René]. (2009). Erfolgsfaktorenforschung. In R. Röderstein (Hrsg.), *Gabler Research. Erfolgsfaktoren im Supply-Chain-Management der DIY-Branche*. Zugl.: Siegen, Univ., Diss., 2009 (1. Aufl., S. 33–46). Gabler. https://doi.org/10.1007/978-3-8349-8456-2_3
- Roediger, T. (2011). Methoden zur Messung des Akquisitionserfolgs. In T. Roediger (Hrsg.), *Gabler Research. Werte schaffen durch M&A-Transaktionen*. Zugl.: Oestrich-Winkel, Europ. Business School, Diss., 2010 (1. Aufl., S. 109–119). Gabler. https://doi.org/10.1007/978-3-8349-6318-5_5
- Rosar, U., Hagenah, J. & Klein, M. (2013). Sind Sportreporter die objektiveren Beobachter? *Sportwissenschaft*, 43(3), 206–221. <https://doi.org/10.1007/s12662-013-0306-8>
- Rossi, P. H., Freeman, H. E. & Lipsey, M. W. (2003). *Evaluation. A systematic approach*. (7. Aufl.). Sage.
- Rossi, P.H., Freeman, H.E. & Hofmann, G. (1988). *Programm-Evaluation: Einführung in die Methoden angewandter Sozialforschung*. Enke.
- Rost, J. (2000). Allgemeine Standards für die Evaluationsforschung. In W. Hager, J.-L. Patry & H. Brezing (Hrsg.), *Handbuch Evaluation psychologischer Interventionsmaßnahmen*. (S. 129–140). Huber.
- Röttger, S. (2013). *Nachwuchsförderung im Wandel der Zeit: Wie Profifußballvereine sich für eine erfolgreiche Zukunft rüsten*. Diplomica-Verl.
- Rowold, J. (2008). Evaluation eines unternehmensweiten Trainings zur Kundenorientierung in Call Centern auf den vier Ebenen nach Kirkpatrick. In J. Rowold, S. Hochholdinger, S. & N. Schaper (Hrsg.), *Evaluation und Transfersicherung betrieblicher Trainings. Modelle, Methoden und Befunde* (S. 90–99). Hogrefe.
- Rowold, J. & Mönninghoff, M. (2005). Strategic human resource utility analysis. *Journal of Human Resource Costing & Accounting*, 9(1), 10–25. <https://doi.org/10.1108/14013380510636676>
- Saborowski, Y. & Muellerbuchhof, R. (2010). Selbstmanagement-Training als Methode der Kompetenzentwicklung bei Berufseinsteigern - am Beispiel von Auszubildenden technischer Fachrichtungen. *Zeitschrift für Arbeits- u. Organisationspsychologie*, 54(2), 83–91.
- Salas, E. & Cannon-Bowers, J. A. (2001). The science of training: a decade of progress. *Annual review of psychology*, 52, 471–499. <https://doi.org/10.1146/annurev.psych.52.1.471>

- Salas, E., Cannon-Bowers, J. A. & Kozlowski, S. W. J. (1997). The science and practise of training: Current trends and emerging themes. In J. K. Ford, S. W. J. Kozlowski, K. Kraiger, M. Salas & M. Teachout (Hrsg.), *Improving training effectiveness in work organizations* (S. 357–368). Lawrence.
- Sanders, J. R. (2000). *Handbuch der Evaluationsstandards: Die Standards des „Joint Committee on Standards for Educational Evaluation“* (2. durchgesehene Auflage). VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-322-93234-1>
- Sattler, C. & Sonntag, K. (2016). Evaluation Güte und Qualität personaler Führung sichern. In K. Sonntag (Hrsg.), *Personalentwicklung in Organisationen* (S. 603–627). Hogrefe.
- Schermann, M., Hemsen, H., Buchmüller, C., Bitter, T., Krcmar, H., Markl, V. & Horenen, T. (2014). Big Data. *WIRTSCHAFTSINFORMATIK*, 56(5), 281–287. <https://doi.org/10.1007/s11576-014-0434-2>
- Schermuly, C. C., Schröder, T., Nachtwei, J., Kauffeld, S [Simone] & Gläs, K. (2012). Die Zukunft der Personalentwicklung. *Zeitschrift für Arbeits- und Organisationspsychologie A&O*, 56(3), 111–122. <https://doi.org/10.1026/0932-4089/a000078>
- Schewe, G [G.] & Littkemann, J [J.]. (1999). Einführung innovativer Spielsysteme im Profifußball: Die Übertragung eines betriebswirtschaftlichen Organisationskonzepts. In H.-D. Horch, J. Heydel & A. Sierau (Hrsg.), *Professionalisierung im Sportmanagement*, (S. 325–336).
- Schiffler, A. & Hübner, S. (2000). Allgemeine Standards für die Evaluationspraxis. Die Standards des „Joint Committee on Standards for Educational Evaluation“ und ihre Anwendung auf praktische Aspekte bei der Evaluation von psychologischen Interventionsmaßnahmen. In W. Hager, J.-L. Patry & H. Brezing (Hrsg.), *Handbuch Evaluation psychologischer Interventionsmaßnahmen*. (S. 141–152). Huber.
- Schmalen, C., Kunert, M. & Weindlmaier, H. (2006). Erfolgsfaktorenforschung: theoretische Grundlagen, methodische Vorgehensweise und Anwendungserfahrungen in Projekten für die Ernährungsindustrie. *Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V.*(41), 351–362. <https://ageconse-arch.umn.edu/record/259272/files/bd41nr30.pdf>
- Schmid, L. (2012). *Erfolgssynamik: Ein Konzept für die Erstellung systemdynamischer Modelle zur Entscheidungsunterstützung im strategischen Management*. Books on Demand.
- Schneider, A. (2011). Professionelle Wirkung zwischen Standardisierung und Fallverstehen: Zum Stand der Wirkungsforschung. In N. Eppler, I. Miethe & A. Schneider (Hrsg.), *Qualitative und quantitative Wirkungsforschung*. Budrich.
- Schnell, R., Hill, P. B. & Esser, E. (2011). *Methoden der empirischen Sozialforschung* (9., aktualisierte Aufl.). Oldenbourg.
- Schoenberg, R. (2006). Measuring the Performance of Corporate Acquisitions: An Empirical Comparison of Alternative Metrics. *British Journal of Management*, 17(4), 361–370. <https://doi.org/10.1111/j.1467-8551.2006.00488.x>
- Scholz, C., Stein, V. & Bechtel, R. (2004). *Human Capital Management: Wege aus der Unverbindlichkeit*. Wolters Kluwer.

- Scholz, C., Stein, V. & Bechtel, R. (2011). *Human Capital Management: Raus aus der Unverbindlichkeit!* Luchterhand.
- Schroeder, W. (2016). Der Spielertransfer in der EU – 20 Jahre nach dem Bosman-Urteil. In G. Schulze (Hrsg.), *Aktuelle Rechtsfragen im Profifußball: Psychologische Faktoren und rechtliche Gestaltung* (1. Aufl., S. 134–151). Nomos Verlagsgesellschaft mbH & Co. KG. <https://doi.org/10.5771/9783845264257-134>
- Schupp, J. & Wolf, C. (Hrsg.). (2015). *Schriftenreihe der ASI - Arbeitsgemeinschaft Sozialwissenschaftlicher Institute. Nonresponse Bias: Qualitätssicherung sozialwissenschaftlicher Umfragen*. Springer VS.
- Schwaninger, M. & Kaiser, C. (2007). Erfolgsfaktoren organisatorischen Wandels. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, 59(2), 150–172. <https://doi.org/10.1007/BF03371691>
- Schweitzer, M. & Küpper, H. U. (2013). *Produktions- und Kostentheorie: Grundlagen — Anwendungen*. Gabler Verlag. <https://books.google.de/books?id=Gqv2BQAAQBAJ>
- Scriven, M. (2007). *The Logic and Methodology of Checklists*. https://wmich.edu/sites/default/files/attachments/u350/2014/logic%26methodology_dec07.pdf
- Seeber, S. (2000). Stand und Perspektiven des Bildungscontrollings. In S. Seeber, E. M. Krekel & J. van & Buer (Hrsg.), *Bildungscontrolling. Ansätze und kritische Diskussionen zur Effizienzsteigerung von Bildungsarbeit* (S. 19–50). Lang.
- Seidel, I. (2011). Trends in der Talentforschung und Talentförderung, 19–23. https://www.researchgate.net/profile/ilka-seidel/publication/281816258_trends_in_der_talentforschung_und_talentforderung_leistungssport_41_2_19-23/links/5601097308aeafc8ac8c7e79/trends-in-der-talentforschung-und-talentfoerderung-leistungssport-41-2-19-23.pdf
- Setaro, J. (2001). *Many happy returns: Calculating e-learning ROI*. <http://www.learningcircuits.org/2001/jun2001/Elearn.ht>
- Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428. <https://doi.org/10.1037//0033-2909.86.2.420>
- Smokvina, V. (2013). The UEFA “Home-grown Players” Rule: Does It Fulfil Its Aim? *The International Journal of Sport and Society*, 3(2), 67–80. <https://doi.org/10.18848/2152-7857/CGP/v03i02/53904>
- Söhnen, S. (2019). *Einstieg in die Intraklassenkorrelation – eine einfache Anleitung für die Auswahl, Berechnung und Interpretation der Intraklassenkorrelation mit IBM SPSS Statistics*. [https://statistik-support-söhnen.de/2019/03/08/einstieg-in-die-intraklassenkorrelation-eine-einfache-anleitung-fuer-die-auswahl-berechnung-und-interpretation-der-intraklassenkorrelation-mit-ibm-spss-statistics/#:~:text=Zusammengesetzt%20hei%C3%9Ft%20das%3A%20Die%20Intraklassenkorrelation,eines%20Subjekt%20\(z.B.%20eines%20Kindes](https://statistik-support-söhnen.de/2019/03/08/einstieg-in-die-intraklassenkorrelation-eine-einfache-anleitung-fuer-die-auswahl-berechnung-und-interpretation-der-intraklassenkorrelation-mit-ibm-spss-statistics/#:~:text=Zusammengesetzt%20hei%C3%9Ft%20das%3A%20Die%20Intraklassenkorrelation,eines%20Subjekt%20(z.B.%20eines%20Kindes)
- Solga, M. (2011). Evaluation der Personalentwicklung. In Ryschka, J., Solga, M. & Mattenklott, A. (Hrsg.), *Praxishandbuch Personalentwicklung* (S. 369–400). Gabler.

- Solga, M., Ryschka, J. & Mattenklott, A. (2011). Personalentwicklung: Gegenstand, Prozessmodell, Erfolgsfaktoren. In Ryschka, J., Solga, M. & Mattenklott, A. (Hrsg.), *Praxishandbuch Personalentwicklung* (S. 19–34). Gabler.
- Sonntag, K. & Stegmaier, R. (2016). HR-Praktiken und Unternehmenserfolg - eine Evaluationsperspektive. In K. Sonntag (Hrsg.), *Personalentwicklung in Organisationen* (S. 663–697). Hogrefe.
- Sontag, B. (2011). *Strategische Erfolgsfaktoren professioneller Sportorganisationen*. Springer Gabler.
- Spiel, C. (2003). Wissenschaftliche Evaluation - eine Maßnahme der Qualitätssicherung. *Erziehung und Unterricht*, 153(7-8), 883-888.
- Spitzer, D. & Conway, M. (2002). *Link training to your bottom-line*. Infoline. American Society for Training and Development.
- sportschau.de. (2020). *Fußball: Nur die wenigsten Talente schaffen den Sprung in den Profifußball - Fußball*. <https://www.sportschau.de/fussball/talente-text-allgemein-100.html>
- Stangl, W. (2021). *summative Evaluation – Online Lexikon für Psychologie und Pädagogik (stangl.eu)*. <https://lexikon.stangl.eu/12982/summative-evaluation>
- Statista. (2021a). *Ausländeranteil der 1. Fußball-Bundesliga bis 2021*. <https://de-statista.com.ezproxy.fhb.fh-swf.de/statistik/daten/studie/1040656/umfrage/auslaenderanteil-der-1-fussball-bundesliga/>
- Statista. (2021b). *Ausländeranteil der großen Fußball-Ligen Europas | Statista*. <https://de.statista.com/statistik/daten/studie/163550/umfrage/auslaenderanteil-der-top-fussball-ligen/>
- Stein, P. (2014). Forschungsdesigns für die quantitative Sozialforschung. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 135–151). Springer VS.
- Stein, V. (2006, 3. April). *Siegener Manifest zur Saarbrücker Formel*. Siegen. Universität Siegen. https://www.wiwi.uni-siegen.de/pmg/veroeffentlichungen/dokumente/stein_hcm_siegener_manifest.pdf
- Stockmann, R. & Meyer, W. (2014). *Evaluation: Eine Einführung* (2. Aufl.). Budrich.
- Strack, R. & Baier, J. (2009). Talent- und Risikomanagement in unruhigen Zeiten. In W. Jäger (Hrsg.), *Talent Management. Strategien. Umsetzung, Perspektiven (Personalwirtschaft Buch)* (S. 25–32). Luchterhand.
- Strauß, S. (2020). *Analyse von Datenbanken auf Fehlerhaftigkeit am Beispiel der Daten der Englischen Liga und Nationalmannschaft von Transfermarkt.de* [Unveröffentlichte Studienarbeit]. Fachhochschule Südwestfalen, Meschede.
- Strunk, K. S. (1999). *Status of and barriers to financial impact evaluations in employer-sponsored training programs. Dissertation Abstract International, A60/06, 148.* (A60/06, 148). University of Arkansas, Arkansas.
- Stufflebeam, D. L. (Hrsg.) (1971). *The Relevance of the CIPP Evaluation Model for Educational Accountability*.
- Stufflebeam, D. L. (2001). The Metaevaluation Imperative. *American Journal of Evaluation*, 22(2), 183–209. <https://doi.org/10.1177/109821400102200204>

- Stufflebeam, D. L. (2004). A Note on the Purposes, Development, and Applicability of the Joint Committee Evaluation Standards. *American Journal of Evaluation*, 25(1), 99–102.
- Stufflebeam, D. L. (2011). Meta-Evaluation: [S.l.], v. 7, n. 15, p. 99-158, nov. 2010. ISSN 1556-8180. *Journal of MultiDisciplinary Evaluation*, 7(15), 99–158. https://journals.sfu.ca/jmde/index.php/jmde_1/article/download/300/295
- Stufflebeam, D. L. & Shinkfield, A. J. (2007). *Evaluation Theory, models and applications*. Wiley.
- Stügelmaier, M. (2016). *Erfolgsquote und Turnoverrate in einem Nachwuchsleistungszentrum: Eine Analyse der Ausbildung und Talentförderung von Nachwuchsspielern in Deutschland*. AV Akademikerverlag.
- Stützer, C. M., Welker, M. & Egger, M. (Hrsg.). (2018). *Computational social science in the age of big data*. Halem, Köln
- Stützer, C. M. (2020). *Innovative Forschungsmethoden in der Evaluation – Text Mining und Data Analytics zur Erfolgsmessung und Wirksamkeitsanalyse*, in: B. Keller et al. (Hrsg.), *Marktforschung für die Smart Data World*, Springer, Wiesbaden. https://doi.org/10.1007/978-3-658-28664-4_12
- Surowiecki, J. (2005). *Die Weisheit der Vielen.: Warum Gruppen klüger sind als Einzelne und wie wir das kollektive Wissen für unser wirtschaftliches, soziales und politisches Handeln nützen können*. (2. Aufl.). Bertelsmann.
- Süßmair, A. (2007). Kosten-Nutzen-Analysen im Bereich Human Resources - eine verhaltensorientierte ökonomische Perspektive. In A. Süßmair & J. Rowold (Hrsg.), *Kosten- Nutzen-Analysen und Human Resources*. (S. 80–93). Beltz.
- Tasliarmut, M. (2022). *Innovative Evaluationsmethoden und -instrumente*. 16.05.2022. Digitale Jahrestagung des Arbeitskreises „Berufliche Bildung“ der DeGEval – Gesellschaft für Evaluation e.V. https://www.degeval.org/fileadmin/users/Arbeitskreise/AK_Berufliche_Bildung/Evaluationsmethoden_und_-instrumente.pdf
- Tausendpfund, M. (2018). *Quantitative Methoden in der Politikwissenschaft*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-20698-7>
- Tirre, S. (2011). *Systemische Organisationsberatung zwischen systemtheoretischen Anspruch und Wirklichkeit in der Beratungspraxis: Unveröffentlichte Seminararbeit*. Freie Universität Berlin.
- Tirre, S. (2012). *Wirksamkeit von Weiterbildung: Evaluation der Ausbildung zum Innovationspromotor* [Masterthesis, Freie Universität Berlin, Berlin]. DataCite.
- Torgler, B., Schmidt, S. L. & Frey, B. S. (2006). Relative Income Position and Performance: An Empirical Panel Analysis. *SSRN Electronic Journal*. Vorab-Onlinepublikation. <https://doi.org/10.2139/ssrn.889328>
- Tornow, H. (2005). Wirkungsorientierte Steuerung der Sozialen Arbeit, 85(Nr. 8), S. 282-287. <https://repository.difu.de/jspui/handle/difu/266425>
- Transfermarkt.de. (2013). *NEU: Leitlinien der MWA - Marktwertanalyse - Forum | Seite 1*. https://www.transfermarkt.de/neu-leitlinien-der-mwa/thread/forum/67/thread_id/233519

- Transfermarkt.de. (2017). *NEU: Marktorientierung und Definition - Marktwertanalyse - Forum | Seite 1*. https://www.transfermarkt.de/neu-marktorientierung-und-definition/thread/forum/67/thread_id/235745
- Transfermarkt.de. (2021a). *Fußball-Transfers, Gerüchte, Marktwerte und News*. <https://www.transfermarkt.de/>
- Transfermarkt.de. (2021b, 28. September). *Datenpflege - "How-to-Guide": FAQ*. <https://www.transfermarkt.de/intern/faq>
- o.A. Transfermarkt.de. (2021, 17. Dezember). *FC Schalke 04 - Wertvollste Abgänge*. <https://www.transfermarkt.de/fc-schalke-04/wertvollsteabgaenge/verein/33>
- UEFA. *Die UEFA*. <https://de.uefa.com/insideuefa/about-uefa/what-uefa-does/>
- UEFA. (2005). *Protection of young players*. UEFA. <https://www.uefa.com/insideuefa/protecting-the-game/protection-young-players/?referrer=%2Fnews%2Fnewsid%3D943393>
- UEFA. (2014). *Finanzielles Fairplay*. <https://de.uefa.com/insideuefa/protecting-the-game/financial-fair-play/>
- UEFA. (2018). *UEFA-Reglement zur Klublizenzierung und zum finanziellen Fairplay Ausgabe 2018*. <https://documents.uefa.com/v/u/x5wm6~rAbUSzY4JhvwousA>
- UEFA (2020a) *Studienergebnisse zu Trainingseinrichtungen und Nachwuchsförderung in Europa veröffentlicht*. <https://de.uefa.com/insideuefa/news/0261-1056fa1e9686-272bde866037-1000--studienergebnisse-zu-trainingseinrichtungen-und-nachwuchsforder/>
- UEFA (2020b) *UEFA Training Facility and Youth Investment Landscape* <https://www.uefa.com/insideuefa/uefatrainingfacilitiesandyouthinvestment-landscape/>
- UEFA. (2021). *Wie kommen die Klubkoeffizienten zustande?* UEFA. <https://de.uefa.com/memberassociations/uefarankings/club/about/>
- van Buren, M. E. & Erskine, W. (2002). *State of the industry: ASTD's annual review of trends in employer-provided training in the United States*. American Society of Training and Development.
- Viemann, K., Filbrich, B. & Tietje, C. (2005). Controlling in der Fußballbranche – Eine empirische Untersuchung zum Implementierungsstand des Controlling in den drei höchsten deutschen Fußballligen / Accounting Practices and Control in the Soccer Industry – An Empirical Study on the State of Accounting Practices and Control Implementation in the Top Three German Soccer Leagues. *Sport und Gesellschaft*, 2(3), 296–313. <https://doi.org/10.1515/sug-2005-0305>
- Vögele, C. (2018). Gerüchteküche Transfermarkt. *Journal für Sportkommunikation und Mediensport*, 3(1), 27–50. <https://doi.org/10.25968/JSKMS.2018.1.27-50>
- Wagner, G. G. (2014). Kapitalmärkte irren immer wieder. *DIW Wochenbericht*, 81(26), 628. <http://hdl.handle.net/10419/98682>
- Waibel, R. & Käppeli, M. (2006). *Betriebswirtschaft für Führungskräfte*. Versus.
- Wallebohr, A. (2019). *Nachwuchsausbildung im Profifußball: eine empirische Analyse von Schweizer Fußballclubs* [, Friedrich-Schiller-Universität Jena]. DataCite.

- Wallebohr, A. & Daumann, F. (2019). Beurteilung des Erfolgs in der Nachwuchsausbildung: Eine empirische Analyse von Schweizer Fußballclubs. *German Journal of Exercise and Sport Research*, 49(1), 65–77. <https://doi.org/10.1007/s12662-018-0557-5>
- Wang, G. G., Dou, Z. & Li, N. (2002). A systems approach to measuring return on investment for HRD interventions. *Human Resource Development Quarterly*, 13(2), 203–224. <https://doi.org/10.1002/hrdq.1024>
- Wang, G. G. & Wilcox, D. (2006). Training evaluation: knowing more than is practiced. *Advances in Developing Human Resources*, 8(4), 528–539.
- Wang, R. Y. & Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12(4), 5–33. <http://www.jstor.org/stable/40398176>
- Warr, P., Allan, C. & Birdi, K. (1999). Predicting three levels of training outcome. *Journal of Occupational and Organizational Psychology*, 72(3), 351–375. <https://doi.org/10.1348/096317999166725>
- Warr, P. & Bunce, D. (1995). Trainee characteristics and the outcomes of open learning. *Personnel Psychology*, 48(2), 347–375. <https://doi.org/10.1111/j.1744-6570.1995.tb01761.x>
- Warr, P. B., Bird, M. & Rackham, N. (1970). *The Evaluation of Management Training*. Gower Press.
- Weiber, R [Rolf] & Mühlhaus, D. (2014). *Strukturgleichungsmodellierung: Eine anwendungsorientierte Einführung in die Kausalanalyse mit Hilfe von AMOS, SmartPLS und SPSS ; [mit Extras im Web]* (2. Aufl.). Springer-Lehrbuch. Springer Gabler.
- Weir, P. L., Smith, K. L., Paterson, C. & Horton, S. (2010). Canadian Women's Ice Hockey -- Evidence of a Relative Age Effect. *Talent Development & Excellence*, 2(2), 209–217. <https://search.ebscohost.com/login.aspx?direct=true&db=bsu&AN=63612979&lang=de&site=ehost-live>
- Weisberg, H. F. (2005). *The total survey error approach: A guide to the new science of survey research*. University of Chicago Press.
- Werner, J. M. & DeSimone, R. L. (2005). *Human resource development* (4. Aufl.). Thomson.
- Wickel-Kirsch, S. (2012). Grundlagen des Human Resource Controlling. In R. Gleich & A. Klein (Hrsg.), *Controlling-Instrumente für das moderne Human Resource Management* (S. 23–42).
- Wickham, H. & Grolemund, G. (2016). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. "O'Reilly Media, Inc."
- Widmer, T [T.]. (2000). Qualität der Evaluation – Wenn Wissenschaft zur praktischen Kunst wird. In R. Stockmann (Hrsg.), *Evaluationsforschung*. (S. 77–102). Leske + Budrich.
- Widmer, T [Thomas]. (1996). *Meta-Evaluation: Kriterien zur Bewertung von Evaluationen*. Zugl.: Zürich, Univ., Diss., 1995 [XVI, 341 S]. Paul Haupt. <https://doi.org/10.5167/UZH-157742>

- Widmer, T. & Rocchi, T. de. (2012). *Evaluation. Grundlagen, Ansätze und Anwendungen*. Rüegger.
- Wirtz, M. & Kutschmann, M. (2007). Analyse der Beurteilerübereinstimmung für kategoriale Daten mittels Cohens Kappa und alternativer Masse [Analyzing interrater agreement for categorical data using Cohen's kappa and alternative coefficients]. *Die Rehabilitation*, 46(6), 370–377. <https://doi.org/10.1055/s-2007-976535>
- Witzel, A. (2010). Längsschnittdesign. In G. Mey & K. Mruck (Hrsg.), *Handbuch qualitative Forschung in der Psychologie* (1. Aufl., S. 290–303). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92052-8_20
- Wottawa, H. & Thierau, H. (2003). *Lehrbuch Evaluation* (3. Aufl.). Huber.
- Würthele, V. G. (2003). *Datenqualitätsmetrik für Informationsprozesse: Datenqualitätsmanagement mittels ganzheitlicher Messung der Datenqualität* [, ETH Zurich]. DataCite.
- Wylleman, P., Alfermann, D. & Lavallee, D. (2004). Career transitions in sport: European perspectives. *Psychology of Sport and Exercise*, 5(1), 7–20. [https://doi.org/10.1016/S1469-0292\(02\)00049-3](https://doi.org/10.1016/S1469-0292(02)00049-3)
- Zülch, H. (2020). Schwache Vereine trotz starker Liga: Wie der deutsche Profifußball nach der Geistersaison krisenfest wird. *manager magazin*. <https://www.managermagazin.de/lifestyle/fitness/fussball-bundesliga-wie-clubs-nach-der-geistersaison-krisenfest-werden-a-1307968.html>
- Zülch, H., Ottenstein, P. & Manz, E. (2020). The German Bundesliga Clubs and Their Future. <https://www.hhl.de/app/uploads/2020/07/Executive-Summary-The-German-Bundesliga-Clubs-and-Their-Future.pdf>

Fußball ist eine beliebte Sportart, die weltweit Millionen von Menschen begeistert. Über Fußball lässt sich vortrefflich diskutieren, über aktuelle Spiele ebenso wie über Taktiken oder Spielertransfers. Die Bedeutung der Fähigkeiten der Spieler im Profifußball ist enorm, ihre Ausbildung für den Profifußball in Deutschland ist ein wichtiger Prozess, der in den Nachwuchsleistungszentren der Vereine stattfindet. Diese Jugendakademien ermöglichen talentierten jungen Spielern, ihre fußballerischen Fähigkeiten zu entwickeln und sich auf eine mögliche Karriere im Profifußball vorzubereiten.

Die Qualität der Ausbildung in einem Nachwuchsleistungszentrum eines Vereins der deutschen Bundesliga wird vom Deutschen Fußball-Bund (DFB), der Deutschen Fußball Liga (DFL), Spielerberatern, Scouts, der Öffentlichkeit und den Medien bisher anhand verschiedener Kriterien bewertet, etwa an den Erfolgen der Nachwuchsmannschaften, der Integration in einen Profikader, an Infrastruktur und Ausstattung. In wirtschaftlich angespannten Zeiten reichen diese Kriterien für die Vereine jedoch nicht, um den Erfolg von Nachwuchsleistungszentren umfassend zu messen.

Die vorliegende Studie befasst sich mit der Messung des Outcomes der Nachwuchsleistungszentren: der hervorgebrachten Spielerqualität bewertet mit Marktpreisen. So wird eine Analyse der Wirksamkeit und Effizienz aus betriebswirtschaftlicher Perspektive erst ermöglicht. Das Ergebnis der Arbeit ist ein Evaluationskonzept für Nachwuchsleistungszentren.